Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genetic Architectures of Smoking-Induced Cardiovascular Disease in East Asians: A Multi-Omics Mendelian Randomization Study of the HP and CYP2A Loci

Shuo Ying
Tianjin Chest Hospital
Qiang Chen
Tianjin Chest Hospital
Li Zhang
Tianjin Chest Hospital
Naikuan Fu
Tianjin Chest Hospital
Peng Zhang
pzhang0822@163.com

Tianjin Chest Hospital

Research Article

Keywords: Smoking, Cardiovascular disease, Mendelian randomization, ALDH2, Multi-omics decoupling, HP and CYP2A loci

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11747256/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

Objective

Smoking is a major modifiable cardiovascular risk factor. However, its causal spectrum and underlying genetic architecture in East Asian populations remain incompletely characterized. A critical challenge is the highly prevalent *ALDH2* rs671 variant, whose profound pleiotropic effects complicate genetic causal inference in this population. Furthermore, it remains unknown whether smoking-induced molecular alterations propagate linearly from DNA methylation to gene expression and circulating proteins (the mQTL–eQTL–pQTL axis). We aimed to dissect the causal spectrum of smoking in East Asians independent of this pleiotropy, and to map the genetic continuity of the multi-omic architecture.

Methods

We performed two-sample Mendelian randomization (MR) for two smoking phenotypes—cigarettes per day (CPD) and smoking initiation—on nine cardiovascular outcomes in East Asians. Instrumental SNPs were rigorously screened, with the *ALDH2* rs671 variant explicitly excluded. Two-step MR was used to assess cardiometabolic and alcohol mediators. Summary-data-based MR (SMR) was integrated with East Asian mQTL, eQTL, and pQTL data. Bayesian colocalization and spatial *cis*-overlap analysis were performed to validate the molecular architecture. Additionally, sample-overlap sensitivity, reverse MR, and Steiger filtering were conducted to ensure causal robustness.

Results

After excluding rs671, genetically proxied CPD robustly increased the risk of peripheral artery disease (PAD) in the BioBank Japan cohort (IVW OR = 2.51, 95% CI 1.19–5.33, $P = 0.016$), which was independently replicated in a separate European Bioinformatics Institute (EBI) East Asian cohort (OR = 2.22, 95% CI 1.10–4.49, $P = 0.026$). Associations with the other eight cardiovascular outcomes were null. Two-step MR showed no causal effect of CPD on lipids, type 2 diabetes, or BMI (all P > 0.18). Multi-omics integration revealed convergent signals at the *HP* and *CYP2A* loci across epigenomic (mQTL, $p_SMR < 1\times10^{-5}$), transcriptomic (eQTL, *HP* $P = 9.86\times10^{-7}$; *CYP2A7* $P = 5.86\times10^{-8}$), and proteomic (pQTL, *HP* OR = 0.876; *CYP2A* OR = 0.949) layers. However, colocalization did not support a shared causal variant (*HP* PP.H4 = 0.066; *CYP2A* PP.H4 = 0.069), contrasting with the positive control *ABO* (PP.H4 = 0.840). Spatial analysis showed CPD instruments were located within 72 kb of *CYP2A7* but were not covered by the pQTL panel. Steiger filtering confirmed the forward direction. Reactome annotation placed CYP2A in xenobiotic metabolism and *HP* in heme scavenging.

Conclusion

In East Asians, smoking exerts a highly specific causal effect on PAD, independent of the *ALDH2*–alcohol axis and traditional cardiometabolic mediators. The *HP* and *CYP2A* loci converge across multi-omic layers but are genetically decoupled. These findings extend the companion European study and highlight the necessity of ancestry-specific confounder control and distinguishing molecular convergence from shared causal continuity in multi-omic MR studies.

Introduction

Cardiovascular disease (CVD) remains the leading cause of mortality worldwide, and smoking is one of the most important modifiable risk factors [1, 2]. The cardiovascular toxicity of smoking is highly heterogeneous, affecting diverse vascular beds and cardiac structures, ranging from atherosclerotic diseases such as coronary artery disease (CAD) and peripheral artery disease (PAD), to structural and electrophysiological disorders such as heart failure (HF) and atrial fibrillation (AF) [3, 4]. Although large-scale genetic and multi-omic studies have substantially advanced our understanding of smoking-related CVD, the vast majority of this evidence has been generated in European ancestry populations [5, 6]. East Asian populations possess distinct genetic architectures and environmental exposures; therefore, a comprehensive dissection of the causal spectrum of smoking across diverse cardiovascular subtypes in East Asians is urgently needed to inform precision prevention.

A major methodological challenge in conducting genetic causal inference in East Asian populations is the highly prevalent *ALDH2* rs671 variant. This variant profoundly alters alcohol metabolism and is classically known as the "alcohol-flushing polymorphism" [7]. Because alcohol consumption and smoking behaviors are highly intertwined and culturally embedded in East Asian societies, rs671 strongly influences behavioral confounders [8]. In Mendelian randomization (MR), where genetic variants are used as instrumental variables, rs671 acts as a severe pleiotropy trap that violates the exclusion-restriction assumption [9]. Failure to rigorously screen and exclude this variant can generate profound spurious associations, such as falsely attributing CAD risk to smoking. Therefore, ancestry-specific confounder control is an absolute prerequisite for any valid genetic study of smoking in East Asians.

Beyond the macroscopic causal spectrum, the underlying molecular architecture of smoking-induced cardiovascular injury remains incompletely characterized. It is classically hypothesized that smoking-induced alterations propagate linearly from DNA methylation (mQTL) to gene expression (eQTL), and ultimately to circulating proteins (pQTL)[10, 11]. However, recent multi-omic evidence from a European cohort suggests that this cascade is not strictly linear. A companion study has revealed a profound "multi-layer decoupling," where epigenetic imprints and protein effectors operate through distinct genetic architectures rather than a single shared causal variant [12]. Given the distinct genetic background of East Asians, it remains entirely unknown whether this molecular decoupling is conserved across ancestries, or whether it manifests through ancestry-specific regulatory architectures at distinct loci.

To address these critical gaps, we performed an integrative multi-omics MR study in East Asians. Our objectives were threefold: (1) to systematically estimate the causal effects of two smoking phenotypes

(cigarettes per day and smoking initiation) on nine cardiovascular outcomes while strictly excluding the *ALDH2* rs671 pleiotropy; (2) to exclude traditional cardiometabolic and alcohol mediators using two-step MR; and (3) to map the genetic continuity of the mQTL–eQTL–pQTL axis at the target loci. By integrating Bayesian colocalization, spatial *cis*-overlap analysis, and Reactome pathway annotation, we formally tested whether the multi-omic signals converge biologically while remaining genetically decoupled. This work provides a cross-ancestry perspective on the molecular architecture of smoking-induced cardiovascular injury and extends the findings of our companion European study.

Methods

Study Design

We performed an integrative multi-omics Mendelian randomization (MR) study using publicly available summary-level genome-wide association study (GWAS) data to dissect the causal effects of smoking on cardiovascular diseases (CVDs) in East Asian populations. The overall study design is summarized in Fig. 1. Ethical approval was not required, as all data were derived from publicly available databases with appropriate institutional review board approvals.

Figure 1 Study design. Two East Asian smoking phenotypes (cigarettes per day and smoking initiation) were used to estimate causal effects on nine cardiovascular outcomes using two-sample Mendelian randomization. Instrumental SNPs were rigorously screened against behavioral confounders, with the highly pleiotropic *ALDH2* rs671 variant explicitly excluded. The primary spectrum analysis revealed a highly specific positive association with PAD, while the other eight CVDs were null. Molecular mechanisms were investigated using mQTL-SMR, eQTL-SMR, pQTL MR, Bayesian colocalization, and spatial *cis*-overlap analysis at the *HP* and CYP2A loci.

