Journal of the American Heart Association

ORIGINAL RESEARCH

Mendelian Randomization Study With Clinical Follow-Up Links Metabolites to Risk and Severity of Pulmonary Arterial Hypertension

Elham Alhathli ⓘ, MSc; Thomas Julian ⓘ, MD; Zain Ul Abideen Girach; A. A. Roger Thompson ⓘ, MBChB, PhD; Christopher Rhodes ⓘ, PhD; Stefan Gräf ⓘ, PhD; Niamh Errington ⓘ, PhD; Martin R. Wilkins ⓘ, MD; Allan Lawrie ⓘ, BSc, PhD; Dennis Wang ⓘ, PhD; Johnathan Cooper-Knock ⓘ, BMBCh, PhD

BACKGROUND: Pulmonary arterial hypertension (PAH) exhibits phenotypic heterogeneity and variable response to therapy. The metabolome has been implicated in the pathogenesis of PAH, but previous works have lacked power to implicate specific metabolites. Mendelian randomization (MR) is a method for causal inference between exposures and outcomes.

METHODS AND RESULTS: Using genome-wide association study summary statistics, we implemented MR analysis to test for potential causal relationships between serum concentration of 575 metabolites and PAH. Five metabolites were causally associated with the risk of PAH after multiple testing correction. Next, we measured serum concentration of candidate metabolites in an independent clinical cohort of 449 patients with PAH to check whether metabolite concentrations are correlated with markers of disease severity. Of the 5 candidates nominated by our MR work, serine was negatively associated and homostachydrine was positively associated with clinical severity of PAH via direct measurement in this independent clinical cohort. Finally we used conditional and orthogonal approaches to explore the biology underlying our lead metabolites. Rare variant burden testing was carried out using whole exome sequencing data from 578 PAH cases and 361 675 controls. Multivariable MR is an extension of MR that uses a single set of instrumental single-nucleotide polymorphisms to measure multiple exposures; multivariable MR is used to determine interdependence between the effects of different exposures on a single outcome. Rare variant analysis demonstrated that loss-of-function mutations within activating transcription factor 4, a transcription factor responsible for upregulation of serine synthesis under conditions of serine starvation, are associated with higher risk for PAH. Homostachydrine is a xenobiotic metabolite that is structurally related to L-proline betaine, which has previously been linked to modulation of inflammation and tissue remodeling in PAH. Our multivariable MR analysis suggests that the effect of L-proline betaine is actually mediated indirectly via homostachydrine.

CONCLUSIONS: Our data present a method for study of the metabolome in the context of PAH, and suggests several candidates for further evaluation and translational research.

Key Words: homostachydrine ■ Mendelian randomization ■ metabolome ■ pulmonary arterial hypertension ■ serine

See Editorial by Provencher et al.

RESEARCH PERSPECTIVE

What Is New?

- Mendelian randomization screen of the metabolome from genome-wide association studies totaling >100 000 participants concludes that serum serine is protective against pulmonary arterial hypertension (PAH), and serum homostachydrine increases risk of PAH.
- To validate these findings, we directly measured plasma serine and homostachydrine in an independent cohort of 449 patients with PAH; strikingly, both metabolites were significantly associated with measures of clinical severity with a consistent direction of effect, and homostachydrine had a significant effect on PAH survival.

What Question Should Be Addressed Next?

- Follow-up analysis suggest that both serine and homostachydrine may be acting via modulation of energy metabolism, but this will require further evaluation.
- Our findings could lead to guidance cautioning against the consumption of certain food products containing homostachydrine, or increasing serine consumption, for patients with PAH.

Nonstandard Abbreviations and Acronyms

ATF4	activating transcription factor 4
BH	Benjamini–Hochberg
FDR	false discovery rate
IVW	inverse-variance weighted
LOF	loss of function
MR	Mendelian randomization
MVMR	multivariable Mendelian randomization
PAH	pulmonary arterial hypertension
REVEAL	Registry to Evaluate Early and Long-Term PAH Disease Management

Pulmonary arterial hypertension (PAH), formerly known as primary pulmonary hypertension, is a rare life-threatening disorder of the pulmonary circulation. It is an archetypal complex disease caused by the interaction of risk genotypes with the environment.[1] Significant advances have been made to understand genetic and environmental risk factors for PAH,[2] but much is still unknown, which limits efforts to develop effective treatments. Certain genetic susceptibility genes have been identified, but it is currently not possible to predict who will be affected or when the disease will occur.

The metabolome is an integrator of genetic and environmental factors; it reflects the state of tissues and cells and can alter their function. As such, the metabolome is a good candidate to explore the biological mechanisms underpinning PAH, and indeed there is evidence for the role of the metabolome in the risk and severity of PAH. For example, Rhodes et al[3] examined 686 biological metabolites in 365 individuals with idiopathic/heritable PAH, 121 healthy controls, and 139 disease controls using ultraperformance liquid chromatography–mass spectrometry. Sets of 20 to 50 metabolites distinguished patients with PAH from controls and disease controls, and correlated with disease severity as measured by survival. Expanding this investigation to patients with chronic thromboembolic pulmonary hypertension demonstrated comparable metabolic abnormalities despite their different underlying biological dysfunction; interestingly, these patients also experienced considerable restoration of their metabolic profiles after therapeutic pulmonary thromboendarterectomy surgery.[4]

Mendelian randomization (MR) is an example of a method for causal inference whereby a significant relationship implies a causal effect of the exposure of interest on the disease trait, providing a series of assumptions are met.[5,6] First, the instrumental single-nucleotide polymorphisms (SNPs) must be robustly associated with the exposure. Second, the SNPs must be associated with the disease risk outcome only through the exposure of interest, rather than affecting the outcome via multiple pathways or exposures. Third, the SNPs must not be associated with confounders. In practice, it is often difficult to be sure that all assumptions are met, and in particular it is difficult to exclude the influence of confounders that are genetically correlated with both the exposure and the outcome. As a result, follow-up orthogonal analyses, particularly those including direct measurement of nominated exposures, are a useful means of supporting an MR result.

