

Non-independence in network-structured Mendelian randomization: a permutation framework and six failure modes

Prayagh Prabhakar

prayagh09@gmail.com

Independent Researcher

Research Article

Keywords: Mendelian randomization, proteomics, colocalization, permutation testing, multiple comparisons, coronary artery disease

Posted Date: October 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11115416/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background.

Protein Mendelian randomization (MR) is now routine, and a common inference treats an estimate built from a ligand's own cis protein quantitative trait locus (pQTL) and one built from its receptor's as evidence about the same pathway. That assumption has not been tested since it means comparing MR estimates across a network of interacting proteins, and network comparisons violate the independence that standard multiple-testing corrections assume.

Methods.

Using cis-pQTLs as instruments, we analyzed 660 of the 661 UK Biobank Pharma Proteomics Project proteins spanning 1,129 curated ligand–receptor pairs, and estimated effects on coronary artery disease in two genome-wide association studies (GWAS): Nikpay 2015 (60,801 cases, no UK Biobank overlap) and Aragam 2022 (181,522 cases, which include the Nikpay data). Discordance between ligand-level and receptor-level estimates was tested against a permutation null holding every protein's estimate and degree fixed and randomizing only the pairing. Estimates were followed up with Bayesian colocalization and with replication on a second proteomic platform.

Results.

Ligand and receptor estimates agreed no more than randomly paired proteins: 48.5% same-direction against a rewired null of 49.8% ($p = 0.64$) in Nikpay, and 47.0% against 50.0% ($p = 0.20$) in Aragam. Counts of significantly discordant pairs rose threefold with outcome power, from 10 to 31, and the null rose with them, from 9 to 30 ($p = 0.85$). Bonferroni correction across 476 deduplicated pairs built from 361 proteins reported ten discoveries that three proteins account for. A node-shuffling null gave $p = 0.014$ because protein degree correlates with effect size in Aragam ($\rho = +0.133$, $p = 0.008$). Flagging protein-altering instruments, a routine sensitivity analysis, flagged PCSK9 and showed no enrichment among causal proteins (Fisher $p \geq 0.44$). Colocalization verdicts survived a change of outcome GWAS at only 11 of 14 loci ($PP.H4 \geq 0.8$), mostly through outcome power; one locus reversed outright.

Conclusions.

Ligand-level and receptor-level MR are not interchangeable evidence about a pathway. More important, the naive analysis that appears to identify specific discordant pairs is reproducible. It returns ten pairs after correction, headed by one at $p = 8.5 \times 10^{-7}$, and we show what removes them.

Background

Nominating drug targets from plasma protein MR [1, 2] is now routine. A single recent study ran 2,940 proteins against 19 cardiovascular diseases [3], and the same design applied to coronary artery disease recurs throughout the recent literature. The summary statistics are public and a genome-wide scan finishes in a day. When a ligand and its receptor both have cis-pQTLs, an estimate from either is read as evidence about the pathway they share [4, 5]. If interleukin-6 raises coronary risk, the reasoning goes, then the interleukin-6 receptor should point the same way, and either instrument licenses a statement about interleukin-6 signaling. At IL6R, the allele that raises circulating interleukin-6 by 9.45% per copy also lowers C-reactive protein, a pattern matching pharmacological receptor blockade [4]. The ligand measurement and the receptor's activity move in opposite directions, so the two instruments answer opposite questions. Whether that case is an abnormality or the norm has not been established, because establishing it requires comparing MR estimates over many pairs at once. Pairs are edges, edges share nodes, and one protein's estimate therefore reappears in every pair it belongs to. The pairs are not independent tests; however, the corrections routinely applied to them assume otherwise. We set out to test the pathway assumption, which required a null that adheres to the network and which exposed many flaws in common practice.

Methods

This study is reported in compliance with the STROBE-MR guidelines [6]. A completed STROBE-MR checklist, giving the manuscript text or page number that supports each item, is provided as Additional file 1. Figure 1 shows the flow of proteins and pairs through the analysis. All analyses ran under Python 3.14.3 with numpy 2.4.6, scipy 1.17.0 and matplotlib 3.10.8, and were run or reproduced under pandas 2.3.3 and 3.0.5. Every stochastic step is seeded: permutation nulls use seed 20260829. The study protocol was not pre-registered. Two predictions, both described in Results, were specified before the analysis was run.

Data sources

This is a two-sample, summary-data Mendelian randomization study. No individual-level data were used. Every source below is published and openly available.

Protein exposures. Olink protein summary statistics come from the sex-combined meta-analysis release of Koprulu et al. 2025 [7], a reanalysis of the UK Biobank Pharma Proteomics Project [8]. It is distributed at https://omicscience.charite.de/Selexspecific_pGWAS/, directory UKBB_MA, one file per assay, named `invn_<OlinkID>_UKBB_proteomics_sexcombined.tsv.gz`, GRCh37. This release has no GWAS Catalog accession. Cross-platform SomaScan instruments come from the same resource's Fenland_MA directory, files `SeqId_<SeqId>_Fenland_proteomics_sexcombined.tsv.gz`, a reanalysis of the original Fenland study [9].

Coronary artery disease outcomes. Nikpay et al. 2015 [10] is GWAS Catalog accession GCST003116, PMID 26343387, <https://www.ebi.ac.uk/gwas/studies/GCST003116>. The file used was the additive-model summary statistics, `cad.add.160614.website.txt`, GRCh37, downloaded from the GWAS Catalog FTP at http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST003001-GCST004000/GCST003116/cad.add.160614.website.txt. Aragam et al. 2022 [11] is GWAS Catalog accession GCST90132314, PMID 36474045. Regions were fetched by tabix from the harmonized file at https://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/GCST90132001-GCST90133000/GCST90132314/harmonised/GCST90132314.h.tsv.gz, GRCh38, and validated back to GRCh37 as described below.

Ligand-receptor pairs. OmniPath [12] was queried at <https://omnipathdb.org/interactions?datasets=ligreextra&genesymbols=yes&organisms=9606&fields=sources,references&format=tsv>. Re-running the identical query later returned the same 6,555 directed gene-symbol pairs.

Constraint metrics. gnomAD v2.1.1 [13], file https://storage.googleapis.com/gcp-public-data-gnomad/release/2.1.1/constraint/gnomad.v2.1.1.lof_metrics.by_gene.txt.bgz, 19,704 rows covering 19,658 unique genes.

Access and permissions. All resources above are openly downloadable, summary-level data. None required an application, a registration, or a data-access committee approval. Each was retrieved anonymously. The CARDIoGRAMplusC4D page requests citation of the original paper, which we provide, but imposes no access condition. No individual-level UK Biobank data were used, so no UK Biobank data-access application was required.

Instrumental variable assumptions

Each protein's effect on coronary artery disease is estimated by the Wald ratio of its single cis instrument, the variant-outcome association divided by the variant-exposure association, with a first-order approximation to the standard error. Variant-exposure and variant-outcome associations are taken as published by each source GWAS, with covariate adjustment as described in those studies. The main analysis rests on the three core instrumental variable assumptions. Relevance, that each instrument associates with its protein, is enforced by the genome-wide significance threshold and the $F \leq 10$ exclusion. The weakest sentinel retained has $F = 30.1$. Independence, that the instrument shares no common cause with the outcome, and the exclusion restriction, that the instrument affects coronary artery disease only through the measured protein, cannot be directly tested from summary statistics. Restricting instruments to a 1 Mb cis window around each protein's own gene reduces the chance that the variant acts through an unrelated pathway. The protein-altering instrument check and the cross-platform replication test whether an instrument's association with the protein reflects true concentration. Colocalization tests whether the same variant plausibly drives both the protein and the disease signal, which bears on exclusion at each locus. None of these analyses can rule out a violation with certainty, and we do not claim that they do.

Exposures and instruments

Protein levels came from the UK Biobank Pharma Proteomics Project [8], measured on the Olink Explore platform (antibody-based proximity extension assay), accessed through the sex-combined meta-analysis release at omicscience.charite.de [7]. Median sample size for the 552 proteins with a cis instrument was 46,859 (range 38,091–47,534). Ligand–receptor pairs were taken from OmniPath [12], restricted to interactions where both partners are assayed on the Olink panel. This gave 1,129 pairs spanning 661 proteins, of which 660 were analyzed; the summary statistics for one protein (CGN) could not be retrieved.