Black arrows denote the main study flow and causal effects. The purple dashed arrow denotes the cross-talk between the macroscopic spectrum and molecular mechanisms. Box colors indicate macro-level exposures (blue), multi-omics data sources (purple), confounder and spectrum analysis (green), multi-omics convergence and decoupling (yellow), and the final conceptual conclusion (dashed gray). SNP, single nucleotide polymorphism; CPD, cigarettes per day; MR, Mendelian randomization; mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; SMR, summary-data-based Mendelian randomization; PP.H4, posterior probability of a shared causal variant; PAD, peripheral artery disease; CAD, coronary artery disease; HF, heart failure; DVT, deep vein thrombosis; IS, ischemic stroke; AF, atrial fibrillation; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia.

Data Sources

Exposures: Two smoking phenotypes were selected as primary exposures in East Asians: cigarettes per day (CPD; GWAS ID: ieu-b-5071; N = 72,655) and smoking initiation (GWAS ID: bbj-a-78; N = 165,436),

both derived from the BioBank Japan (BBJ).

Outcomes: Nine cardiovascular outcomes were analyzed: peripheral artery disease (PAD), coronary artery disease (CAD), heart failure (HF), deep vein thrombosis (DVT), ischemic stroke (IS), atrial fibrillation (AF), aortic aneurysm (AA), dilated cardiomyopathy (DCM), and ventricular arrhythmia (VA). Summary statistics were obtained from the BBJ, the GWAS Catalog, and published East Asian meta-analyses. Detailed dataset characteristics are provided in Table 1.

Multi-omics Data

Blood *cis*-mQTL data were obtained from the East Asian meta-analysis (Hatton et al. 2024), blood *cis*-eQTL data from the 1kcp East Asian dataset, and plasma *cis*-pQTL data from the Nagahama cohort. The 1000 Genomes Project East Asian (EAS) reference panel was used for linkage disequilibrium (LD) calculations.

Instrumental Variables (IVs) Selection

For each smoking behavior, independent genetic instruments were selected at $P < 5\times10^{-8}$ and clumped at $r^2 < 0.001$ within a 10,000 kb window. The *F*-statistic was calculated to quantify instrument strength, and SNPs with $F < 10$ were excluded.

Given the unique genetic architecture of East Asians, the *ALDH2* rs671 variant was stringently screened against alcohol intake and educational attainment. It was excluded a priori due to its extreme pleiotropy ($P < 1 \times 10^{-300}$ for alcohol intake). Furthermore, each selected SNP was queried against five behavioral/social confounders: alcohol intake, educational attainment, household income, physical activity, and mental disorders. After stringent screening, five independent SNPs for CPD and seven independent SNPs for smoking initiation were retained for downstream analyses (Supplementary Table S1).

Statistical Analysis

Univariable MR (UVMR)

UVMR analyses were primarily performed using the inverse-variance weighted (IVW) method with a random-effects model. To ensure robustness, the weighted median (WM) estimator and MR-Egger regression were performed as sensitivity analyses. Effect sizes are presented as odds ratios (ORs) with 95% confidence intervals (CIs). Multiple testing was controlled using the Benjamini-Hochberg false discovery rate (FDR) within each exposure-outcome matrix. Heterogeneity was assessed using Cochran's *Q* statistic, and horizontal pleiotropy was evaluated using the MR-Egger intercept and MR-PRESSO.

Mediation and Pleiotropy Assessment

To determine whether the smoking-CVD effects were mediated by traditional cardiometabolic traits, two-step MR was performed using the retained CPD instruments against five mediators: LDL-C, HDL-C, triglycerides, type 2 diabetes, and BMI. None of the retained CPD SNPs were significantly associated with alcohol intake (all P > 0.09), and these SNPs were largely absent from the UK Biobank alcohol GWAS after standard quality control. Consequently, an alcohol-adjusted multivariable MR (MVMR) was not feasible due to an insufficient number of overlapping instruments. This lack of association strongly supports the absence of horizontal pleiotropy via the alcohol pathway.

Multi-Omics Integration (SMR and pQTL MR)

Summary-data-based MR (SMR) was used to integrate East Asian mQTL and eQTL data with CVD GWAS. SMR *P*-values were calculated using the Wald ratio, and heterogeneity was assessed using the HEIDI test. For proteomic layers, two-sample MR was performed using *cis*-pQTL instruments for *HP* and the *CYP2A* cluster (*CYP2A7/CYP2A13/CYP2B6*) to estimate their causal effects on peripheral artery disease (PAD), the primary cardiovascular outcome identified in our UVMR analysis.

Bayesian Colocalization and Spatial Overlap

To test whether the multi-omic signals shared a common causal variant, Bayesian colocalization was performed using the coloc.abf method. PP.H4 > 0.75 was considered strong evidence for a shared causal variant. *ABO* (an established causal locus for thrombotic diseases) was used as a positive control. To further explore the regulatory architecture, we performed a spatial *cis*-overlap analysis. The physical genomic distance between the retained CPD instruments and the target genes (*HP* on chr16, *CYP2A* cluster on chr19) was calculated, with ± 1 Mb defined as the *cis*-region.

Reverse MR and Sample Overlap Sensitivity

To rule out reverse causation, bidirectional MR was performed using ischemic stroke (IS) as the exposure and CPD as the outcome, as the East Asian IS GWAS provided a sufficient number of genome-wide significant independent instruments. Steiger filtering was applied to mathematically confirm the causal direction. Because the exposure and outcome GWAS were both derived from East Asian populations, sample overlap was formally evaluated using the mrSampleOverlap R package. The expected bias and type I error rates were estimated under a wide range of hypothetical overlap proportions (0% to 50%) (Supplementary Table S5).

Functional Enrichment

To contextualize the biological functions of the identified targets, targeted Reactome pathway annotation was performed using the ReactomePA package (version 1.54.0). Target genes (*HP*, *CYP2A6*, *CYP2A7*, and *CYP2A13*) were mapped to Entrez IDs, and an over-representation analysis (ORA) was conducted using all annotated genes in org.Hs.eg.db as the background. Multiple testing was controlled

using the Benjamini-Hochberg false discovery rate (FDR), and pathways with an adjusted $P < 0.05$ were considered significant.

Results

Instrumental variables and ALDH2 pleiotropy screening

In the East Asian population, a total of 12 independent SNPs were identified for the two smoking phenotypes (five for CPD and seven for smoking initiation) after strict clumping and confounder screening. All retained instruments had an F-statistic > 10, indicating negligible weak instrument bias. Crucially, the *ALDH2* rs671 variant was excluded a priori due to its extreme pleiotropy with alcohol intake ($P \approx 0$) and educational attainment ($P = 4.63\times10^{-5}$). As shown in the phenome-wide screening, failing to exclude this variant would generate profound spurious associations (Supplementary Table S1). For instance, the inclusion of rs671 falsely attributed a strong protective effect to CPD on coronary artery disease (CAD), which was entirely nullified after its exclusion (Fig. 2). Summary data for all exposures and outcomes were successfully retrieved (Table 1).