Genetic variation is fixed at conception, and therefore SNPs, which form the instruments used in MR, can be considered upstream of adult-onset disease[7] such as PAH. We propose that previous studies of the metabolome in PAH have suffered from limited power because of relatively small sample size for the effect sizes they are trying to measure; as an example, the study described above highlighted a panel of metabolites[3] but was underpowered to implicate individual metabolites in PAH pathogenesis. This is largely a result of a requirement for direct measurement of metabolite concentrations in serum from a large number of patients with PAH. By using MR we overcome this key bottleneck by using genetic correlates to infer metabolite concentrations within PAH patients, thereby negating the requirement for direct measurement. By using different populations

to evaluate exposure and outcome, so-called 2-sample MR, can achieve larger sample sizes without the need to perform all measurements in the same sample set.[5] In the same way, MR also alleviates any bias in the selection of subjects, which often hampers observational studies.[5] We have previously used MR to perform a search for metabolites linked to disease, and here we aimed to apply the same methodology to PAH. We have supported our findings through a series of conditional and orthogonal approaches including direct measurement of serum metabolites in a clinical cohort of patients with PAH. The findings of our study are summarized in Figure 1. Our work opens important new avenues in the search for an understanding of the biological mechanisms underpinning PAH.

METHODS

The authors declare that all supporting data are available within the article and its online supplementary files. Data sets used in our analysis are cited below including reference to ethical approval, but no new institutional review board approval or informed consent was required for this article.

Exposure and Outcome Genome-Wide Association Study

We used 2-sample MR to determine whether metabolite exposures are causally linked to the risk of PAH in a FinnGen PAH genome-wide association study (GWAS) of 125 cases and 162 837 controls (https://www.finngen.fi/en). A more recent overlapping FinnGen PAH GWAS of 208 cases and 243 756 controls was used to confirm metabolites, which passed Benjamini–Hochberg (BH) false discovery rate (FDR) multiple testing correction in the first GWAS.[8]

We performed an analysis of the effect of 575 serum metabolites on PAH risk using publicly available GWAS studies of the serum metabolome.[9,10] The 2 source studies differed in sample size and methodology:
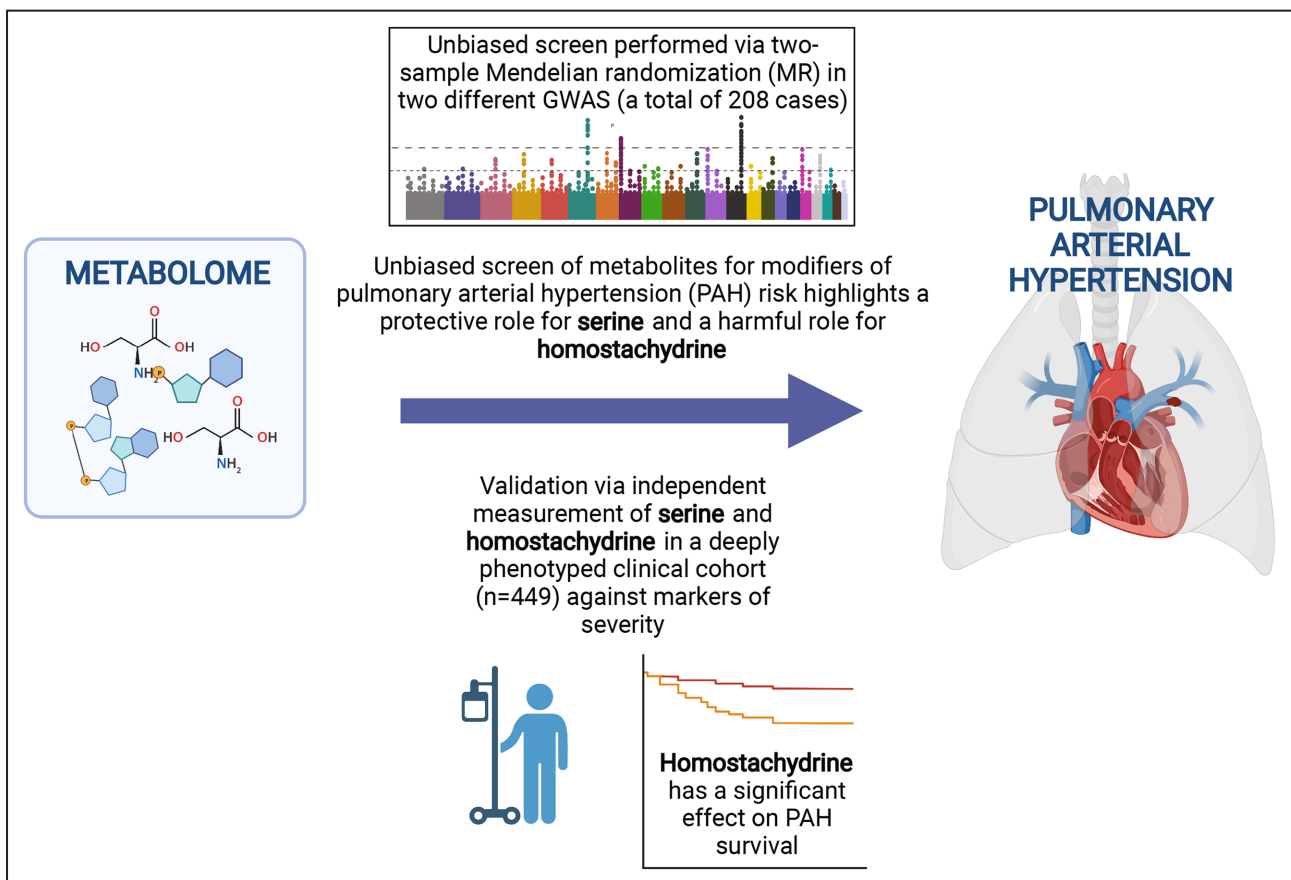
Figure 1. Overview of the metabolome-wide screen for metabolites that modify risk of pulmonary arterial hypertension (PAH).
We performed a 2-sample Mendelian randomization (MR) screen of 575 metabolites to identify those causally linked to PAH in a FinnGenn cohort. Metabolites that the passed Benjamini–Hochberg false discovery rate multiple testing correction were further evaluated by robust MR measures, sensitivity analysis, and in a validation FinnGenn PAH genome-wide association study (GWAS). Finally, we directly measured these metabolites in an independent UK PAH cohort of 449 patients and demonstrated a correlation with disease severity.