For each protein we retained every genome-wide significant variant ($p < 5 \times 10^{-8}$) within 1 Mb of its own gene and took the lead variant (smallest p-value) as the instrument, the design used by recent single-variant protein MR. Because values below about 10^{-308} cannot be represented as floating-point numbers, and the source prints such values as text (for example, 1.20×10^{-7186}), p-values were compared as printed, and among variants with identical p-values a non-palindromic variant was preferred. Variants with $F \leq 10$ were to be dropped, the conventional, though criticized, weak-instrument threshold [14]. The weakest sentinel retained has $F = 30.1$, because a variant reaching $p < 5 \times 10^{-8}$ in a sample of this size is necessarily strong. The rule does bind on the SomaScan instruments used for cross-platform comparison, which were selected differently (Cross-platform replication). Because each of the 661 protein files is roughly 240 MB and the full set exceeds available disk, files were processed in a download–extract–delete cycle, retaining only cis variants and a compact genome-wide signal summary per protein. We also recorded, for every protein, where its genome-wide significant signal actually sits: how much falls in cis, how much falls inside the extended MHC (chr6:28,477,797–33,448,354, GRCh37), and where the single strongest variant lies. This audit is reported in its own right because it determines which proteins can be studied at all. The instrument for each of the 552 proteins, with its association with the protein and with coronary artery disease in each outcome GWAS, is given in Additional file 2.

Outcomes

Two coronary artery disease GWAS were used. Nikpay 2015 [10] contributes 60,801 cases and includes no UK Biobank cohort, so it shares no participants with the exposures, and it is our primary analysis for the pair-level comparison. The colocalization analyses and Table 1 are computed against Aragam 2022, which has the outcome power they require. Aragam 2022 [11] contributes 181,522 cases and is the higher-powered secondary analysis. However, the two are not independent: Aragam is a CARDIoGRAMplusC4D meta-analysis that incorporates the Nikpay data alongside UK Biobank and other cohorts, so agreement between them is consistency across nested samples. Heterogeneity across the cohorts meta-analyzed within each outcome GWAS was not assessed here. The 181,522 is a headline count, and the realized case count per harmonized instrument in the Aragam file has median 178,731 and minimum 72,859 (460 distinct variants in 461 proteins). Within the 28 colocalization regions the median per-variant case count is 98,456, well below the headline figure. Aragam also includes UK Biobank, whose exact contribution to the 1,165,690 participants is not reported separately. The 33,941 cases and 438,394 controls quoted in that paper [11] are its exome burden test. Taking the exposure sample as entirely contained in it, the overlap fraction is $46,859/1,165,690 = 4.0\%$. Burgess et al. [15] show that bias from overlap is a linear function of the overlap proportion, scaled by the instrument strength, so that it is approximately the overlap fraction divided by F. Burgess et al. also treat the case-control setting, and report that when risk-factor data cover both cases and controls, as ours do, bias is similar with a binary outcome as with a continuous one, so the approximation applies here. Burgess et al. derive it under a null causal effect. At the median F of 855 it gives 4.7×10^{-5} of the observational confounding. Even at the worst case we could construct, complete overlap with our weakest retained instrument at $F = 30$, the approximation gives 3.3×10^{-2} .

Table 1
Colocalization and cross-platform replication for the 28 proteins nominally significant in both GWAS

Protein	OR Nikpay (95% CI)	OR Aragam (95% CI)	CAD GW- sig in region	PP.H4	PP.H3	Independent signals	Verdict	SomaScan F	PP.H4 SomaScan	PAV instrument	Adjusted p Aragam
PCSK9	1.26 (1.14– 1.39)	1.25 (1.20– 1.31)	23	1.000	0.000	2	shared	246	1.000	yes	4.8e-26
FGF5	1.08 (1.05– 1.11)	1.06 (1.05– 1.08)	13	0.998	0.002	4	shared	8	0.025	no	2.6e-11
TGFB1	1.30 (1.18– 1.45)	1.25 (1.19– 1.31)	264	0.997	0.003	2	shared	9	0.012	no	4.6e-17
COL6A3	1.35 (1.07– 1.71)	1.31 (1.17– 1.47)	0	0.976	0.018	2	shared (no CAD signal)	58	0.001	no	8.7e-04
ANGPTL1	0.90 (0.84– 0.96)	0.93 (0.90– 0.96)	0	0.909	0.033	3	shared (no CAD signal)	905	0.903	no	0.008
PROCR	0.92 (0.88– 0.96)	0.93 (0.91– 0.96)	168	0.896	0.104	3	shared	no assay	-	yes	9.7e-07
DAG1	0.50 (0.36– 0.69)	0.58 (0.48– 0.69)	201	0.877	0.123	4	shared	8	0.019	no	2.9e-07
FLT1	2.04 (1.30– 3.19)	1.77 (1.40– 2.25)	85	0.841	0.159	2	shared	12	0.326	no	8.5e-04
ERBB4	0.89 (0.80– 0.99)	0.90 (0.86– 0.95)	0	0.752	0.044	7	unresolved (no CAD signal)	11	0.005	no	0.049
TIE1	0.89 (0.81– 0.99)	0.91 (0.86– 0.96)	0	0.613	0.040	3	unresolved (no CAD signal)	357	0.571	no	0.102
CNTN2	0.96 (0.94– 0.99)	0.98 (0.96– 0.99)	0	0.466	0.156	4	unresolved (no CAD signal)	2205	0.466	no	0.253
IL1RN	1.10 (1.01– 1.21)	1.07 (1.03– 1.13)	0	0.176	0.093	3	unresolved (no CAD signal)	186	0.171	no	0.904
TFPI	0.92 (0.87– 0.98)	0.93 (0.90– 0.96)	0	0.169	0.823	5	distinct (no CAD signal)	54	0.000	no	0.002
EFNA1	1.08 (1.02– 1.15)	1.04 (1.01– 1.08)	0	0.072	0.092	3	unresolved (no CAD signal)	no assay	-	no	1.000

Mendelian randomization estimates, colocalization, and cross-platform replication for the 28 proteins with a nominally significant effect on coronary artery disease in both outcome GWAS in the same direction. Odds ratios are per standard deviation of inverse-normal transformed protein level. PP.H4 and PP.H3 are posterior probabilities of a shared and of two distinct causal variants against Aragam 2022. Independent signals counts independent genome-wide significant exposure signals in the region by distance clumping at 250 kb, an upper bound. PAV instrument marks a protein-altering sentinel, or one tagging such a variant at r -squared 0.8 or above. Adjusted p is the Bonferroni-corrected MR p -value against Aragam 2022 across the 359 proteins comparable in both GWAS. For FST, the instrument lies outside the ± 500 kb colocalization region.

Additional files

Additional file 1 (.docx). STROBE-MR checklist. A completed STROBE-MR checklist for this manuscript, giving the manuscript text or page number that supports each item.

Additional file 2 (.csv). Instrument associations. For each of the 552 proteins with a cis instrument, the sentinel variant (GRCh37 position and alleles), its association with the protein (beta, standard error, p -value, $-\log_{10}$ p -value, F , sample size) and with coronary artery disease in Nikpay 2015 and Aragam 2022 (beta, standard error, p -value, harmonization status), with outcome effects aligned to the protein's effect allele. The p -value column states the protein association as the source prints it, including values below floating-point precision, and the $-\log_{10}$ p -value column is computed from that printed value.