Table 1
Baseline characteristics of the GWAS datasets and instrumental variables

Trait Category	Specific Trait	Sample size (n)	Case (n)	Control (n)	SNP (n)	Ethnicity	Data source
Exposures	CPD	72,655	-	-	6,108,828	EAS	BBJ
	Smoking initiation	165,436	83,810	81,626	5,961,480	EAS	BBJ
Outcomes	PAD	177,713	4,112	173,601	24,186,090	EAS	Sakaue
	CAD	212,453	29,319	183,134	8,881,048	EAS	BBJ
	IS	174,686	4,018	170,668	13,435,541	EAS	BBJ
	HF	185,457	13,385	172,072	5,887,003	EAS	Han
	AF	159,690	8,180	151,510	13,433,295	EAS	BBJ
	DVT	309,628	2,418	307,210	15,588,353	EAS	Taiwan Biobank
	AA	174,756	2,825	171,931	13,435,467	EAS	BBJ
	DCM	179,000	1,444	177,556	19,906,330	EAS	BBJ
	VA	157,213	1,673	155,540	13,426,630	EAS	BBJ
Multi-omics	Blood *cis*-mQTL	2,099	-	-	80,394	EAS	Hatton
	Blood *cis*-eQTL	1,101	-	-	26,300,000	EAS	1kcp EAS
	Plasma *cis*-pQTL	1,823	-	-	4,600,000	EAS	Nagahama Cohort

GWAS, genome-wide association study; CPD, cigarettes per day; BBJ, BioBank Japan; PAD, peripheral artery disease; CAD, coronary artery disease; IS, ischemic stroke; HF, heart failure; AF, atrial fibrillation; DVT, deep vein thrombosis; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia; mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; EAS, East Asian.

Vascular-bed specificity of smoking in East Asians

We evaluated the causal spectrum of genetically proxied CPD across nine cardiovascular outcomes. A striking vascular-bed specificity was observed. CPD robustly increased the risk of peripheral artery disease (PAD) (IVW OR = 2.51, 95% CI 1.19–5.33, P = 0.016 in BBJ; and OR = 2.22, 95% CI 1.10–4.49, P = 0.026 in European Bioinformatics Institute (EBI); weighted median OR = 3.03, 95% CI 1.98–4.62, P < 0.001). In stark contrast, CPD demonstrated null associations with all other cardiovascular outcomes

evaluated, including CAD (OR = 0.996, *P* = 0.960), ischemic stroke (OR = 0.991, *P* = 0.917), atrial fibrillation (OR = 0.960, *P* = 0.828), aortic aneurysm (OR = 1.164, *P* = 0.635), dilated cardiomyopathy (OR = 1.032, *P* = 0.899), ventricular arrhythmia (OR = 0.869, *P* = 0.659), heart failure (OR = 0.936, *P* = 0.607), and deep vein thrombosis (OR = 1.163, *P* = 0.504) (Fig. 2; Supplementary Table S2). The smoking initiation phenotype yielded directionally consistent but statistically underpowered estimates, constrained by limited instrument variables (Table 2).

Figure 2 Cardiovascular outcome spectrum of smoking in East Asians. Forest plot displaying the causal effects of genetically proxied cigarettes per day (CPD) on nine cardiovascular outcomes in East Asian populations. The estimates are inverse-variance weighted (IVW) odds ratios (ORs) with 95% confidence intervals (CIs). All analyses were performed after the rigorous exclusion of the highly pleiotropic *ALDH2* rs671 variant. A highly specific vascular-bed tropism was observed: CPD was significantly associated with an increased risk of peripheral artery disease (PAD) (highlighted in red; *P* = 0.016 for the BBJ cohort). A consistent replication was observed in an independent EBI cohort (OR = 2.22, *P* = 0.026; data not shown in the forest plot). In stark contrast, genetically proxied CPD showed null associations with all other eight cardiovascular outcomes (highlighted in black; all *P* > 0.05). The dashed vertical line indicates the null hypothesis (OR = 1). The exact P-values for each outcome are annotated on the right panel. CPD, cigarettes per day; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; PAD, peripheral artery disease; CAD, coronary artery disease; IS, ischemic stroke; AF, atrial fibrillation; HF, heart failure; DVT, deep vein thrombosis; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia.

Table 2
Univariable MR associations between smoking phenotypes and nine cardiovascular outcomes in East Asians

Exposure	Outcome	SNP (n)	Method	OR (95% CI)	*P*
CPD	PAD (BBJ)	5	IVW	2.51(1.19–5.33)	0.02*
	PAD (EBI)	5	IVW	2.22(1.10–4.49)	0.03*
	CAD	5	IVW	0.996(0.86–1.16)	0.96
	IS	5	IVW	0.99(0.84–1.17)	0.92
	AF	5	IVW	0.96(0.67–1.39)	0.83
	HF	3	IVW	0.94(0.73–1.20)	0.61
	DVT	5	IVW	1.16(0.75–1.81)	0.51
	AA	5	IVW	1.16(0.62–2.18)	0.64
	DCM	5	IVW	1.03(0.63–1.69)	0.90
	VA	5	IVW	0.87(0.47–1.62)	0.66
Smoking initiation	PAD (BBJ)	4	IVW	2.80(0.12–64.79) **	0.52
	PAD (EBI)	4	IVW	6.07(0.31-119.14) **	0.23
	CAD	4	IVW	2.25(0.45–11.19) **	0.32
	IS	7	IVW	0.42 (0.10–1.69) **	0.22
	AF	7	IVW	0.94 (0.10–9.11) **	0.96
	HF	3	IVW	2.32(0.07–79.68) **	0.64
	DVT	4	IVW	0.80(0.02–37.82) **	0.91
	AA	7	IVW	0.07 (0.001–4.59) **	0.21
	DCM	7	IVW	2.12 (0.03-162.32) **	0.74
	VA	7	IVW	1.09 (0.02–55.50) **	0.97

Estimates represent the causal effects of genetically proxied smoking phenotypes (CPD and smoking initiation) on nine cardiovascular outcomes in East Asians. All analyses were performed after excluding the highly pleiotropic *ALDH2* rs671 variant. The inverse-variance weighted (IVW) method was used. For smoking initiation, four SNPs were retained for primary outcomes (PAD, CAD, HF, DVT) after strict confounder screening, while seven SNPs were available for the extended outcomes (IS, AF, AA, DCM,

VA). * Indicates statistical significance (P < 0.05). ** Indicates highly imprecise estimates (variance explosion) due to insufficient instrumental variables, interpreted as exploratory. CPD, cigarettes per day; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; BBJ, BioBank Japan; EBI, European Bioinformatics Institute; PAD, peripheral artery disease; CAD, coronary artery disease; IS, ischemic stroke; AF, atrial fibrillation; HF, heart failure; DVT, deep vein thrombosis; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia.

Independence from cardiometabolic and alcohol pathways

To rule out traditional mediation, two-step MR was performed. Genetically proxied CPD showed no causal effect on LDL-C (*P* = 0.186), HDL-C (*P* = 0.771), triglycerides (*P* = 0.470), type 2 diabetes (*P* = 0.609), or BMI (*P* = 0.384) (Supplementary Table S3), indicating that the CPD-PAD effect is independent of cardiometabolic disturbances. We further rigorously excluded alcohol pathway pleiotropy. None of the five retained CPD SNPs were associated with alcohol intake (all *P* > 0.09), and they were largely absent from the UK Biobank alcohol GWAS after standard quality control. Consequently, an alcohol-adjusted MVMR was not feasible due to insufficient overlapping instruments (Supplementary Table S4). To address the potential bias introduced by sample overlap (since both exposure and outcome GWAS were derived from BioBank Japan), a comprehensive sensitivity analysis using the mrSampleOverlap package was conducted. Even under an extreme hypothetical 50% sample overlap, the estimated bias for all 18 exposure-outcome pairs remained zero, and the type I error rate was strictly maintained at the nominal 0.05 level. Furthermore, reverse MR for ischemic stroke identified only the rs671 variant. Steiger filtering confirmed the causal direction was forward (steiger_dir = FALSE, *P* = 0.0037) (Supplementary Table S6), ruling out reverse causation and further validating the *ALDH2* pleiotropy trap.