Kettunen et al[10] employed nuclear magnetic resonance spectroscopy to quantify metabolites in serum samples from up to 24925 individuals, whereas Shin et al[9] employed liquid-phase and gas chromatography coupled with tandem mass spectroscopy in serum from 7824 individuals. The sample population for both studies was almost exclusively European; both studies included ≈2000 individuals from the Cooperative Health Research in the Region of Augsburg cohort. The different methodologies employed were complementary: Nuclear magnetic resonance is more easily applied in a high-throughput manner, but sensitivity is lower than mass spectroscopy-based quantification.[11] A total of 16 metabolites (4% of the total number) were measured in both source studies not including any of our 5 top candidates.

As a further confirmation, we used a new GWAS study of the metabolome[12] in 50000 Europeans from the INTERVAL study (https://www.intervalstudy.org.uk).[13] Plasma concentrations of serine were measured using the Metabolon HD4 Discovery platform (Metabolon, Durham, NC), which is a liquid chromatography coupled with mass spectroscopy method similar to that used in the Shin et al study.[9]

We performed a second MR screen of the effect of 150 immune traits on PAH risk. Here, we used a separate publicly available GWAS of blood concentration of a set of human immune system traits in 669 female twins from the United Kingdom.[14] Immune cells were quantified via specific antibodies and flow cytometry.

Two-Sample MR

Genetic instruments were selected on the basis of a P value threshold of either $P=5e-06$, $P=5e-07$, or $P=5e-08$; the lowest threshold was chosen to ensure >10 instrumental SNPs. When the cutoff is too low, informative instruments will be lost, but when it is too high, noninformative instruments will be introduced and instrument pleiotropy is more likely to occur.[15] For each metabolite and immune trait studied in our MR screens, we used multiplicative random-effects inverse-variance weighted (IVW) estimate for significance testing, as this method has the greatest statistical power,[16] and it allows us to make accurate causal inferences for >4 instrumental SNPs under the assumption of limited balanced pleiotropy. IVW estimates were FDR[17] corrected for multiple testing by the BH method[17] assuming that the condition of "positive regression dependency on each one from a subset"[18] is met; a threshold of FDR<0.05 was applied to identify metabolites that are causally associated with PAH. In practice, it is difficult to be sure that the condition of positive regression dependency on each one from a subset is met, and therefore we also applied both a Bonferroni and a Benjamini–Yekutieli correction for multiple testing, which assume either no interdependence or are agnostic to the nature of dependence between statistical tests.[18] Independent SNPs were clumped using a stringent cutoff of $R<0.001$ within a 10000-kb window in the European reference panel. For clumped SNPs in linkage disequilibrium, the SNP with the lowest P value was retained. When an exposure SNP was unavailable in the outcome data set, a proxy with a high degree of linkage disequilibrium ($R^2>0.9$) within a European reference population was identified. SNP effects on outcomes and exposures were harmonized so that the beta values are based on the same alleles. To reduce the risk of errors due to strand issues, palindromic alleles with minor allele frequency >0.42 were excluded from the analysis.[15]

Robust Tests and Sensitivity Analysis

To increase confidence in the IVW results from our unbiased screening, we performed a series of robust MR measures and sensitivity analyses. We used an F-statistic to measure the strength of the association between instrumental SNPs and the exposure of interest. An F-statistic >10 indicates that an SNP-derived estimate has a bias of <10% of its intragroup variability and signifies an acceptable instrument. Pleiotropy occurs between SNPs for which the difference in effect size for the exposure is not proportional to the difference in effect size for the outcome, and is usually due to a violation of one of the key assumptions underlying MR: the assumption that instrumental SNPs should be associated with the outcome only through the exposure.[5] To account for pleiotropy, we removed SNPs for which the P value for the association with the outcome was lower than for the association with the exposure of interest. The MR–Egger intercept test was also used to identify directional horizontal pleiotropy. To avoid false positives, we limited our analysis to exposures with >10 instrumental SNPs and used a quantile–quantile plot to determine whether the median P value was significantly inflated. A test with too few SNPs gives excessive weight to each individual SNP.[19] As IVW estimates are vulnerable to pleiotropic SNPs, we used Cochran's Q test ($P>0.05$) as a sensitivity measure to detect heterogeneity indicating pleiotropy. Moreover, radial MR[20] was used to remove statistically significant outlier SNPs. The I^2 statistic was used to measure the heterogeneity between variant-specific causal estimates, with a low I^2 indicating bias toward the null hypothesis.[21] A leave-one-out analysis was applied to identify results in which ≥1 SNPs exert a disproportionate effect. TwoSampleMR (version 0.5.6), Mendelian Randomization (version 0.5.1) and RadialMR (version 1.0) R packages were used for all

MR analyses (R Foundation for Statistical Computing, Vienna, Austria).

Clinical Validation

We evaluated the serum metabolites that were nominated by our 2-sample MR study using measurements of the same molecules within plasma samples from the UK PAH cohort[3,4] including 449 PAH patients. In this cohort, metabolic profiling was performed using ultra-performance liquid chromatography–mass spectrometry.[3,4] We examined the concentration of serine and homostachydrine and tested for correlation between identified candidates and clinical measures of PAH severity including survival. In all analyses, we adjusted for other covariates such as age, sex, body mass index (BMI), population ancestry (Asian, Black, White, not stated, and other) and geographic location of the sample collection.

Clinical measures of severity in the UK PAH cohort included World Health Organization functional class; NT- proBNP (N-terminal pro-B-type natriuretic peptide), which correlates with myocardial stress and provides prognostic information; 6-minute walking distance; and a series of measurements taken at right heart catheterization including pulmonary arterial wedge pressure, pulmonary vascular resistance, right atrial pressure, cardiac index, cardiac output, and venous oxygen saturation. We also included the percentage of predicted forced vital capacity as a measure of lung function. The Registry to Evaluate Early and Long-Term PAH Disease Management (REVEAL) is an aggregate score used for prognostication in PAH.[22] Many of these measures are not specific to PAH but have been associated with clinical outcomes including survival. In the analysis of clinical measurements we used linear/logistic regression adjusted for covariates including age, sex, BMI, ethnicity, and geographic location of the sample collection. In the analysis of survival, we applied Cox regression adjusted for age, sex, BMI, ethnicity, and geographic location of the sample collection. Kaplan–Meier curves were plotted by dividing patients with PAH on the basis of their serum metabolite concentrations into equal-sized percentiles.