Protein	OR Nikpay (95% CI)	OR Aragam (95% CI)	CAD GW-sig in region	PP.H4	PP.H3	Independent signals	Verdict	SomaScan F	PP.H4 SomaScan	PAV instrument	Adjusted p Aragam
EPHA2	0.80 (0.69–0.93)	0.91 (0.84–0.99)	0	0.030	0.177	3	unresolved (no CAD signal)	79	0.011	no	1.000
TIMP3	0.95 (0.93–0.98)	0.97 (0.96–0.99)	0	0.023	0.924	4	distinct (no CAD signal)	5524	0.022	no	0.551
NRP1	0.92 (0.87–0.98)	0.95 (0.92–0.99)	0	0.021	0.836	3	distinct (no CAD signal)	425	0.026	no	1.000
CSF1	1.13 (1.02–1.26)	1.09 (1.03–1.15)	0	0.016	0.898	1	distinct (no CAD signal)	13	0.018	no	1.000
BMP6	0.87 (0.76–0.99)	0.93 (0.87–1.00)	0	0.012	0.125	2	unresolved (no CAD signal)	19	0.006	yes	1.000
FST	1.18 (1.03–1.36)	1.11 (1.03–1.19)	0	0.009	0.283	3	unresolved (no CAD signal)	20	0.005	no	1.000
LTBP3	1.06 (1.02–1.11)	1.05 (1.02–1.07)	52	0.001	0.999	3	distinct	no assay	-	no	0.022
MMP12	0.94 (0.90–0.97)	0.97 (0.95–0.99)	16	0.000	0.999	3	distinct	1057	0.000	no	1.000
CSF2	0.90 (0.82–0.99)	0.94 (0.89–0.99)	22	0.000	1.000	3	distinct	6	0.017	yes	1.000
LPL	0.92 (0.87–0.97)	0.93 (0.90–0.95)	267	0.000	1.000	3	distinct	no assay	-	yes	2.8e-05
APOE	0.91 (0.88–0.94)	0.91 (0.90–0.92)	97	0.000	1.000	3	distinct	468	0.000	yes	1.3e-35
ANGPTL4	0.92 (0.86–0.98)	0.94 (0.90–0.97)	7	0.000	1.000	1	distinct	27	1.000	no	0.226
COL4A1	0.87 (0.78–0.96)	0.87 (0.82–0.92)	157	0.000	1.000	2	distinct	no assay	-	no	2.5e-04
LPA	1.22 (1.17–1.26)	1.23 (1.21–1.26)	915	0.000	1.000	4	distinct	no assay	-	no	1.4e-115
Mendelian randomization estimates, colocalization, and cross-platform replication for the 28 proteins with a nominally significant effect on coronary artery disease in both outcome GWAS in the same direction. Odds ratios are per standard deviation of inverse-normal transformed protein level. PP.H4 and PP.H3 are posterior probabilities of a shared and of two distinct causal variants against Aragam 2022. Independent signals counts independent genome-wide significant exposure signals in the region by distance clumping at 250 kb, an upper bound. PAV instrument marks a protein-altering sentinel, or one tagging such a variant at r-squared 0.8 or above. Adjusted p is the Bonferroni-corrected MR p-value against Aragam 2022 across the 359 proteins comparable in both GWAS. For FST, the instrument lies outside the ± 500 kb colocalization region.											
Additional files											
Additional file 1 (.docx). STROBE-MR checklist. A completed STROBE-MR checklist for this manuscript, giving the manuscript text or page number that supports each item.											
Additional file 2 (.csv). Instrument associations. For each of the 552 proteins with a cis instrument, the sentinel variant (GRCh37 position and alleles), its association with the protein (beta, standard error, p-value, -log10 p-value, F, sample size) and with coronary artery disease in Nikpay 2015 and Aragam 2022 (beta, standard error, p-value, harmonization status), with outcome effects aligned to the protein's effect allele. The p-value column states the protein association as the source prints it, including values below floating-point precision, and the -log10 p-value column is computed from that printed value.											

All analyses are carried out in GRCh37 coordinates. The exposures and Nikpay 2015 are distributed on that build while Aragam 2022 is not. The harmonized Aragam release is distributed on GRCh38, so regions fetched from it were validated against the original GRCh37 identifier retained in its

markername field, and rows whose identifier did not resolve to the requested locus were discarded. Across the 28 regions fetched this way, 866 rows were rejected.

Estimation and the permutation null

The MR estimators were implemented in Python and validated against cases with known answers: a hand-computed Wald ratio, agreement between inverse-variance weighting [16] on a single variant and the Wald ratio, recovery of an injected pleiotropic offset by MR-Egger [17], and the weighted median [18] returning the true effect under 27% contamination where inverse-variance weighting does not. Twenty-one such tests pass, and seven more check variant selection. Because each protein enters the main analysis through a single cis-instrument, the main analysis has only one genetic variant per exposure. Cochran's Q, I^2 , MR-Egger and the weighted median do not apply to it. They are used here only as validated estimators in the test suite above. The colocization code was checked the same way, by 14 tests, and then against the reference implementation: on the R coloc package's own test dataset it reproduces that package's published posterior probabilities to within 3.4×10^{-7} relative error on all five hypotheses.

Our harmonization drops palindromic variants outright [19]. In a single-sentinel design, this removes the protein, and 74 of the 552 sentinels (13.4%) are palindromic. Outcome rows were retrieved by rsID. Eight instruments present in Nikpay at the same GRCh37 position and alleles under a different identifier were therefore not matched. A ninth, TRAF2, has a malformed rsID in the source summary statistics, where the rsID field reads "9" instead of the correct rs7045531, so it was matched in neither outcome GWAS, removing two directed pairs from Aragam, TNF–TRAF2 and GSTP1–TRAF2. MRC1's Aragam variant is present at the same position, but its rsID has no GRCh38 mapping, so MRC1 was excluded from the Aragam analysis. HJV's Aragam variant is reported on the opposite strand, a non-palindromic complement of alleles (A/G against T/C) that harmonization does not resolve, so HJV was excluded from the Aragam analysis. As a sensitivity analysis, Nikpay outcome rows were re-extracted by GRCh37 position instead of rsID, which raised comparable unordered pairs from 476 to 494 and the proteins entering them from 361 to 377, raised the Bonferroni-significant pairs from 10 to 11 (adding TGFB1–ITGB6), added LTB to the seven proteins significant after Bonferroni correction in both GWAS, and left the same-direction fraction at 48.8% against a rewiring-null p of 0.71, compared with 48.5% and p = 0.64 in the main analysis, so the conclusion is unchanged. Of the 552 proteins, 361 enter a comparable pair against Nikpay and 399 against Aragam. Effects are per standard deviation of inverse-normal transformed protein level on the log-odds of coronary artery disease. For a pair, the difference between the ligand-based and receptor-based estimate was tested as $z = (b_L - b_R) / \sqrt{(se_L^2 + se_R^2)}$, the two instruments being at different loci. Protein-level MR p-values were adjusted across the 359 proteins comparable in both GWAS, by Bonferroni and by the Benjamini-Hochberg false discovery rate.

OmniPath records some interactions in both directions. Because the difference statistic is symmetric in the two proteins, the pair with A as ligand and B as receptor is the same comparison as the reverse, and we deduplicated to unordered pairs (removing 14 pairs in the Nikpay set and 22 in the Aragam set). Two permutation schemes were used: *Edge rewiring* keeps every protein's own estimate and its own degree and randomizes only which receptor each ligand is paired with. This raises the question: is the observed pairing special? *Node shuffling* moves the (beta, se) pairs to different proteins, holding the network fixed. Both used 20,000 permutations with seed 20260829. Empirical p-values are two-sided with a + 1 correction so that p is never exactly zero.

Because the headline result is a null, we computed what it could have detected. The standard deviation of the same-direction fraction under the rewiring null gives the smallest excess detectable at 80% power and a two-sided alpha of 0.05. We also simulated the alternative directly on the observed network: each protein is assigned a sign; a fraction of edges is then made concordant by copying one partner's sign to the other, with edges taken in random order, so that a protein already assigned keeps its sign and the dependence induced by shared proteins is preserved. Power is the fraction of 2,000 simulated datasets exceeding the null's upper 2.5% point.

Colocalization and cross-platform replication

For colocization, we used the approximate Bayes factor method of Giambartolomei et al. [20], keeping the default variant priors ($p_1 = p_2 = 1 \times 10^{-4}$, $p_{12} = 1 \times 10^{-5}$) and coloc's default effect priors, which differ by trait type: 0.15 on the inverse-normal protein scale and 0.2 on the log-odds scale for the binary outcome. This too was written in Python and checked on constructed regions whose answer is fixed in advance: a shared variant must return H4, two distinct variants must return H3, a single-trait signal must return H1 with the symmetric result on swapping traits, and posteriors must sum to one without the H3 log-difference going negative. Regions are ± 500 kb around the gene.

We also counted, per region, how many independent genome-wide significant exposure signals are present, because the method assumes one causal variant per trait. Signals were counted by distance clumping: the lead variant is taken, every variant within 250 kb of it is removed, and the step repeats. Without a linkage disequilibrium reference this counts signals separated by distance, which can only overemphasize how many are independent, so it is an upper bound. The Fenland study [9] measures plasma proteins on SomaScan, an aptamer-based platform, and is distributed by the same resource in an identical file format. Median sample size is 8,348, roughly a fifth of the UK Biobank Pharma Proteomics Project. Since a failure to replicate at that sample size may reflect power, every comparison is reported alongside the SomaScan instrument F-statistic. SomaScan F-statistics are taken from the lead variant of the SomaScan–Nikpay harmonized overlap, while SomaScan posteriors are computed against Aragam 2022.