Multi-omics convergence at the HP and CYP2A loci

To explore the underlying molecular architecture, we integrated East Asian mQTL, eQTL, and pQTL data. We identified convergent signals across three biological layers at the *HP* and *CYP2A* loci. At the epigenomic layer, mQTL-SMR identified significant associations for *HP* (cg23815491, *p*_SMR = 1.71×10^{-6}) and the *CYP2A* cluster (*CYP2B7P* cg10575842, *p*_SMR = 5.06×10^{-6}; *CYP2A13* cg13482327, *p*_SMR = 7.21×10^{-4}) (Supplementary Table S8a). At the transcriptomic layer, eQTL-SMR revealed significant causal links for *HP* (*P* = 9.86×10^{-7}) and *CYP2A7* (*P* = 5.86×10^{-8}) (Supplementary Table S8b). At the proteomic layer, pQTL MR demonstrated that genetically proxied higher levels of plasma *HP* (OR = 0.876, 95% CI 0.82–0.93, *P* = 2.65×10^{-5}) and *CYP2A* (OR = 0.949, 95% CI 0.90–0.996, *P* = 0.035) were protective against PAD (Fig. 3A; Table 3; Supplementary Table S9).

Genetic decoupling despite multi-layer convergence

Despite the multi-layer convergence, formal Bayesian colocalization did not support a shared causal variant for *HP* (PP.H4 = 0.066; PP.H1 = 0.893) or *CYP2A* (PP.H4 = 0.069; PP.H1 = 0.903) (Table 3), indicating that the mQTL and eQTL/pQTL signals at these loci are driven by distinct genetic architectures. In contrast, the positive control *ABO* showed strong colocalization with PAD (PP.H4 = 0.840), confirming the validity of our pipeline (Fig. 3B; Supplementary Table S10). Spatial *cis*-overlap

analysis further clarified the regulatory boundaries. None of the CPD instruments fell within the ± 1 Mb *cis*-window of *HP* (nearest distance > 6.8 Mb). Interestingly, three CPD instruments (rs112673025, rs2545769, rs8100418) were located within 13–72 kb of *CYP2A7*, physically placing them in the *cis*-region. However, these specific SNPs were not covered in the Nagahama pQTL dataset, precluding direct smoking-to-protein MR at this locus (Table 3; Supplementary Table S7). These data collectively demonstrate a multi-layer decoupling: the multi-omic signals converge at the pathway level but are driven by independent genetic variants.

Figure 3 Multi-omics convergence but genetic decoupling at the HP and CYP2A loci. (A) Bubble plot illustrating the convergent signals across epigenomic (mQTL), transcriptomic (eQTL), and proteomic (pQTL) layers for *HP* and *CYP2A*. Bubble size represents the statistical significance (-log10 P); red indicates a positive effect and blue indicates a protective effect. The exact P-values or odds ratios (ORs) are annotated adjacent to the respective bubbles. (B) Bayesian colocalization results evaluating the posterior probability of a shared causal variant (PP.H4) between mQTL/eQTL signals and CVD outcomes for *HP*, *CYP2A*, and the positive control *ABO*. Despite the multi-layer convergence observed in (A), *HP* (PP.H4 = 0.066) and *CYP2A* (PP.H4 = 0.069) did not support a shared causal variant, whereas the positive control *ABO* showed strong colocalization (PP.H4 = 0.840). The dashed red line indicates the threshold for strong colocalization (PP.H4 = 0.75). These findings define a multi-layer decoupling architecture, where biological signals converge at the pathway level but are driven by distinct genetic variants. mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; PP.H4, posterior probability of a shared causal variant; OR, odds ratio; CVD, cardiovascular disease.

Table 3
Multi-omics convergence and genetic decoupling at the *HP* and *CYP2A* loci

Locus	Omics Layer	Evidence / Metric	Result (OR or p-value)	Colocalization (PP.H4)	Spatial *cis*-Overlap (± 1 Mb)
HP	Epigenomic	mQTL-SMR (cg23815491)	$P = 1.71 \times 10^{-6}$	-	-
	Transcriptomic	eQTL-SMR	$P = 9.86 \times 10^{-7}$	-	-
	Proteomic	pQTL-MR (PAD)	OR = 0.876 ($P = 2.65 \times 10^{-5}$)	-	-
	Integration	Bayesian Colocalization	PP.H1 = 0.893	0.066	Non-overlap (nearest > 6.8 Mb)
CYP2A	Epigenomic	mQTL-SMR (cg10575842)	$P = 5.06 \times 10^{-6}$	-	-
	Transcriptomic	eQTL-SMR (CYP2A7)	$P = 5.86 \times 10^{-8}$	-	-
	Proteomic	pQTL-MR (PAD)	OR = 0.949 ($P = 0.035$)	-	-
	Integration	Bayesian Colocalization	PP.H1 = 0.903	0.069	*Cis*-overlap (13–72 kb)*
ABO	Positive Control	eQTL vs PAD GWAS	-	0.84	-

This table summarizes the epigenomic (mQTL), transcriptomic (eQTL), and proteomic (pQTL) signals, formal Bayesian colocalization, and spatial *cis*-overlap analysis for *HP*, CYP2A, and the positive control *ABO*. The Result column displays the specific p-values or odds ratios for each layer, while PP.H1 and PP.H4 indicate the posterior probabilities of association with only one trait versus a shared causal variant, respectively. Dashes (-) indicate that the analysis was not applicable to that specific layer. * The lack of direct smoking-to-protein MR at the CYP2A locus was due to the Nagahama pQTL panel not covering these specific SNPs, rather than a biological absence of regulation. mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; SMR, summary-data-based Mendelian randomization; PP.H4, posterior probability of a shared causal variant; PP.H1, posterior probability of association with only one trait; OR, odds ratio; PAD, peripheral artery disease; GWAS, genome-wide association study.

Functional pathway annotation

To contextualize the biological functions of the identified loci, Reactome pathway annotation was performed. The *CYP2A* cluster (*CYP2A6/2A7/2A13*) mapped predominantly to xenobiotic metabolism

(p.adjust = 1.95×10^{-7}), cytochrome P450 substrate classification (p.adjust = 2.59×10^{-6}), and biological oxidations (p.adjust = 5.10×10^{-5}). In parallel, *HP* mapped to heme scavenging from plasma (p.adjust = 6.61×10^{-3}) and scavenger receptor-mediated ligand uptake (p.adjust = 1.66×10^{-2}). These annotations provide a biologically coherent context for the convergence of *HP* and *CYP2A* in smoking-related PAD (Supplementary Table S11; Supplementary Figure S1).