Multivariable MR

Multivariable MR (MVMR) is a conditional MR analysis technique[23] that we used to determine whether the protective effect of serine was mediated via regulation of immune cell numbers and whether the toxic effect of homostachydrine was detectable via coffee consumption or mediated via L-proline betaine. The P value cutoffs used to select instrumental SNPs for each exposure were chosen so as to achieve adequate instrument strength for both exposures (conditional F-statistic >10 for each exposure). Reported results showed no evidence of instrument heterogeneity (Cochran's Q test $P>0.05$). Exposures were derived from independent cohorts and therefore a correction for the covariance between the effect of the genetic variants on each exposure was not necessary. MVMR was implemented using the MVMR (version 0.3) and MendelianRandomization (version 0.5.1) R packages.

Rare Variant Analysis

MR is based on common genetic variants (minor allele frequency $>1\%$), which are largely orthogonal to large-effect rare missense variants. We therefore used the AstraZeneca PheWAS portal[24] to analyze the Union Primary Pulmonary Hypertension data set (361 675 controls and 578 PAH cases) to determine whether rare variants within enzymes responsible for metabolism of our candidate metabolites are also associated with risk of PAH. Identified rare variants with a common biological effect were collapsed into a single Fisher's exact 2-sided test to determine whether the burden of variants is different in PAH cases and controls. We focused on loss-of-function (LOF) variants including protein-truncating variants, missense variants, and in-frame insertions and deletions that are damaging to the protein (Rare Exome Variant Ensemble Learner score ≥ 0.25)[25] that were rare (minor allele frequency <0.001 if protein-truncating or minor allele frequency <0.0005 if non–protein-truncating, in both UK Biobank and GnomAD[26]); in addition, we considered that variants were high-quality variant calls on the basis of coverage, mapping quality, genotype quality, and Hardy–Weinberg equilibrium (ptvraredmg as described in Wang et al[24]).

RESULTS

Metabolome-Wide MR Causally Associates 5 Metabolites With PAH Risk

We hypothesized that serum metabolites may act upon pulmonary tissue to modify the biological processes leading to PAH. We used MR to perform a study of the entire set of serum metabolites to determine which metabolites are causally related to PAH. To achieve this, we obtained genetic variants associated with the serum concentration of 575 metabolites in 7284 individuals.[9,10] Next, we applied a series of MR tests in which the exposure was formed from metabolite-associated genetic variants, and genetic liability to PAH was the outcome measure. Genetic associations with PAH were determined within the
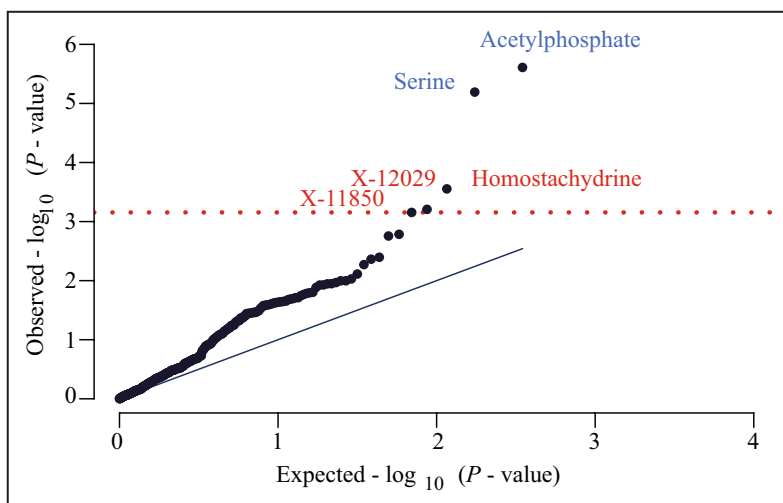
Figure 2. Five metabolites were significantly associated with risk of pulmonary arterial hypertension (PAH) by Mendelian randomization.
Quantile–quantile plot of inverse-variance weighted P values demonstrates that 5 metabolites passed a Benjamini–Hochberg false discovery rate multiple testing correction (red line). Blue text represents a protective association, and red text represents a harmful association.

FinnGen cohort consisting of 125 patients with PAH and 162 837 controls.[8]

The complete results of our full metabolome screen are shown in Table S1. After BH FDR correction of the IVW estimate P value, 5 metabolites were significantly associated with risk of PAH. Two metabolites were protective against PAH: acetylphosphate (IVW P=2.46e-06, β=−16.06, SE=3.4) and serine (IVW P=6.43e-06, β=−6.7, SE=1.49; Figures 2 and 3A and 3B). Three metabolites increased the risk of PAH: homostachydrine (IVW P=0.0003, β=1.09, SE=0.3), X-12029 (IVW P=0.0006, β=8.9, SE=2.6), and X-11850 (IVW P=0.0007, β=0.95, SE=0.28) (Figures 2 and 3C through 3E). Based on the sensitivity tests performed, none of these tests was invalidated by instrument pleiotropy or weak instruments (Table 1). Acetylphosphate was also significantly associated with PAH risk by the weighted median estimate (P=0.04, β=−13.7), and X-11850 was significant by the weighted mode estimate (P=0.02, β=1.93). Only homostachydrine was significant in multiple robust MR estimates (weighted median P=0.03, β=1.28; MR Egger P=0.003, β=1.38; weighted mode P=0.006, β=1.26, Table 1), which lends significant support to the validity of this metabolite as a mediator of toxicity leading to PAH.

There is some uncertainty as to which multiple testing correction to apply to the IVW estimate P value given the correlated nature of serum metabolite concentrations. The BH correction assumes the condition of "positive regression dependency on each one from a subset,"[18] but it is difficult to be certain that this condition is fulfilled. The Benjamini–Yekutieli correction is agnostic to the nature of dependence between tests,[18] and a Bonferroni correction assumes no interdependence. Under either the Benjamini–Yekutieli or Bonferroni multiple testing correction, then, only serum concentrations of acetylphosphate and serine are significantly associated with risk of PAH at a threshold corrected P<0.05.