Protein-altering instruments

Each instrument was annotated with the Ensembl Variant Effect Predictor [21]. An instrument was called protein-altering only if the consequence fell on the gene encoding the protein being measured. We also tested whether a lead variant might tag such a variant, restricting candidates to genome-wide significant variants inside the gene body and requiring $r^2 \geq 0.8$ in 1000 Genomes European samples [22]. At most six candidate variants per protein were tested. For six proteins, linkage disequilibrium could not be retrieved, and these were counted as not flagged (none is among the 28).

Results

Of 660 proteins, 552 (83.6%) have a cis instrument that meets genome-wide significance. The remaining 108 (16.4%) do not. They are absent from any proteome-wide MR that defines instruments this way. Among the 629 proteins with any genome-wide significant signal, the strongest variant lies in cis for 384 (61.0%) and in trans for 245 (39.0%) (Fig. 2). Signal inside the extended MHC is common. Over all 660 proteins the median share of genome-wide significant variants falling there is 2.1% (3.4% among the 629 with any signal), but 217 of 660 (32.9%) exceed 20%, 122 (18.5%) exceed 50%, and 14 (2.1%) exceed 90%. Of the 1,129 ligand–receptor pairs, 397 (35.2%) cannot be compared because at least one partner has no usable instrument; harmonization losses, mostly palindromic sentinels, remove more, leaving 476 comparable unordered pairs against Nikpay and 534 against Aragam.

Next, we asked whether the proteins without a cis instrument are a biologically significant set. Common variation that moves a protein's level should be purged at genes that cannot tolerate it. This prediction was specified in advance; it failed, and the effect reversed. The comparison uses the 641 of 660 proteins with gnomAD constraint data, 537 with an instrument and 104 without. Median LOEUF [13] was 0.865 for proteins without an instrument and 0.801 for proteins with one (one-sided $p = 0.85$; common-language effect size 0.468). Neither pLI (one-sided $p = 0.11$) nor missense z (one-sided $p = 0.54$) supported it either. Mutational target size is associated with it instead: genes lacking an instrument are shorter, median coding sequence 840 bp against 1,284 bp ($p = 4.5 \times 10^{-4}$), expected loss-of-function count 12 against 19 ($p = 1.0 \times 10^{-3}$). Split the proteins into quartiles of coding-sequence length and no constraint difference survives anywhere (one-sided $p = 0.06$ –0.44). None of which makes the set insignificant. It contains LRP1 (LOEUF 0.058), CACNA1C, DLL4, IL2RG, ITGB1, and CTLA4, which is a licensed drug target.

Ligand and receptor estimates

Where both partners carry a usable instrument, the two estimates share no information about coronary artery disease beyond what the individual proteins carry. Against Nikpay, 48.5% of pairs pointed the same way. The configuration model null gave 49.8% ($p = 0.64$). Against Aragam, 47.0% against 50.0% ($p = 0.20$). Counts of pairs differing at nominal significance were 53 observed against 54 expected ($p = 0.85$) and 133 against 125 ($p = 0.25$). The Bonferroni counts resulted in ten pairs differing after correction in Nikpay against a null expectation of 9 ($p = 0.72$); thirty-one differed in Aragam against 30 ($p = 0.85$). Figure 3 shows each observed statistic against its null. Tripling the case count multiplies the apparent yield threefold and the null by a similar factor, because precision rises equally for random pairings as for real ones.

Causal proteins are also not clustered in the network. In Aragam, assortativity of the MR z-score across edges was -0.009 against a null of -0.005 ($p = 0.92$), and edges joining two proteins with $|z| > 1.96$ numbered 25 against an expected 32 ($p = 0.22$), slightly lower than chance. One pre-specified mechanism was established before the run: pairs in which either partner's strongest signal maps to the other partner's locus, a clearance signature, should show greater ligand-versus-receptor divergence than pairs whose strongest signals are both in cis. Six pairs carried the signature against 232 cis–cis reference pairs in Nikpay; the median $|z|$ of the difference was 0.83 against 0.87 (one-sided Mann-Whitney $p = 0.48$), and opposite-direction estimates occurred in 4 of 6 against 131 of 232 (one-sided Fisher $p = 0.48$). In Aragam the same test gave 6 against 232, medians 1.62 against 1.21 ($p = 0.21$), and 4 of 6 against 119 of 232 (one-sided $p = 0.37$). The Aragam test covers the 486 of 556 directed Aragam pairs that are also classified in the Nikpay pair set. The prediction failed at both sample sizes.

False positives

The standard analysis resulted in ten pairs differing after Bonferroni correction across 476 deduplicated pairs, headed by the TGFB1–CD109 pair: ligand OR 1.30 ($p = 3.4 \times 10^{-7}$), receptor OR 1.00 ($p = 0.79$), difference $p = 8.5 \times 10^{-7}$. However, this is not a significant finding for three reasons. The 476 pairs were built from 361 proteins. SDC1, SDC4 and ITGB2 each appear in 18 pairs, TGFB1 in 13. Each estimate is reused in every pair containing it. Bonferroni over 476 therefore corrects for a family of tests that does not exist. The ten pairs are built from 13 distinct proteins, but three account for all ten. TGFB1 drives six, DAG1 three and FGF5 one. Of their ten partners, none has a detectable coronary effect (all $p > 0.16$). The reported yield is inflated roughly threefold by reuse. Against a null that preserves the network, the expected count is 8.7. Node shuffling gave $p = 0.014$ for the count of nominally different pairs in Aragam (133 observed against 100 expected; in Nikpay, 53 against 45, $p = 0.33$). That null is unsuitable for this network.

Node shuffling assumes a protein's estimate is exchangeable with any other protein's even though it is not. Degree correlates with effect size (Spearman $\rho = +0.133$, $p = 0.008$), so hub proteins carry larger effects, and moving estimates to random nodes destroys that structure and deflates the null from the initial value of 125 to 100. Edge rewiring preserves both each protein's estimate and its degree, changes only the pairing, and returns $p = 0.25$. We checked that rewiring is not itself biased by a role asymmetry. Ligand and receptor instruments do not differ in strength (median F 702 versus 580 over the 490 directed Nikpay pairs, $p = 0.90$; median standard error 0.0612 versus 0.0621).

Protein-altering instruments

Twenty-eight of the 359 proteins that enter a comparable pair in both GWAS had an effect on coronary artery disease that reached nominal significance in both, in the same direction (Table 1), led in Aragam by LPA ($z = 23.1$), APOE (-12.9), PCSK9 (11.1) and TGFB1 (9.06). After Bonferroni correction across those 359 proteins, seven of the twenty-eight remain significant in both GWAS: APOE, DAG1, FGF5, LPA, PCSK9, PROCRA and TGFB1. Fifteen remain significant in Aragam. A Benjamini-Hochberg false discovery rate of 5% selects the same seven in both GWAS and twenty-one of the twenty-eight in Aragam. The twenty-eight are carried forward as the nominal set because that is what the routine analysis reports. A missense variant in an antibody's epitope changes measured binding rather than protein concentration, so those instruments are routinely flagged as a sensitivity analysis. Eighty-four of 552 sentinels (15.2%) are protein-altering in their own gene, rising to 16.7% once instruments tagging such a variant at $r^2 \geq 0.8$ are included. This flag is not enriched among causal proteins. Five of 28 (17.9%) carry a protein-altering sentinel, against 79 of the other 524 (15.1%; odds ratio 1.22, Fisher $p = 0.60$); counting LD-tagged variants as well, 6 of 28 (21.4%) against 86 of the other 524 (16.4%; odds ratio 1.39, $p = 0.44$). Worse, the six flagged are APOE, PCSK9, PROCRA, LPL and CSF2, whose sentinels alter their own protein, and BMP6, whose sentinel tags such a variant. PCSK9's instrument is rs11591147, the R46L loss-of-function variant, and PCSK9 is the best-validated target in cardiovascular medicine. Structural categorization alone cannot separate a missense instrument that influences biology from one that influences binding.