Discussion

This integrative multi-omics MR study provides a comprehensive dissection of the causal spectrum and molecular architecture of smoking-induced cardiovascular disease (CVD) in East Asian populations. Our findings demonstrate three key insights: (1) the highly prevalent *ALDH2* rs671 variant acts as a severe pleiotropy trap that, if unaddressed, falsely attributes protective effects to smoking; (2) after its exclusion, smoking exhibits a striking vascular-bed specificity by selectively driving PAD while sparing other vascular beds; and (3) despite the convergence of epigenomic, transcriptomic, and proteomic signals at the *HP* and *CYP2A* loci, the underlying genetic architectures are fundamentally decoupled. These findings provide a cross-ancestry perspective and extend the companion European study[12] by defining an ancestry-specific yet mechanistically conserved model of smoking-induced vascular injury.

In East Asian populations, the causal inference of smoking is profoundly complicated by the *ALDH2* rs671 variant. Because rs671 strongly influences alcohol-flushing and drinking behaviors, which are culturally intertwined with smoking, its inclusion violates the exclusion-restriction assumption of MR. Furthermore, this pleiotropy may extend beyond alcohol metabolism, as *ALDH2* has been shown to regulate foam cell formation through interactions with LDLR and AMPK, providing a direct molecular link to atherosclerosis independent of drinking behavior[13]. Additionally, the rs671 variant has been associated with a broader spectrum of cardiometabolic risk factors, including BMI, triglycerides, and metabolic syndrome [8], further complicating its use as a valid instrument for smoking. Our data clearly demonstrate that failing to exclude rs671 falsely attributes a protective effect to smoking on CAD. Once this pleiotropy trap is removed, a robust and highly specific causal effect on PAD emerges (OR = 2.51 in BBJ, independently replicated in EBI with OR = 2.22). This specificity is biologically plausible, as smoking-induced endothelial dysfunction, oxidative stress, and inflammation are particularly potent drivers of the pathophysiology underlying PAD[14, 15]. This is highly consistent with the macroscopic findings of the companion European study, which also identified PAD as a principal target. However, the East Asian data provide a cleaner dissection: while European studies must rely on complex MVMR to disentangle environmental confounders, the unique genetic landscape of East Asians allows us to isolate the pure vascular toxicity of smoking by strictly controlling the *ALDH2*-alcohol axis.

The robustness of the CPD-PAD effect is further highlighted by our mediation analyses. Two-step MR revealed that traditional cardiometabolic traits—such as LDL-C, HDL-C, triglycerides, T2D, and BMI—do not mediate this association. Furthermore, none of the retained CPD instruments were associated with alcohol intake, and comprehensive sample-overlap sensitivity analysis confirmed a bias of zero across all outcomes. This suggests that the smoking-PAD link is driven by direct endothelial injury, oxidative

stress, and systemic inflammation, rather than traditional metabolic disturbances. This finding perfectly aligns with the European companion study, reinforcing the concept that the atherogenic properties of tobacco extend far beyond traditional metabolic dysregulation. A recent multi-omics study also examined the link between tobacco exposure and PAD, proposing IL6/STAT3 as a potential cotinine-related inflammatory mediator[16]. However, our data indicate that the CPD-PAD effect persists independently of traditional cardiometabolic factors, and that the multi-omic signals converge at the HP and *CYP2A* loci rather than through a single linear inflammatory cascade, suggesting a more complex, decoupled molecular architecture.

A major strength of this study is the systematic mapping of the mQTL–eQTL–pQTL axis. We found that the *HP* and *CYP2A* loci converge across all three omics layers. Biologically, Reactome annotation placed *CYP2A* (*CYP2A6/2A7/2A13*) in xenobiotic and nicotine metabolism, and *HP* in heme scavenging and anti-inflammatory pathways, providing a coherent mechanistic context for their roles in PAD. Beyond pathway annotation, direct evidence supports this axis: serum *HP* levels have been causally linked to macroangiopathy in Chinese patients with type 2 diabetes through Mendelian randomization[17], and *HP*'s primary function is to bind cell-free hemoglobin and prevent heme-iron–mediated oxidative damage to the vascular endothelium[18]. Furthermore, the *HP* polymorphism itself has been directly associated with peripheral arterial occlusive disease in clinical cohorts[19]. For *CYP2A*, genetic variation is well established to influence nicotine metabolism and smoking behaviors, providing a mechanistic link between this locus and the exposure itself[20]. However, formal Bayesian colocalization revealed that these multi-omic signals do not share a single causal variant (*HP* PP.H4 = 0.066; *CYP2A* PP.H4 = 0.069), contrasting with the robust shared signal at the positive control *ABO* (PP.H4 = 0.840). This defines a multi-layer decoupling architecture: the biological signals converge at the pathway level, but are driven by distinct genetic variants. This finding is strikingly concordant with the companion European study, which reported a profound transcriptome-proteome decoupling and the absence of mQTL-eQTL colocalization at the classical smoking marker *AHRR*. Together, these two studies across different ancestries establish that smoking-induced molecular alterations do not propagate linearly. Instead, epigenetic imprints and circulating protein effectors operate in parallel, representing constitutional risk pathways rather than a direct upstream-downstream cascade.

We implemented several rigorous defensive analyses. Reverse MR for ischemic stroke identified only rs671, and Steiger filtering mathematically proved the forward direction ($P = 0.0037$), conclusively eliminating reverse causation. The spatial *cis*-overlap analysis elegantly demonstrated that while three CPD instruments are located within 72 kb of *CYP2A7*, the lack of direct smoking-to-protein MR was due to the pQTL panel not covering these specific SNPs, rather than biological non-overlap. Several limitations should be acknowledged. First, the East Asian GWAS for some outcomes (e.g., DCM, VA) have limited sample sizes, reducing statistical power. Second, the smoking instruments capture lifelong genetic liability rather than cumulative exposure. Third, due to the cross-sectional nature of summary-level data, we could not perform tissue-specific multi-omics analysis (e.g., arterial wall or liver), which may reveal different regulatory architectures. Methodologically, our findings of genetic decoupling underscore the limitation of assuming a linear mQTL–eQTL–pQTL cascade. Future integrative analyses

may need to adopt causal network frameworks to capture the parallel and independent pathways that operate across biological layers[21].

Clinically, our findings underscore the necessity of phenotype-specific prevention. For PAD, smoking abstinence is irreplaceable due to its strong, direct, and *ALDH2*-independent causal effect. Furthermore, the identification of *HP* and *CYP2A* pathways suggests potential targets for mitigating smoking-induced vascular toxicity.

Conclusions

In conclusion, in East Asians, smoking exerts a highly specific causal effect on PAD. The *HP* and *CYP2A* loci converge across multi-omic layers but are genetically decoupled. These findings, combined with our companion European study, redefine the molecular architecture of smoking-induced cardiovascular injury and highlight the critical need to distinguish molecular enrichment from shared causal continuity in precision medicine.

Declarations

Ethics approval and consent to participate

As this study involved no primary data collection, separate ethical approval was not required. All datasets used in this study were obtained from publicly available sources with appropriate institutional review board approvals.

Competing interests

The authors declare no competing interests.

Primary funding

The study was conducted independently without any involvement or influence from any of the funding agencies. The study was funded by Tianjin Key Medical Discipline Construction Project (TJYXZDXK-3-017B), Tianjin Health Science and Technology Project (TJWJ2023QN063), Tianjin Health and Science Project (TJWJ2024MS032).

Author Contribution

Formal analysis, Methodology, Software, Data curation, Investigation, Validation, Visualization, Writing – original draft: Shuo Ying, and Qiang Chen; Conceptualization, Supervision, Writing – review & editing: Li Zhang, Naikuan Fu, Peng Zhang.