Validation in Additional GWAS Supports the Validity of Identified Metabolites

To validate the set of metabolites we associated with risk of PAH, we obtained a different larger FinnGen GWAS consisting of 208 PAH cases and 243 756 controls[8] that overlaps with our original sample set. Repeating the MR tests for the association of serum metabolite concentration with risk of PAH in this data set confirmed our previous results for serine, acetylphosphate, and homostachydrine but not for X-12029 and X-11850. In this analysis, both acetylphosphate (IVW P=0.0026, β=−11.3, SE=3.76; Figure S1A) and serine (IVW P=0.0004, β=−5.23, SE=1.49; Figure S1B) were significantly protective against PAH, while homostachydrine increased risk of PAH (IVW P=0.004, β=0.67, SE=0.24; Figure S1C). X-12029 (IVW P=0.1, β=3.6, SE=2.2) and X-11850 (IVW P=0.056, β=0.5 SE=0.26) were not significantly associated with the increased risk of PAH but maintained the same direction of effect. Serum serine was also significantly protective against PAH when measured in a different GWAS of plasma serine
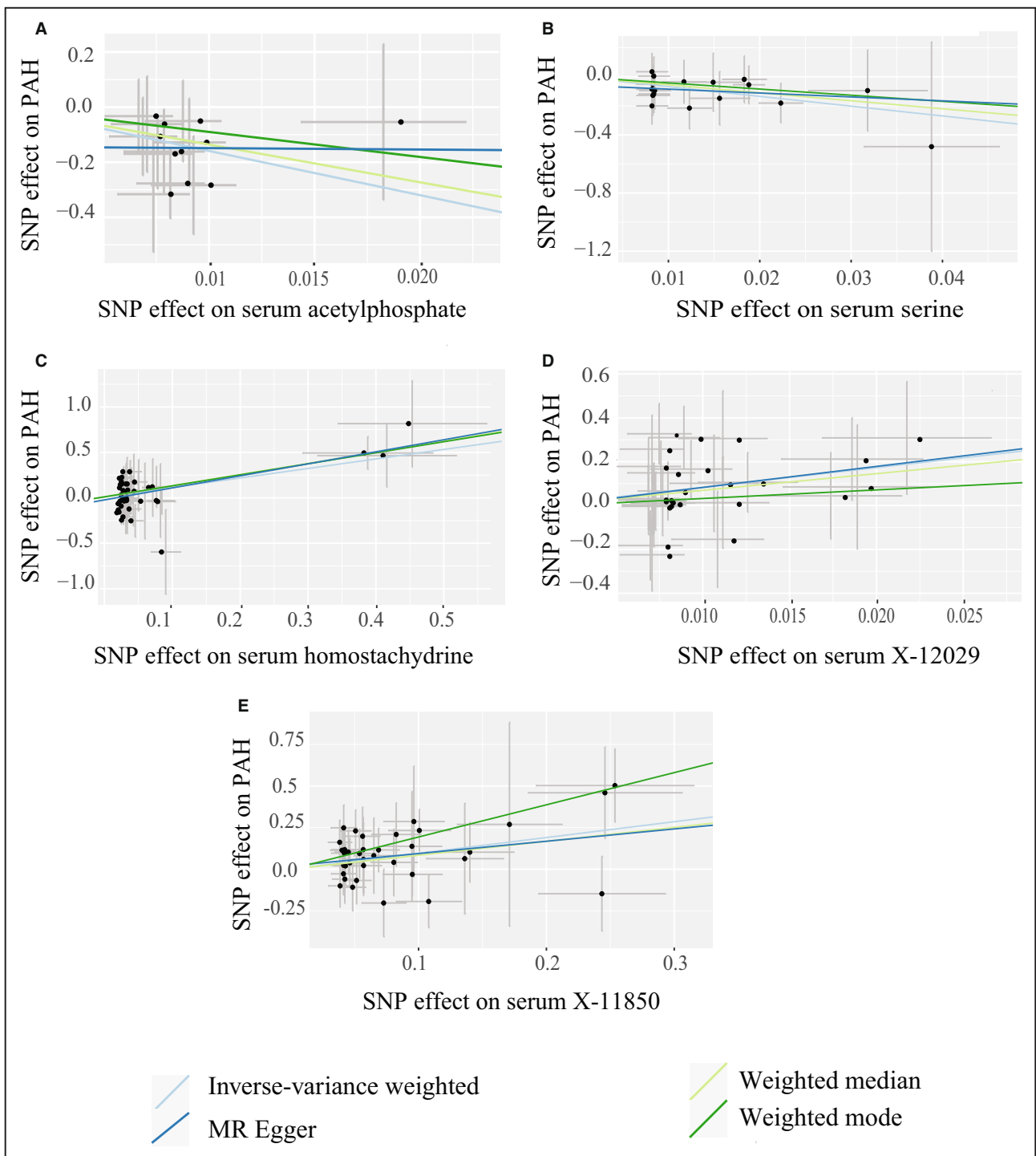
Figure 3. Two-sample Mendelian randomization (MR) tests for a causal effect of metabolites on risk of pulmonary arterial hypertension (PAH).
Scatter plots demonstrate a significant association of serum metabolite concentrations with risk of PAH including in robust MR tests for acetylphosphate (**A**), serine (**B**), homostachydrine (**C**), X-12029 (**D**), and X-11850 (**E**). Each point represents the effect size (β) and standard errors for each single-nucleotide polymorphism (SNP)–outcome relationship.

concentration[12] using the original FinnGen outcome GWAS and using the same instrumental SNPs (IVW $P=0.001$, β=−1.08, SE=0.339; Figure S1D). After these analyses X-12029 and X-11850 were excluded from further analysis, as these results were deemed to be nonreproducible.

Table 1. Robust MR Measures for the Effect of Significant Metabolites on Risk of PAH

Test	Serine	Acetylphosphate	Homostachydrine	X-12029	X-11850
IVW P value	6.43e-6*	2.46e-06*	0.0003*	0.0006*	0.0007*
IVW beta	−6.7*	−16.06*	1.09*	8.9*	0.95*
Weighted median P value	0.13	0.04*	0.03*	0.07	0.1
Weighted median β	−5.5	−13.7*	1.28*	7.6	0.84
Egger P value	0.69	0.98	0.003*	0.33	0.23
Egger β	−2.68	−0.51	1.38*	9.58	0.74
Weighted mode P value	0.36	0.4	0.006*	0.5	0.02*
Weighted mode β	−4.2	−9.08	1.26*	3.9	1.93*
Mean F test	35	25	18	22	18
IVW Cochran's Q test P value	1	0.93	0.8	0.8	0.9
Radial MR outlier SNPs	0	0	2	1	4
Egger intercept test	0.51	0.53	0.3	0.9	0.7
I^2	0.97	0.96	0.94	0.96	0.95
Number of SNPs LOO>0.05	0	0	0	0	0
Total number of SNPs	17	11	42	25	36
Number of pleiotropic SNPs	0	0	0	0	0

IVW indicates inverse-variance weighted; MR, Mendelian randomization; PAH, pulmonary arterial hypertension; and SNP, single-nucleotide polymorphism.
*Statistically significant results.