Colocalization verdicts

Only 14 of the 28 proteins sit at a locus with any genome-wide significant coronary signal in Aragam. In Nikpay, it is 4. The remaining estimates rest on sub-genome-wide outcome associations and cannot colocalize no matter their pQTL. Of those 14, six colocalize ($PP.H4 \geq 0.8$: PCSK9 1.000, FGF5 0.998, TGFB1 0.997, PROCRA 0.896, DAG1 0.877, FLT1 0.841) and eight are refuted ($PP.H3 \geq 0.8$: LTBP3, MMP12, CSF2, LPL, APOE, ANGPTL4, COL4A1, LPA). Eight of fourteen, 57%, reach $PP.H3 \geq 0.8$, against the 31.5% of MR associations that Zheng et al. report as unsupported by colocalization [23]. The denominators differ, theirs being 413 associations over many phenotypes and ours fourteen coronary loci pre-filtered for genome-wide outcome signal.

Running all 28 regardless returns 8 colocalizations and 12 refutations (Fig. 4). COL6A3 reaches $PP.H4 = 0.976$ at a locus whose strongest coronary variant is $p = 4.5 \times 10^{-7}$, so no genome-wide signal is present and the posterior is being driven by sub-threshold noise in both traits. A high $PP.H4$ is not evidence of a shared causal variant when neither trait has a genome-wide significant signal, and colocalization should be reported alongside the outcome signal at the locus. Swapping the outcome GWAS alters the result, defined as $PP.H4 \geq 0.8$, at 3 of 14 loci. The same parameters we applied to COL6A3 apply here, and only 4 of these 14 loci carry genome-wide significant signal in Nikpay as well as Aragam. Therefore, at two of the three changes (DAG1, FLT1), the Nikpay posterior is itself resting on sub-genome-wide signal. APOE is the one reversal at a locus where both GWAS carry genome-wide signal, from 0.000 in Aragam to 0.902 in Nikpay. PCSK9 falls from 1.000 to 0.825 for the same power reason. The Nikpay region contains exactly one variant at $p < 5 \times 10^{-8}$. Twenty-six of the 28 regions contain more than one independent genome-wide significant exposure signal, between two and seven, so the single-causal-variant assumption fails in almost all; only ANGPTL4 and CSF1 carry a single signal. The correct treatment is conditional or SuSiE-based colocalization [24], which requires a linkage disequilibrium reference we do not have. A $PP.H3$ of 1.000 states that the variant influencing the measured protein differs from the variant influencing disease. At LPA, where the allelic architecture is dominated by KIV-2 copy number that single-nucleotide instruments do not capture, and where the region carries 915 genome-wide significant coronary variants, this is not evidence that lipoprotein(a) is non-causal.

Twenty-two proteins have assays on both Olink and SomaScan. The platforms reached the same colocalization verdict for 16 and disagreed for 6. Disagreement is associated with instrument strength (Fig. 5). The absolute difference in $PP.H4$ falls as the SomaScan instrument strengthens (Spearman $\rho = -0.630$, $p = 0.002$), and proteins whose platforms agree carry stronger ones than those that disagree (median F 216 against 10, two-sided Mann-Whitney $p = 0.017$). The $F > 10$ rule was applied when selecting Olink instruments; on SomaScan the strongest cis variant was taken regardless, so that the comparison is between platforms, and four of the 22 fall at or below $F = 10$ (TGFB1 at 9, FGF5 at 8, DAG1 at 8, CSF2 at 6). The SomaScan F -statistic is taken from the SomaScan–Nikpay harmonized overlap, while $PP.H4$ SomaScan is computed against Aragam 2022. Dropping them leaves 15 of 18 agreeing and the association in the same direction at similar strength ($\rho = -0.562$, $p = 0.015$), so the four weak instruments are not carrying it. PCSK9, at $F = 246$, replicates almost exactly: OR 1.25 on Olink against 1.31 on SomaScan, $PP.H4 = 1.000$ on both. ANGPTL4 reverses completely ($PP.H4$ 0.000 on Olink against 1.000 on SomaScan) at $F = 27$, COL6A3 goes from 0.976 to 0.001 at $F = 58$, and FLT1 goes from 0.841 to 0.326 at $F = 12$. One protein with a weak instrument agrees anyway (CSF2, $F = 6$). Instrument strength explains part of the platform disagreement among these 22 proteins and cannot explain all of it. Check F before attributing a platform disagreement to assay chemistry. Separately, LPA, PROCRA, LPL, COL4A1, LTBP3 and EFNA1 have no SomaScan assay in the Fenland release, so for two major lipid targets cross-platform validation was not available.

Pipeline findings

PCSK9 recovers throughout: OR 1.26 in Nikpay and 1.25 in Aragam, $z = 11.1$, $PP.H4 = 1.000$, and cross-platform agreement. The two outcome GWAS agree with each other over the 359 proteins comparable in both, correlating log-odds estimates at Spearman $\rho = 0.575$, $p = 5.4 \times 10^{-33}$ (on z statistics, $\rho = 0.633$). PCSK9 shows the pipeline can detect a real effect, so the nulls above are not a failure to measure.

Discussion

An estimate built from a ligand's cis-pQTL and one built from its receptor's do not provide interchangeable evidence about a pathway. Across 476 (Nikpay) and 534 (Aragam) comparable pairs and two coronary GWAS differing threefold in case count, ligand and receptor estimates agreed at the rate of randomly paired proteins, and causal proteins showed no tendency to sit beside one another in the interaction network. The design is not, however, sensitive to weak concordance. Against the same rewiring null, at 80% power and a two-sided alpha of 0.05, the smallest detectable same-direction fraction is 0.560 in Aragam and 0.562 in Nikpay. When concordance is simulated on the real network, which propagates the constraint that partners share proteins, about 14% of pairs in both GWAS would have to be truly concordant before being detected by this design. The null excludes pathway concordance in more than roughly one pair in seven.

Multiple-testing correction over units assembled from a smaller set of entities is invalid. Three proteins accounted for all ten reported discoveries. Any analysis comparing estimates across the edges of a graph (protein–protein, gene–gene, metabolite pathways) faces this obstacle. In order to overcome this, print the number of distinct entities and the degree distribution before correction. Node shuffling and edge rewiring differ by whether the entity keeps its own value, and where degree correlates with effect size, only rewiring is valid. Ours returned $p = 0.014$ and $p = 0.25$ on the same data. Annotating instruments as protein-altering is a widely used sensitivity analysis, and on this evidence it does not identify measurement artifacts. Cross-platform comparison [25] is the better test, though it too is partly confounded by instrument strength and does not settle every case. Colocalization is treated as a property of a locus and reported from a single outcome GWAS. Three of fourteen verdicts changed when we changed the GWAS. Two of those three sit at loci with no genome-wide signal in the second GWAS, so they are attributed to differences in outcome power, and only one locus with signal in both reversed. At minimum, a colocalization claim should state the outcome GWAS and, where possible, be shown against a second.

Limitations

The single-causal-variant assumption underlying our colocalization fails in 26 of the 28 regions examined, and we lack the linkage disequilibrium reference needed [24], so the colocalization results are indicative. Cross-platform comparison rests on 22 proteins. Fenland is roughly a fifth the size of the UK Biobank Pharma Proteomics Project, so absence of replication there is weak evidence. Nikpay 2015 is 77% European ancestry and Aragam 2022 predominantly European. We did not formally test the direction of causation. No bidirectional Mendelian randomization or Steiger filtering was run, though every instrument is cis to the protein's own gene, which limits, but does not exclude, reverse causation from disease to protein level. Olink measures relative abundance, and its assays are subject to epitope effects that we could only partially characterize [26]. Fourteen of twenty-eight nominally causal proteins sit at loci with no genome-wide coronary signal at all, so that set is weaker than its replication in two GWAS suggests. Finally, a null result over 476 (Nikpay) and 534 (Aragam) comparable pairs does not exclude pathway-level concordance for particular pairs with strong instruments on both sides. We are not claiming that the proteins refuted by colocalization are non-causal for coronary artery disease, that platform disagreements are assay artifacts, or that Mendelian randomization is unreliable in general. PCSK9 was recovered in every analysis.