Acknowledgement

We gratefully acknowledge the investigators and participants of the numerous genome-wide association studies whose summary statistics were used in this study. We thank the BioBank Japan (BBJ) for providing the East Asian GWAS summary statistics for smoking behaviors and cardiovascular outcomes. We thank the MRC-IEU OpenGWAS project and the GWAS Catalog for facilitating public data access. We also extend our gratitude to the Hatton et al. team for providing the East Asian blood cis-mQTL data, the 1kcp consortium for the East Asian blood cis-eQTL data, and the Nagahama cohort for the plasma cis-pQTL summary statistics. We thank the 1000 Genomes Project for the East Asian (EAS) linkage disequilibrium reference panel. Finally, we acknowledge the developers of the R packages used in this study, including TwoSampleMR, coloc, ReactomePA, and mrSampleOverlap, for making robust statistical tools publicly available.

Data availability

All GWAS summary statistics used in this study are publicly available. The smoking exposure datasets were obtained from the MRC-IEU OpenGWAS platform (https://gwas.mrcieu.ac.uk/), specifically CPD (ID: ieu-b-5071) and smoking initiation (ID: bbj-a-78), both derived from BioBank Japan. Cardiovascular outcome summary statistics were obtained from BioBank Japan via the NBDC Human Database (https://humandbs.dbcls.jp/en/), the GWAS Catalog (https://www.ebi.ac.uk/gwas/), and published East Asian meta-analyses. Blood *cis*-mQTL summary statistics were obtained from the East Asian meta-analysis by Hatton et al. (2024). Blood *cis*-eQTL summary statistics were obtained from the 1kcp East Asian dataset. Plasma *cis*-pQTL summary statistics were derived from the Nagahama cohort. LD reference data were obtained from the 1000 Genomes Project East Asian (EAS) reference panel (https://www.internationalgenome.org/). Analysis scripts and the generated Supplementary Tables are available from the corresponding author upon reasonable request. No new data were generated in this study.

References

1. Reitsma MB, Kendrick PJ, Ababneh E, Abbafati C, Abbasi-Kangevari M, Abdoli A, et al. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990–2019: a systematic analysis from the Global Burden of Disease Study 2019. Lancet. 2021;397(10292):2337–60. 10.1016/S0140-6736(21)01169-7.
2. Magnussen C, Ojeda FM, Leong DP, Alegre-Diaz J, Amouyel P, Aviles-Santa L, et al. Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality. N Engl J Med. 2023;389(14):1273–85. 10.1056/NEJMoa2206916.
3. Larsson SC, Mason AM, Bäck M, Klarin D, Damrauer SM, Program MV, et al. Genetic predisposition to smoking in relation to 14 cardiovascular diseases. Eur Heart J. 2020;41(35):3304–10. 10.1093/eurheartj/ehaa193.

4. Pujades-Rodriguez M, George J, Shah AD, Rapsomaniki E, Denaxas S, West R, et al. Heterogeneous associations between smoking and a wide range of initial presentations of cardiovascular disease in 1937360 people in England: lifetime risks and implications for risk prediction. Int J Epidemiol. 2014;44(1):129–41. 10.1093/ije/dyu218.
5. Levin MG, Klarin D, Assimes TL, Freiberg MS, Ingelsson E, Lynch J, et al. Genetics of Smoking and Risk of Atherosclerotic Cardiovascular Diseases: A Mendelian Randomization Study. JAMA Netw Open. 2021;4(1):e2034461. 10.1001/jamanetworkopen.2020.34461.
6. He J, Shen X, Wang M, Gu D, Lu X. Improving Polygenic Risk Prediction for Atherosclerotic Cardiovascular Disease in East Asian Populations. JACC Asia. 2026;6(7):1033–45. 10.1016/j.jacasi.2026.04.038.
7. Zhang J, Guo Y, Zhao X, Pang J, Pan C, Wang J, et al. The role of aldehyde dehydrogenase 2 in cardiovascular disease. Nat Reviews Cardiol. 2023;20(7):495–509. 10.1038/s41569-023-00839-5.
8. Millwood IY, Walters RG, Mei XW, Guo Y, Yang L, Bian Z, et al. Conventional and genetic evidence on alcohol and vascular disease aetiology: a prospective study of 500 000 men and women in China. Lancet (London England). 2019;393(10183):1831–42. 10.1016/S0140-6736(18)31772-0.
9. Au Yeung SL, Jiang C, Cheng KK, Liu B, Zhang W, Lam TH, et al. Is aldehyde dehydrogenase 2 a credible genetic instrument for alcohol use in Mendelian randomization analysis in Southern Chinese men? Int J Epidemiol. 2012;42(1):318–28. 10.1093/ije/dys221.
10. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic Signatures of Cigarette Smoking. Circulation: Cardiovasc Genet. 2016;9(5):436–47. 10.1161/CIRCGENETICS.116.001506.
11. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature. 2018;558(7708):73–9. 10.1038/s41586-018-0175-2.
12. Ying S, Chen Q, Liu X, Zhang P, Fu N, Liu Y. Transcriptome-Proteome Decoupling in Smoking-Induced Cardiovascular Disease: A Multi-Omics Mendelian Randomization Study. Res Square (Preprint). 2026. 10.21203/rs.3.rs-11475023/v1.
13. Zhong S, Li L, Zhang Y-L, Zhang L, Lu J, Guo S, et al. Acetaldehyde dehydrogenase 2 interactions with LDLR and AMPK regulate foam cell formation. J Clin Investig. 2018;129(1):252–67. 10.1172/JCI122064.
14. Behrooz L, Abumoawad A, Rizvi SHM, Hamburg NM. A modern day perspective on smoking in peripheral artery disease. Front Cardiovasc Med. 2023;10:1154708. 10.3389/fcvm.2023.1154708.
15. Wang W, Zhao T, Geng K, Yuan G, Chen Y, Xu Y. Smoking and the Pathophysiology of Peripheral Artery Disease. Front Cardiovasc Med. 2021;8:704106. 10.3389/fcvm.2021.704106.
16. Wei R, Huang M, Zhang Y, He T, Zhu C, Piao H, et al. Tobacco exposure and peripheral artery disease: A multi-omics investigation integrating epidemiology, genetics, toxicology, and single-cell transcriptomics. Ecotoxicol Environ Saf. 2026;314:120083. 10.1016/j.ecoenv.2026.120083.
17. Wang S, Wang J, Zhang R, Wang T, Yan D, He Z, et al. Mendelian randomization analysis to assess a causal effect of haptoglobin on macroangiopathy in Chinese type 2 diabetes patients. Cardiovasc

Diabetol. 2018;17(1):14. 10.1186/s12933-018-0662-7.

18. Schaer CA, Deuel JW, Bittermann AG, Rubio IG, Schoedon G, Spahn DR, et al. Mechanisms of haptoglobin protection against hemoglobin peroxidation triggered endothelial damage. Cell Death Differ. 2013;20(11):1569–79. 10.1038/cdd.2013.113.
19. Delanghe J, Langlois M, Duprez D, De Buyzere M, Clement D. Haptoglobin polymorphism and peripheral arterial occlusive disease. Atherosclerosis. 1999;145(2):287–92. 10.1016/S0021-9150(99)00079-9.
20. Oscarson M. Genetic Polymorphisms in the Cytochrome P450 2A6 (CYP2A6) Gene: Implications for Interindividual Differences in Nicotine Metabolism. Drug Metab Dispos. 2001;29(2):91–5. 10.1124/dmd.29.2.91.
21. Yazdani A, Yazdani A, Mendez-Giraldez R, Samiei A, Kosorok MR, Schaid DJ. From classical mendelian randomization to causal networks for systematic integration of multi-omics. Front Genet. 2022;13:990486. 10.3389/fgene.2022.990486.