Direct Measurement of Serine and Homostachydrine Supports a Link to PAH Clinical Severity

MR is a powerful tool that leverages large sample sets. However, MR relies on genetic correlates of measured exposures and outcomes, which is not equivalent to a direct measurement. We therefore sought to further validate our results by direct measurement of metabolites in a set of clinically profiled PAH patients from the PAH UK cohort of 449 patients with PAH including 260 patients with idiopathic PAH.[3,4] Unlike our MR analysis, this experiment cannot exclude reverse causality; however, where direct measurement supports our MR analyses, we conclude that there is convergent evidence that a particular metabolite modifies biological processes leading to PAH.

In this clinical data set, the plasma concentration of both serine and homostachydrine were significantly correlated with measures of PAH severity after adjusting for age, sex, BMI, population ancestry, and geographic location. Serum serine was positively correlated with measures of cardiac function (Table 2; Figure 4A): cardiac index (linear regression, coefficient=0.56, P=0.005) and cardiac output (coefficient=0.73, P=0.027) and the correlation with cardiac index remained significant after BH FDR correction. Serum homostachydrine was positively associated with NT-proBNP (coefficient=0.035, P=0.05) and pulmonary arterial wedge pressure (coefficient=0.069, P=0.047) and negatively correlated with venous oxygen saturation (coefficient=−0.21, P=0.025) (Table 2; Figure 4A). In each case, and even for results that did not reach statistical significance, the direction of effect was consistent with our MR analyses; that is, serine was protective and homostachydrine was harmful. In addition, increased serum homostachydrine was associated with a reduction in survival (Cox regression, coefficient=6.76e-02, P=0.03) after adjusting for age, sex, BMI, population ancestry, and geographic location (Figure 4B). Serum serine concentration was correlated with an increase in survival,

Table 2. Linear and Logistic Regression of Biomarkers of PAH Severity Against Plasma Concentration of Metabolites

Variable	Serine		Homostachydrine	
	Coefficient	P value	Coefficient	P value
WHO-FC	−3.9	0.72	1.66	0.56
NT-proBNP	−0.5	0.22	0.035	0.051
6MWD	−0.5	0.22	−2.15	0.07
mPAP	−2	0.45	0.19	0.09
PAWP	−0.05	0.94	0.069	0.047*
SvO$_2$	2.5	0.27	−0.21	0.025*
Cardiac index	0.56	0.005*	−0.01	0.19
Cardiac output	0.73	0.027*	−0.017	0.2
PVRdynes	−72	0.48	8.27	0.049*
REVEAL scores	−4.9	0.71	0.09	0.72

All tests were adjusted for age, sex, body mass index, geographic location, and population ancestry. 6MWD indicates 6-minute walking distance; mPAP, mean pulmonary arterial pressure; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PAH, pulmonary arterial hypertension; PAWP, pulmonary arterial wedge pressure; PVRdynes, pulmonary vascular resistance in dynes; SvO$_2$, venous oxygen saturation; and WHO-FC, World Health Organization functional class.
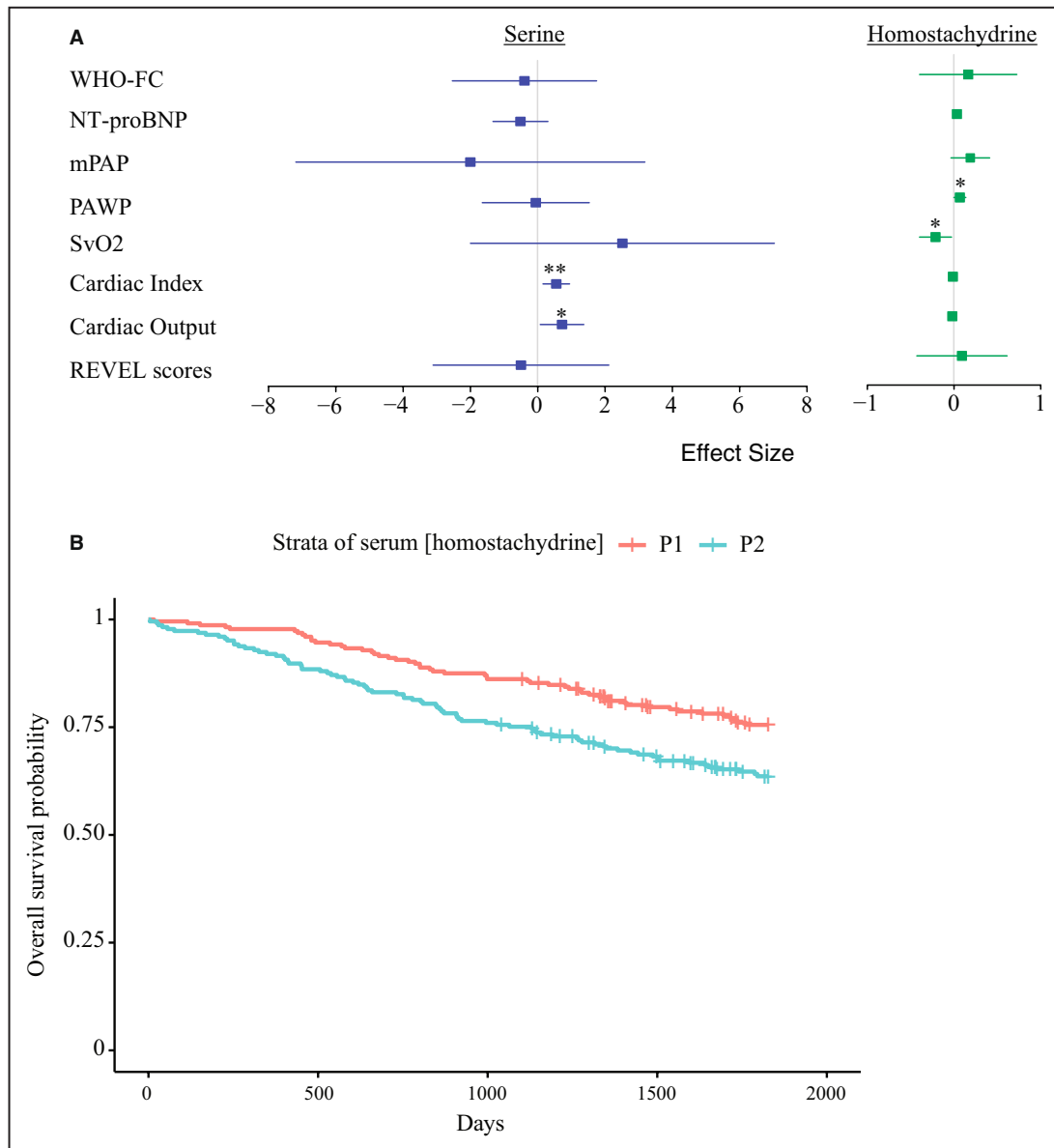
*Statistically significant results.