Conclusions

The six failure modes documented here are a coverage blind spot in which instrument availability drops a sixth of the proteins; multiple-testing correction over pairs assembled from fewer proteins; a node-shuffling null that is biased when degree correlates with effect size; protein-altering-variant flags that do not identify artifacts; single-causal-variant colocalization in regions with several signals, whose verdicts change with the outcome GWAS; and platform disagreement confounded by instrument strength. Ligand-level and receptor-level Mendelian randomization answer different questions, and treating either as evidence about a shared pathway is not supported by data at current sample sizes.

Abbreviations

GWAS
genome-wide association study
LOEUF
loss-of-function observed/expected upper bound fraction
MHC
major histocompatibility complex
MR
Mendelian randomization
OR
odds ratio
pQTL
protein quantitative trait locus
PP
posterior probability.

Declarations

Ethics approval and consent to participate.

Not applicable. All analyses used publicly available summary statistics from studies with their own ethical approvals.

Consent for publication.

Not applicable.

Clinical trial number

not applicable.

Competing interests.

The author declares no competing interests.

Funding.

None.

Author Contribution

PP conceived the study, wrote all analysis code, performed all analyses, prepared all figures and tables, and wrote the manuscript. PP is the sole author and approves the submitted version.

Acknowledgements.

Not applicable.

Data Availability

No new data were generated. All datasets analyzed are publicly available third-party resources. Plasma protein summary statistics (UK Biobank Pharma Proteomics Project, Olink; and Fenland, SomaScan) are at https://omicscience.charite.de/Sexspecific_pGWAS/. Coronary artery disease summary statistics are from the GWAS Catalog. Nikpay 2015 is accession GCST003116 (<https://www.ebi.ac.uk/gwas/studies/GCST003116>) and Aragam 2022 is accession GCST90132314 (<https://www.ebi.ac.uk/gwas/studies/GCST90132314>). Ligand-receptor interactions are from OmniPath (<https://omnipathdb.org/>). Constraint metrics are gnomAD v2.1.1 (<https://gnomad.broadinstitute.org/downloads>). Analysis code, including unit tests for the Mendelian randomization and colocalization estimators and the random seeds used for all permutations, is available from the author on reasonable request.

References

1. Schmidt AF, Finan C, Gordillo-Maranon M, et al. Genetic drug target validation using Mendelian randomisation. *Nat Commun.* 2020;11(1):3255. 10.1038/s41467-020-16969-0.
2. Davey Smith G, Ebrahim S. Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003;32(1):1–22. 10.1093/ije/dyg070.
3. Li C, De Jay N, Zhang SS, et al. Proteome-Wide Mendelian Randomization Identifies Candidate Causal Proteins for Cardiovascular Diseases. *Adv Genet (Hoboken).* 2025;6(2):2500003. 10.1002/ggn2.202500003.
4. Swerdlow DI, Holmes MV, Kuchenbaecker KB, et al. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. *Lancet.* 2012;379(9822):1214–24. 10.1016/S0140-6736(12)60110-X.
5. IL6R Genetics Consortium Emerging Risk Factors Collaboration. Interleukin-6 receptor pathways in coronary heart disease: a collaborative meta-analysis of 82 studies. *Lancet.* 2012;379(9822):1205–13. 10.1016/S0140-6736(11)61931-4.
6. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: The STROBE-MR Statement. *JAMA.* 2021;326(16):1614–21. 10.1001/jama.2021.18236.
7. Koprulu M, Wheeler E, Kerrison ND, et al. Sex differences in the genetic regulation of the human plasma proteome. *Nat Commun.* 2025;16:4001. 10.1038/s41467-025-59034-4.
8. Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature.* 2023;622(7982):329–38. 10.1038/s41586-023-06592-6.

9. Pietzner M, Wheeler E, Carrasco-Zanini J, et al. Mapping the proteo-genomic convergence of human diseases. *Science*. 2021;374(6569):eabj1541. 10.1126/science.abj1541.
10. Nikpay M, Goel A, Won HH, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47(10):1121–30. 10.1038/ng.3396.
11. Aragam KG, Jiang T, Goel A, et al. Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet*. 2022;54(12):1803–15. 10.1038/s41588-022-01233-6.
12. Turei D, Korcsmaros T, Saez-Rodriguez J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat Methods*. 2016;13(12):966–7. 10.1038/nmeth.4077.
13. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434–43. 10.1038/s41586-020-2308-7.
14. Burgess S, Thompson SG. Avoiding bias from weak instruments in Mendelian randomization studies. *Int J Epidemiol*. 2011;40(3):755–64. 10.1093/ije/dyr036.
15. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol*. 2016;40(7):597–608. 10.1002/gepi.21998.
16. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658–65. 10.1002/gepi.21758.
17. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–25. 10.1093/ije/dyv080.
18. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–14. 10.1002/gepi.21965.
19. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *Elife*. 2018;7:e34408. 10.7554/eLife.34408.
20. Giambartolomei C, Vukcevic D, Schadt EE, et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet*. 2014;10(5):e1004383. 10.1371/journal.pgen.1004383.
21. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122. 10.1186/s13059-016-0974-4.
22. 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526(7571):68–74. 10.1038/nature15393.
23. Zheng J, Haberland V, Baird D, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet*. 2020;52(10):1122–31. 10.1038/s41588-020-0682-6.
24. Wallace C. A more accurate method for colocalisation analysis allowing for multiple causal variants. *PLoS Genet*. 2021;17(9):e1009440. 10.1371/journal.pgen.1009440.
25. Kuliesius J, Badonyi M, Navarro P, Marsh JA, Klaric L, Wilson JF. Epitope Effect Prevalence in Affinity-based pQTL studies. *bioRxiv [Preprint]*. 2025. 10.1101/2025.06.20.660695
26. Eldjarn GH, Ferkingstad E, Lund SH, et al. Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature*. 2023;622(7982):348–58. 10.1038/s41586-023-06563-x.