Figures

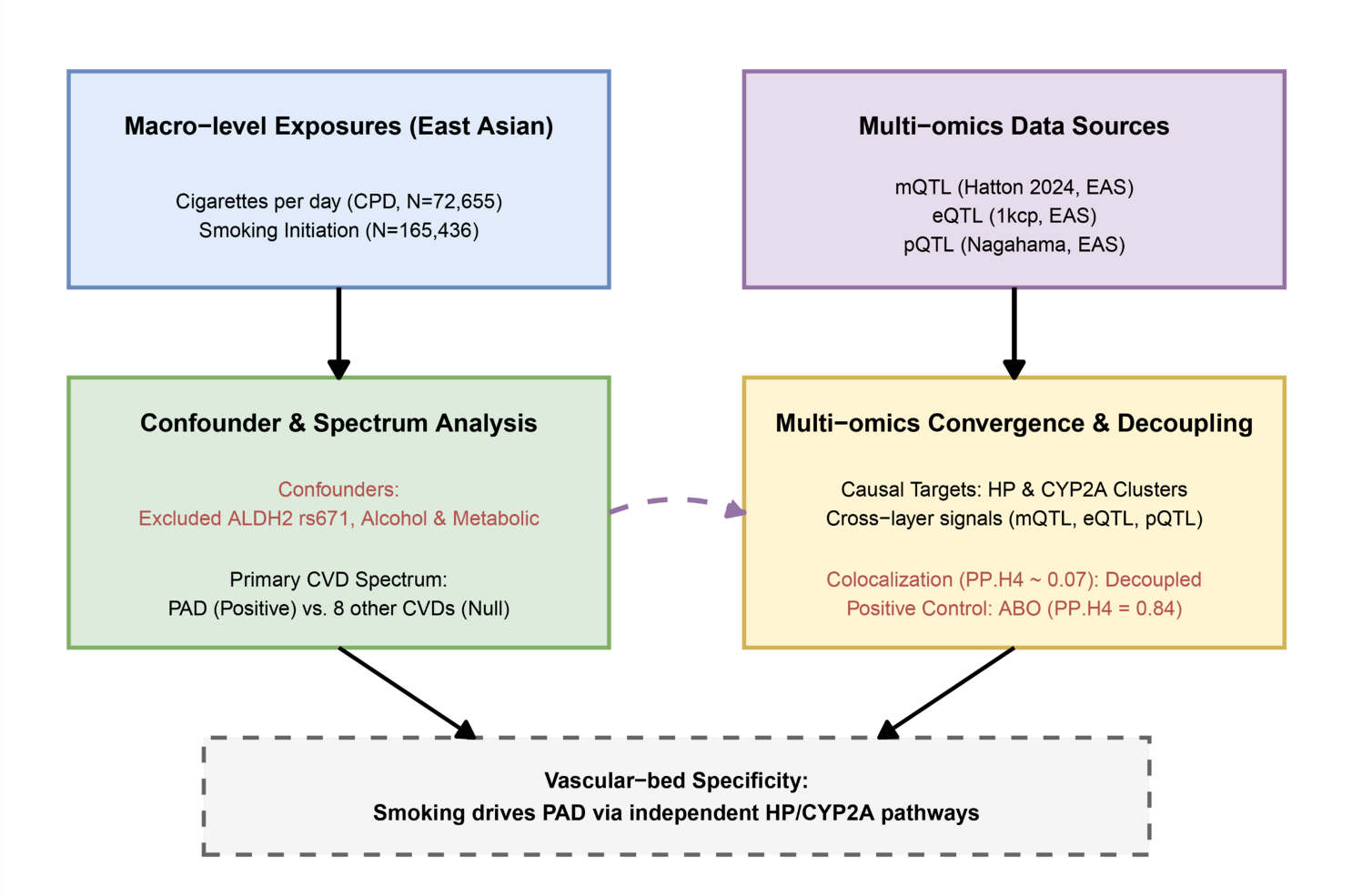

Figure 1

Study design. Two East Asian smoking phenotypes (cigarettes per day and smoking initiation) were used to estimate causal effects on nine cardiovascular outcomes using two-sample Mendelian randomization. Instrumental SNPs were rigorously screened against behavioral confounders, with the highly pleiotropic *ALDH2* rs671 variant explicitly excluded. The primary spectrum analysis revealed a highly specific positive association with PAD, while the other eight CVDs were null. Molecular mechanisms were investigated using mQTL-SMR, eQTL-SMR, pQTL MR, Bayesian colocalization, and spatial *cis*-overlap analysis at the *HP* and CYP2A loci.

Black arrows denote the main study flow and causal effects. The purple dashed arrow denotes the cross-talk between the macroscopic spectrum and molecular mechanisms. Box colors indicate macro-level exposures (blue), multi-omics data sources (purple), confounder and spectrum analysis (green), multi-omics convergence and decoupling (yellow), and the final conceptual conclusion (dashed gray). SNP, single nucleotide polymorphism; CPD, cigarettes per day; MR, Mendelian randomization; mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; SMR, summary-data-based Mendelian randomization; PP.H4, posterior probability of a shared causal variant; PAD, peripheral artery disease; CAD, coronary artery disease; HF, heart failure; DVT, deep vein thrombosis; IS, ischemic stroke; AF, atrial fibrillation; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia.