Figure 4. Serum metabolite levels are correlated with clinical severity in the UK pulmonary arterial hypertension (PAH) cohort.

A, Forest plot to demonstrate linear and logistic regression of biomarkers of PAH severity against plasma concentration of metabolites. Plots indicate mean and 95% CIs. Effect size refers to regression coefficient. All tests were adjusted for age, sex, body mass index, geographic location and population ancestry. ******$P<0.01$, *****$P<0.05$. **B**, Higher serum homostachydrine is associated with a significant reduction in survival (censored at 5 years) in the UK PAH cohort (Cox regression, cor=6.76e-02, $P=0.03$) after adjusting for age, sex, body mass index, geographic site, and ethnicity. A Kaplan–Meier curve is shown where patients with PAH are divided into subgroups based on the following percentiles of their serum concentration of homostachydrine: P1: <50%, P2: >50%. mPAP indicates mean pulmonary arterial pressure; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PAWP, pulmonary arterial wedge pressure; REVEL, Rare Exome Variant Ensemble Learner; SvO$_2$, venous oxygen saturation; and WHO-FC, World Health Organization functional class.

but after adjusting for covariates, the results were not significant (Figure S2). Direct measurement of acetylphosphate was not possible, and therefore there was no clinical evidence to support or refute the role of acetylphosphate in PAH; to be cautious, we excluded acetylphosphate from further analysis.

Homostachydrine Association With PAH Is Independent of Coffee Consumption

Homostachydrine is a xenobiotic metabolite (also known as pipecolic acid betaine) that is highly homologous to L-proline betaine (also called stachydrine).

Initially, it was discovered in leaves of the *Medicago* (alfalfa) and *Achillea* genera. Homostachydrine is also present within frequently used robusta coffee beans.[27] Indeed, serum homostachydrine has been proposed as a marker of coffee consumption.[28] We have demonstrated that serum homostachydrine is causally associated with PAH risk and severity; however, it may not be directly related to PAH. Instead, it may act via an intermediate that is also present within coffee. To investigate this possibility, we performed an MR test for the effect of coffee consumption on risk of PAH.

We used a GWAS for questionnaire-reported coffee consumption consisting of 428 860 individuals.[29] There was no observed effect of coffee consumption on risk of PAH (IVW *P*=0.87, β=0.27). In addition, we performed a conditional analysis of the effect of homostachydrine on risk of PAH independent of coffee consumption using MVMR. MVMR revealed that the association of homostachydrine with PAH was not invalidated when conditioned on coffee consumption (β=0.97, *P*=0.006). These data suggest that homostachydrine is acting directly and is not serving as a proxy for another component of coffee, although an important caveat is that questionnaire-reported coffee consumption may be inaccurate.

Effect of L-Proline Betaine Effect on PAH Is Dependant on Homostachydrine

Homostachydrine is structurally similar to L-proline betaine, but it contains a different alkyl chain length and substitution pattern. It is likely that the 2 share a common metabolic pathway and their serum concentrations may be correlated. Interestingly, an increase in serum L-proline betaine has been associated with a poor prognosis in PAH.[30] Our MR analysis supports this conclusion: We found a borderline significant toxic effect of L-proline betaine on PAH (IVW *P*=0.055, β=5.14, SE=2.68). This raised a question as to whether the effects of homostachydrine and L-proline betaine are independent. To investigate this further, we performed an MVMR in which we discovered that the effect of L-proline betaine is removed (β=1.8, *P*=0.39) when we conditioned on homostachydrine; whereas homostachydrine was associated with increased risk of PAH (β=2.04, *P*=0.054) even when conditioned on L-proline betaine. We conclude that the effect of homostachydrine on risk of PAH is independent of the effect of L-proline betaine, and homostachydrine may even account for previously reported observations regarding the effect of L-proline betaine on PAH.

There Is no MR Evidence That the Effect of Serine on PAH Is Mediated via Immunomodulation

We have identified a protective effect of serine on risk and severity of PAH. We interrogated the literature to determine a likely mechanism for this effect. Activated immune cells are rapidly proliferating and are thus dependent on serine to function.[31] Moreover, patients with PAH exhibit immune system abnormalities.[32–34] We hypothesized that the protective effect of serine on PAH may be mediated via immunomodulation.

To investigate this further, we performed a second MR screen of immune traits[14] as determinants of PAH risk (Table S2). After BH FDR multiple testing correction, only 2 traits were significantly associated with risk of PAH: the proportion of Vδ1+ T cells (IVW, *P*=0.00006, β=−1.02) and the proportion of CD1c- myeloid dendritic cells (IVW, *P*=0.00076, β=0.84). Vδ1+ T cells are tissue-resident lymphocytes involved in immune surveillance where they receive antigen presentation from dendritic cells including CD1c- myeloid dendritic cells. Serum serine was not significantly associated with either of these traits (IVW, *P*>0.05). Nineteen immune traits were nominally (IVW *P*<0.05) associated with PAH risk; but only 2 of these immune traits: the concentration of CD8+ T cells (IVW, *P*=0.013, β=0.6), and the concentration of CD3 on double-negative T cells (IVW, *P*=0.042, β=0.9), were also significantly associated with serum concentration of serine (Table S3). CD8+ T cells are cytotoxic, whereas double-negative T cells are associated with immunomodulation; the notable exception here is CD4+ T cells, which suggests that these helper T cells may be relatively less dependent on serine for their function. We applied MVMR to determine whether the effect of serine on PAH was mediated via either of these immune traits; to the contrary, in MVMR the effect of serine on PAH remained significant (*P*<0.05), but both immune traits were nonsignificant when conditioned on serine. Our MR analysis thereby suggests that the effect of serine is independent of immune function and not mediated via CD8 T cells or CD3+ double-negative T cells.