Figures

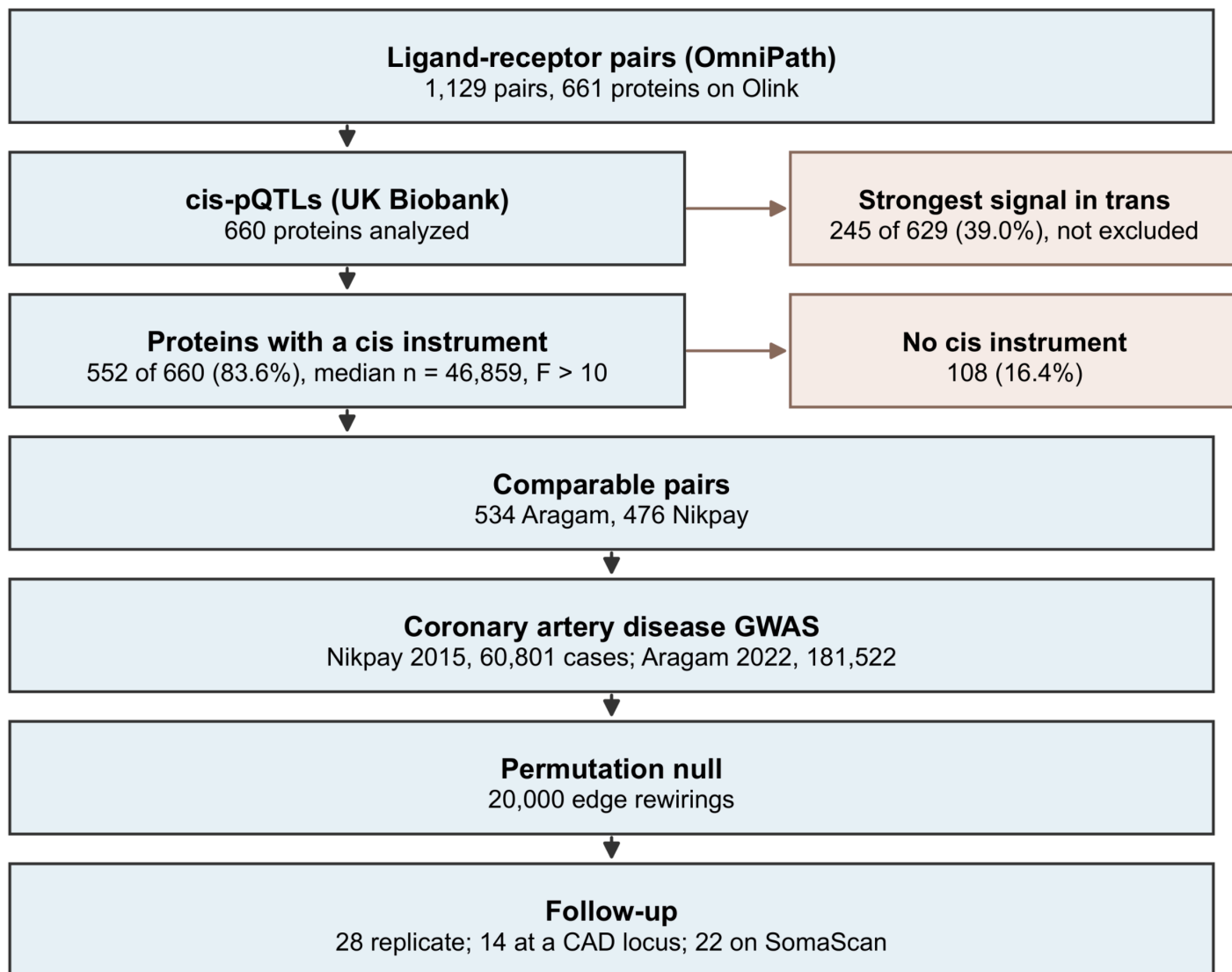


Figure 1

Study design and attrition from the ligand-receptor proteome to the analyzed pairs. Flow of proteins and pairs through the analysis. The main column, in blue, carries the set forward at each step; the branches to the right, in warm gray, record what each step sets aside or flags and why. Color encodes only that distinction. Of the 660 proteins analyzed, 552 have a genome-wide significant cis instrument and are retained at $F > 10$. The 108 without one are excluded, and are absent from any proteome-wide MR that defines instruments this way. Comparable pairs number 534 against Aragam 2022 and 476 against Nikpay 2015, after deduplication to unordered pairs.

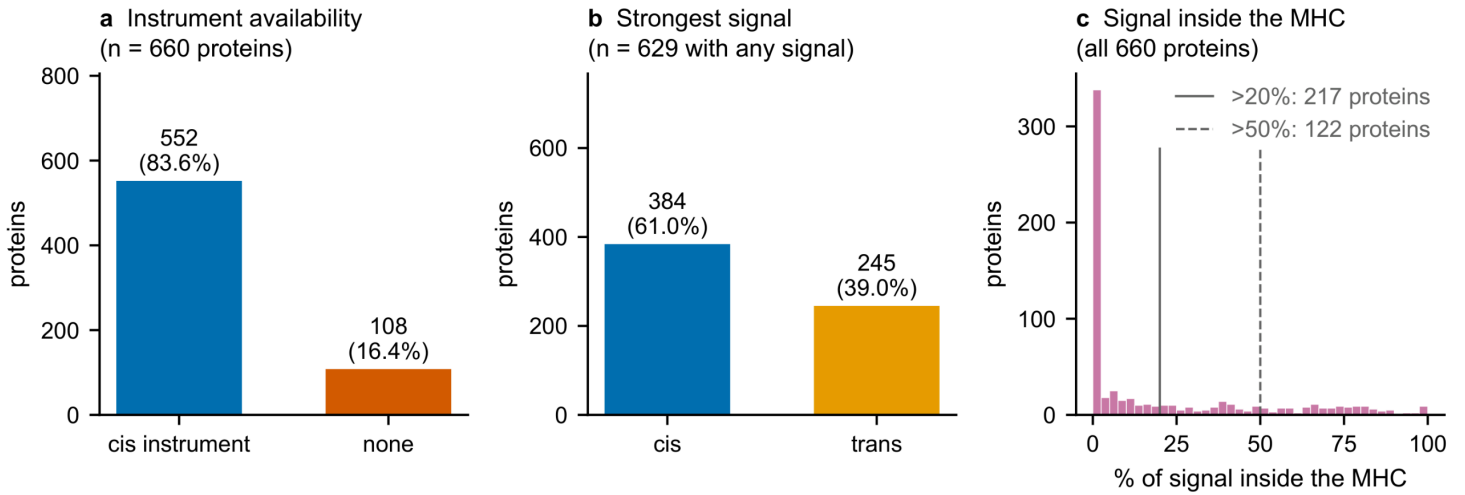


Figure 2

Cis-instrument availability and the location of each protein's genetic signal. (a) Proteins with and without a cis instrument, of 660 audited. (b) Among the 629 proteins carrying any genome-wide significant signal, whether the single strongest variant falls in cis or in trans; 384 are cis. (c) Distribution of the percentage of each protein's genome-wide significant signal that falls inside the extended MHC (chr6:28,477,797–33,448,354, GRCh37), across all 660 proteins; 217 carry more than 20% there.

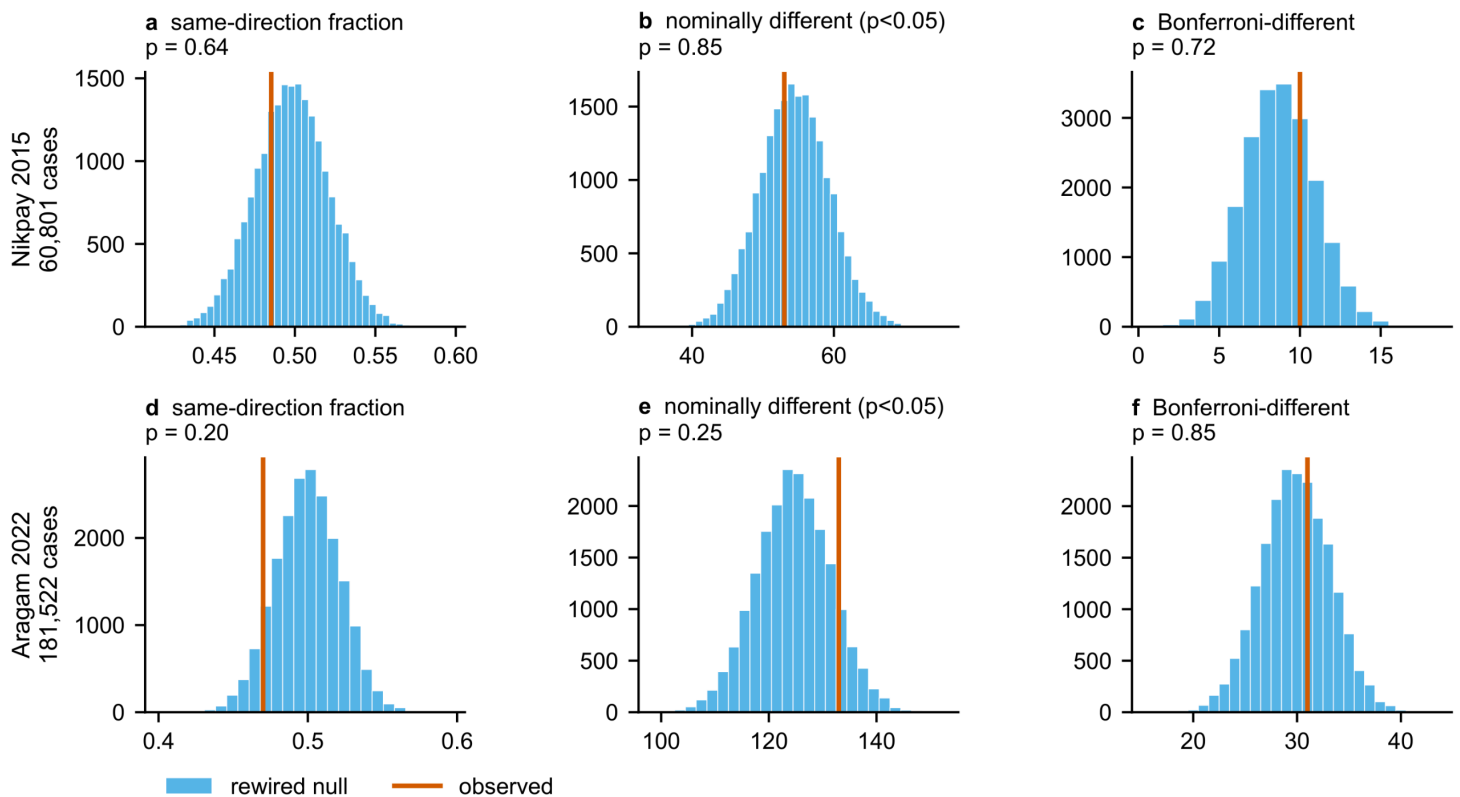


Figure 3

Observed discordance statistics against their edge-rewiring permutation nulls. Each panel shows the null distribution from 20,000 edge rewirings (blue histogram) with the observed value overlaid (orange line). Panels a–c are Nikpay 2015 and d–f Aragam 2022. Columns are the three discordance statistics: same-direction fraction, count of nominally different pairs ($p < 0.05$), and count of Bonferroni-different pairs. Two-sided permutation p-values are given above each panel.