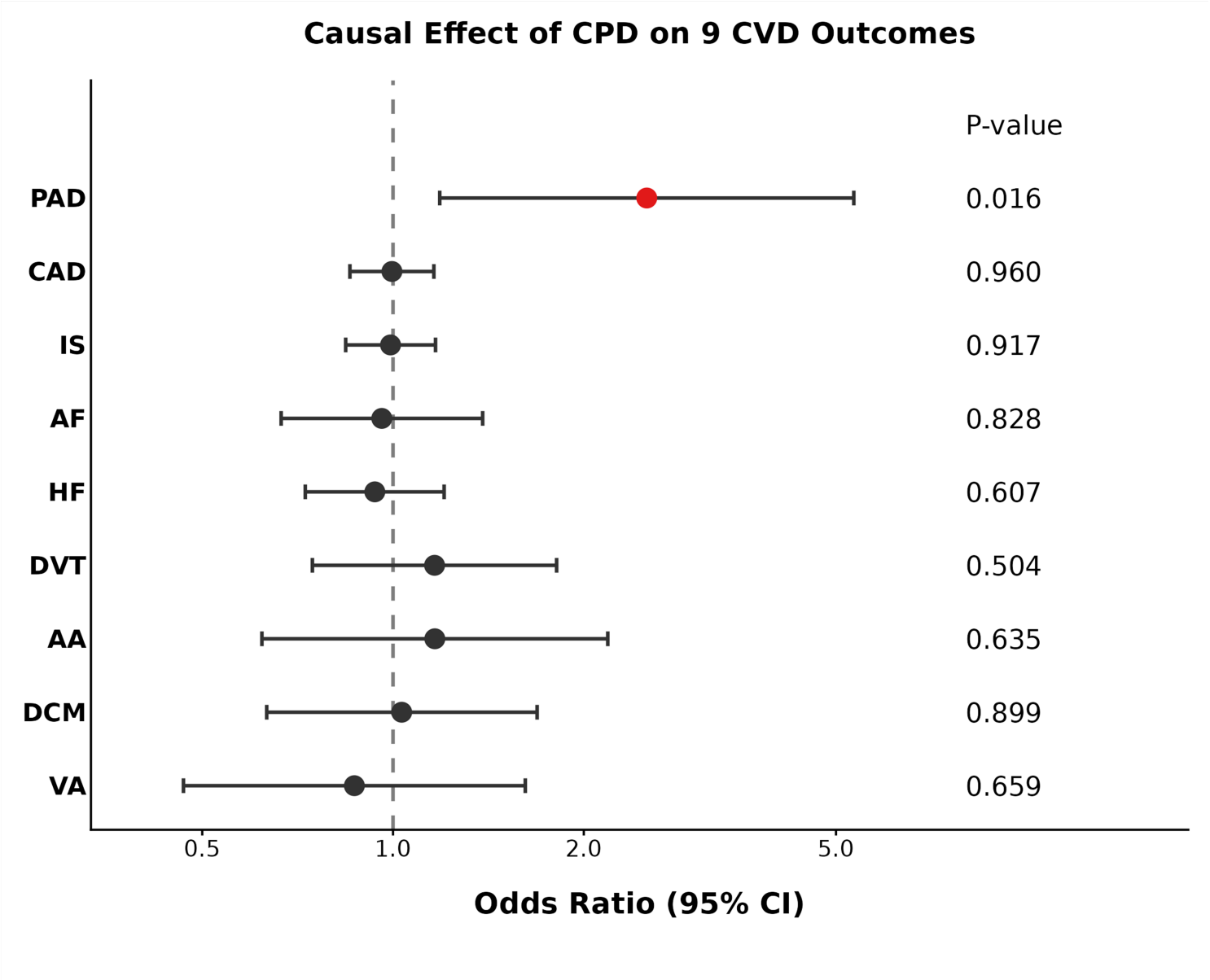

Figure 2

Cardiovascular outcome spectrum of smoking in East Asians. Forest plot displaying the causal effects of genetically proxied cigarettes per day (CPD) on nine cardiovascular outcomes in East Asian populations. The estimates are inverse-variance weighted (IVW) odds ratios (ORs) with 95% confidence intervals (CIs). All analyses were performed after the rigorous exclusion of the highly pleiotropic *ALDH2* rs671 variant. A highly specific vascular-bed tropism was observed: CPD was significantly associated with an increased risk of peripheral artery disease (PAD) (highlighted in red; *P*= 0.016 for the BBJ cohort). A consistent replication was observed in an independent EBI cohort (OR = 2.22, *P* = 0.026; data not shown in the forest plot). In stark contrast, genetically proxied CPD showed null associations with all other eight cardiovascular outcomes (highlighted in black; all *P* > 0.05). The dashed vertical line indicates the null hypothesis (OR = 1). The exact P-values for each outcome are annotated on the right panel. CPD, cigarettes per day; IVW, inverse-variance weighted; OR, odds ratio; CI, confidence interval; PAD, peripheral artery disease; CAD, coronary artery disease; IS, ischemic stroke; AF, atrial fibrillation; HF,

heart failure; DVT, deep vein thrombosis; AA, aortic aneurysm; DCM, dilated cardiomyopathy; VA, ventricular arrhythmia.

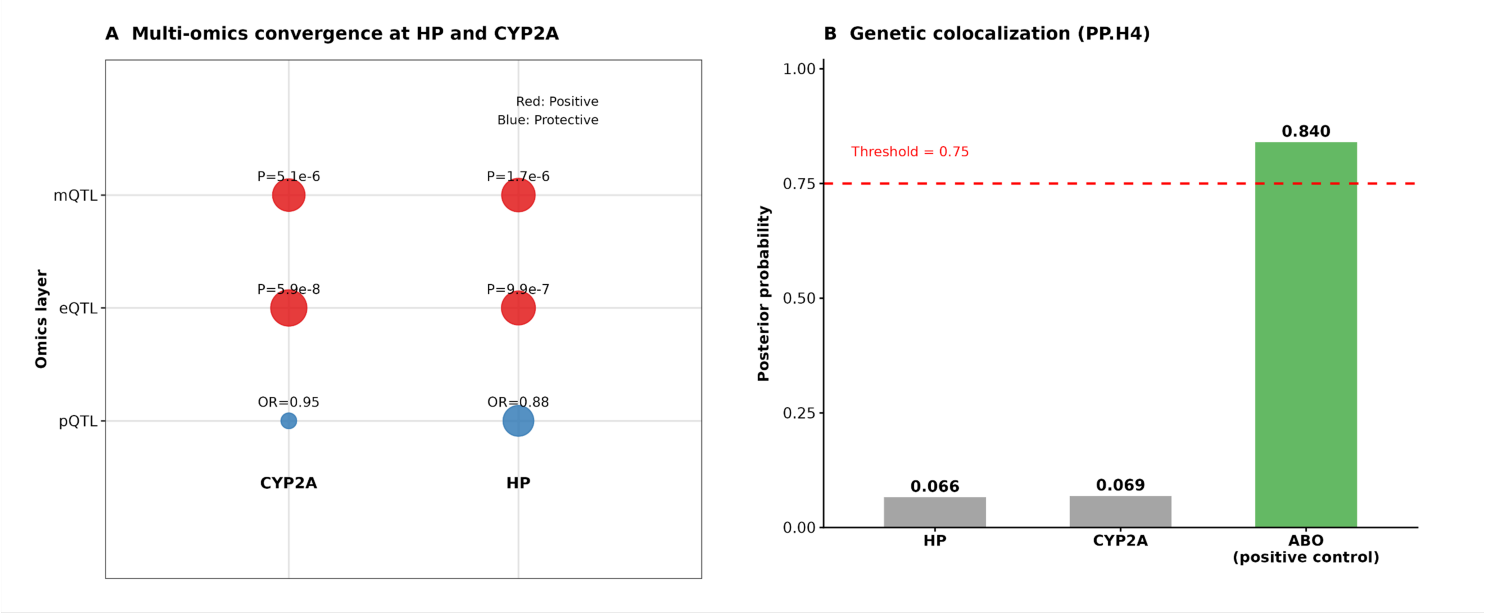

Figure 3

Multi-omics convergence but genetic decoupling at the *HP* and *CYP2A* loci. (A) Bubble plot illustrating the convergent signals across epigenomic (mQTL), transcriptomic (eQTL), and proteomic (pQTL) layers for *HP* and *CYP2A*. Bubble size represents the statistical significance (-log10 P); red indicates a positive effect and blue indicates a protective effect. The exact P-values or odds ratios (ORs) are annotated adjacent to the respective bubbles. (B) Bayesian colocalization results evaluating the posterior probability of a shared causal variant (PP.H4) between mQTL/eQTL signals and CVD outcomes for *HP*, *CYP2A*, and the positive control *ABO*. Despite the multi-layer convergence observed in (A), *HP* (PP.H4 = 0.066) and *CYP2A* (PP.H4 = 0.069) did not support a shared causal variant, whereas the positive control *ABO* showed strong colocalization (PP.H4 = 0.840). The dashed red line indicates the threshold for strong colocalization (PP.H4 = 0.75). These findings define a multi-layer decoupling architecture, where biological signals converge at the pathway level but are driven by distinct genetic variants. mQTL, DNA methylation quantitative trait locus; eQTL, expression quantitative trait locus; pQTL, protein quantitative trait locus; PP.H4, posterior probability of a shared causal variant; OR, odds ratio; CVD, cardiovascular disease.

Supplementary Files

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

- SupplementaryTable.xlsx
- SupplementaryFigureS1.tif