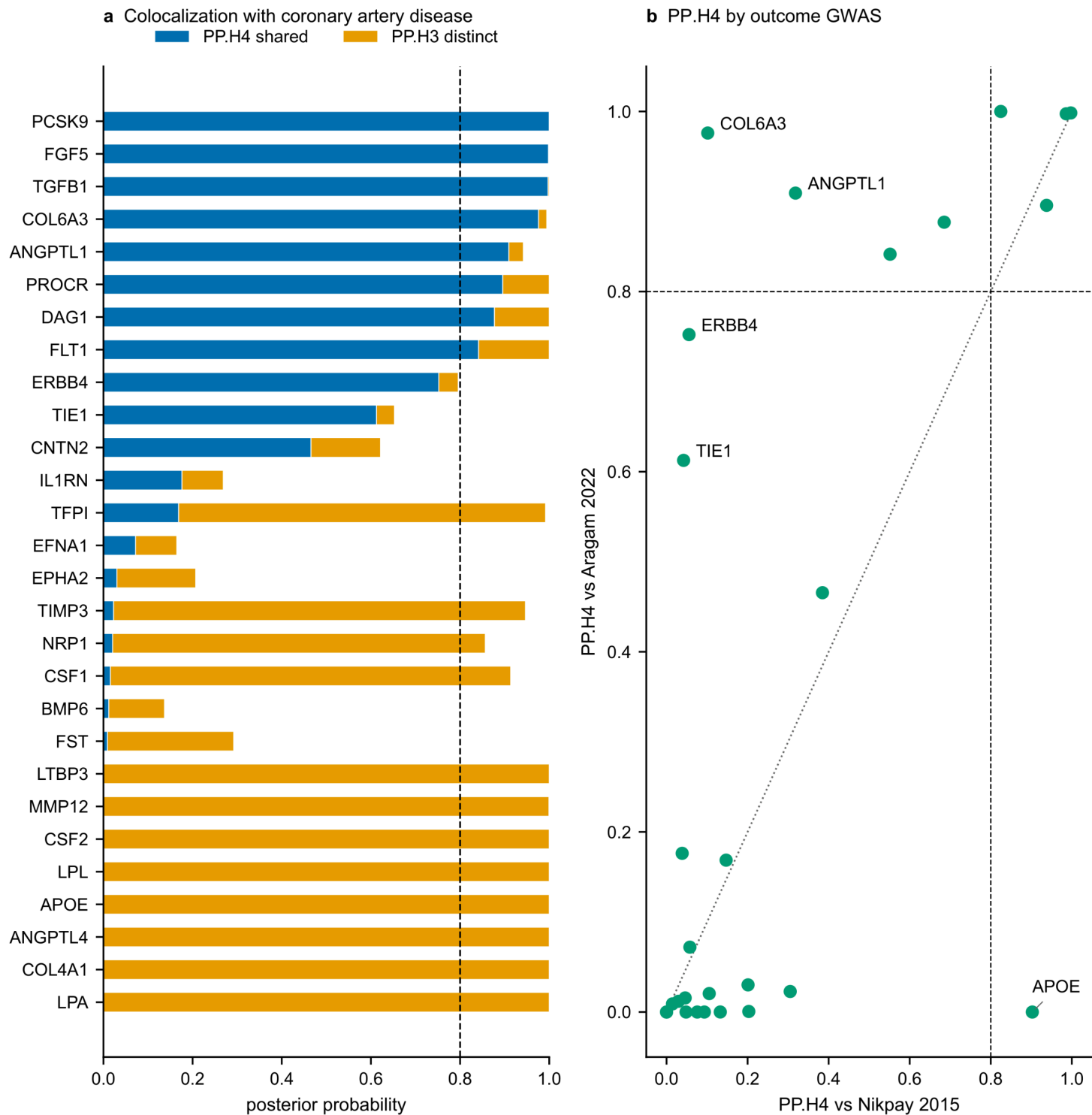


Figure 4

Colocalization of protein and coronary signals, and its dependence on the outcome GWAS. (a) Posterior probabilities for the 28 proteins carried into follow-up, against Aragam 2022. Bars are stacked: PP.H4, one causal variant shared by protein and disease, and PP.H3, two distinct causal variants. The dashed line marks the conventional 0.8 threshold. (b) The same PP.H4 computed against Nikpay 2015 and against Aragam 2022, with the identity line dotted and the 0.8 thresholds dashed. Points off the identity line are loci whose posterior moves with the outcome GWAS. Only those crossing a threshold line change verdict, which five do. The five whose PP.H4 differs by more than 0.4 between the two GWAS are labeled, and these are not the same five.

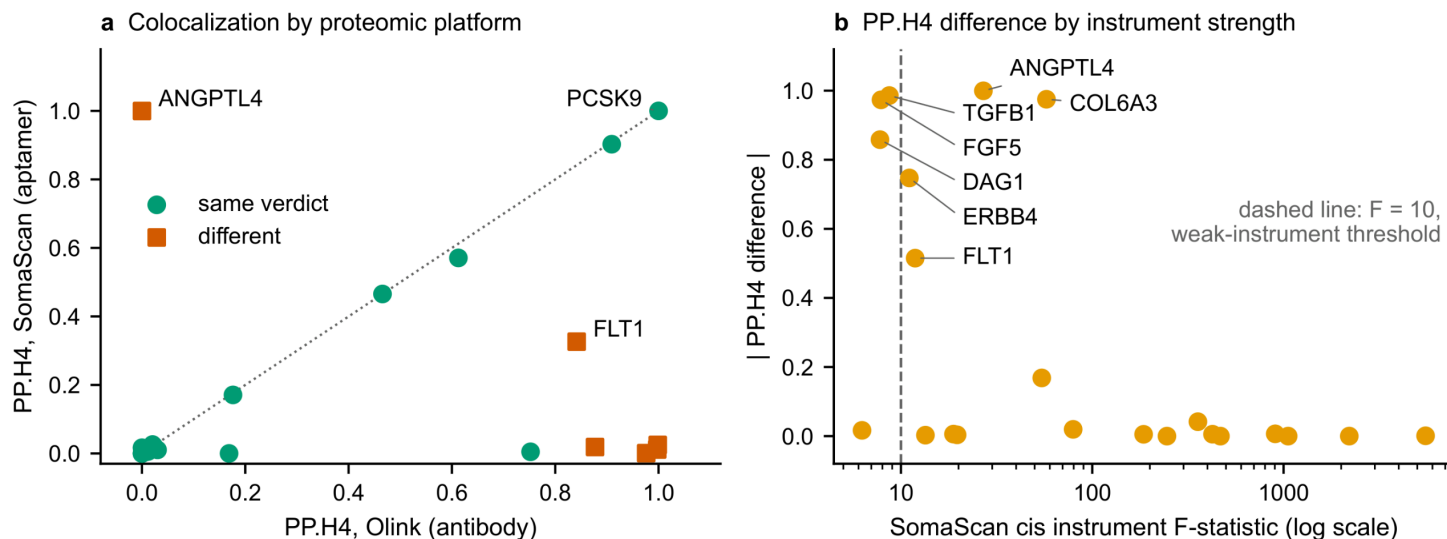


Figure 5

Cross-platform agreement between Olink and SomaScan, and its dependence on instrument strength. (a) PP.H4 against Aragam 2022 computed from Olink (antibody) and from SomaScan (aptamer) instruments, for the 22 proteins assayed on both. Points are colored by whether the two platforms return the same verdict at the 0.8 threshold (two overlap); the identity line is dotted. (b) The absolute difference between the two platforms' PP.H4 against the SomaScan cis instrument F-statistic, on a log scale, with the $F = 10$ weak-instrument threshold dashed. Proteins whose platforms disagree by more than 0.4 are labeled.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1STROBEMRchecklist.docx](#)
- [Additionalfile2instrumentassociations.csv](#)