Rare Variant Analysis Supports the Validity of Serine as Protective Against PAH

If serum serine is protective against PAH, it follows that genetic variants that reduce synthesis of endogenous serine might modify the risk of PAH. Activating transcription factor 4 (ATF4) is a transcription factor that promotes the expression of enzymes involved in endogenous synthesis of serine under conditions of

Table 3. LOF Mutations Within the ATF4 Transcription Factor Involved in Control of Serine Biosynthesis Are Associated With Increased Risk of PAH

Gene	P value	Number of cases	Number of controls	OR	Biological role
ATF4	0.047*	3 (0.53%)	498 (0.14%)	3.78	TF that upregulates expression of PSPH, PSAT, PHGDH and SHMT2
PSPH	>0.1	…	…	…	Serine biosynthesis
PSAT	>0.1	…	…	…	
PHGDH	>0.1	…	…	…	
SHMT2	>0.1	…	…	…	Conversion between glycine and serine

This analysis used whole-exome sequencing from 578 patients with PAH and 361 675 controls.[24] Numbers refer to the total number of PAH cases and controls who carried a LOF variant under the ptvraredmg classification.[24] LOF indicates loss-of-function; OR, odds ratio; PAH, pulmonary arterial hypertension; PHGDH, phosphoglycerate dehydrogenase; PSAT, phosphohydroxythreonine aminotransferase; PSPH, phosphoserine phosphatase; SHMT2, serine hydroxymethyltransferase; and TF, transcription factor.

*Statistically significant results.

serine starvation[35] and has previously been linked to a related disease, dilated cardiomyopathy.[36] We performed rare variant genetic burden testing in a cohort of 578 PAH cases compared with 361 675 controls to determine whether patients with PAH carried an excess of LOF mutations within ATF4 or any of the enzymes under its control (Table 3). There was an excess of LOF mutations within ATF4 in PAH cases compared with controls (Fisher's exact test, odds ratio=3.78, P=0.047), which is consistent with our MR data suggesting a protective role for serum serine. There was no evidence for an excess of LOF variants in any of the individual enzymes downstream of ATF4 involved in serine biosynthesis.

DISCUSSION

PAH is an archetypal complex disease determined by gene–environment interactions that is currently incurable. The role of the metabolome in the development of PAH is poorly understood, although groups of metabolites have been observed to associate with PAH risk and severity in previous observational studies.[3] Here, we have applied a metabolome-wide screen to search for causal links between serum metabolites and PAH risk, leveraging the significant statistical power provided by 2-sample MR. As a comparison with the cited study of Rhodes et al,[3] which used 365 patients with PAH and 121 controls, the increase in same size in our 2-sample MR screen (125 patients with PAH and 162 837 controls) achieves equal power after multiple testing correction, to detect a 20% smaller change in concentration of an individual metabolite.[37] We have previously used MR to perform a similar search for metabolites that determine risk for amyotrophic lateral sclerosis,[15] Alzheimer disease,[38] and age-related macular degeneration.[39]

We have linked 3 metabolites to the risk of PAH using MR in 2 different GWASs for PAH: acetylphosphate, serine, and homostachydrine. Serine and homostachydrine are supported by orthogonal analysis via direct measurement of serum metabolites in a clinical cohort of patients with PAH. In our MR results, higher serum serine reduces risk for PAH, and in the clinical cohort, higher serum serine was associated with reduced PAH severity and increased survival. In contrast, in our MR results, higher serum homostachydrine increased risk for PAH, and in the clinical cohort, higher serum homostachydrine was associated with increased PAH severity and reduced survival. The orthogonal nature of the tests we have applied provides significant evidence that we have identified new biological mechanisms underpinning PAH.

Serine biosynthesis has been previously implicated in dilated cardiomyopathy via the transcription factor ATF4[36]; dilated cardiomyopathy and PAH may share underlying pathobiology, particularly related to vascular remodeling.[40] Our rare variant analysis associated LOF mutations within ATF4 with increased risk of PAH, which supports a role for this pathway in PAH. In the study of the effect of serine biosynthesis on dilated cardiomyopathy,[36] positive treatment effects were correlated with correction of mitochondrial respiration chain defects, suggesting a boost to energy production. Our analyses suggest that serine is not linked directly to PAH via immunomodulation and we conclude that a role in energy metabolism is more likely. Promoting endogenous serine synthesis via the activity of ATF4 may be a novel therapeutic strategy for PAH.

Homostachydrine is a xenobiotic metabolite that is exclusively produced in plants including within robusta coffee beans. Homostachydrine has been proposed as a marker of quality for coffee, which is thought to be improved by a higher proportion of arabica coffee beans.[27] Indeed, serum homostachydrine has been proposed as a marker of coffee consumption.[28] However, our analysis suggests that the effect of homostachydrine on PAH is independent of coffee consumption and therefore not likely to be a by-product of some other component of coffee.

Homostachydrine is structurally similar to ʟ-proline betaine, and our MR analysis suggests that ʟ-proline

betaine may be indirectly linked to risk of PAH via homostachydrine. Higher fasting blood concentrations of L-proline betaine have been linked to poor prognosis in PAH, including a rise in NT-proBNP, which is an important biomarker of PAH severity.[30] Unlike homostachydrine, there is some understanding of the specific activities and physiological roles of L-proline betaine, for example, increased concentrations of L-proline betaine may disrupt mitochondrial structure and function leading to a toxic energy deficit.[41] It is possible that this effect on energy metabolism is actually mediated via homostachydrine, and that an energy deficit is responsible for the link between serum homostachydrine and the risk and severity of PAH.

Our study has a number of limitations. First, the study population for the immune GWAS consisted entirely of female twins from the United Kingdom. While the focus on monozygotic and dizygotic twins may improve power for the identification of genetic correlates of immune traits, which will aid our MR work, the differences in the study population from both our PAH and our serine GWAS may have led to false negatives. All studies are based on European individuals, but clearly female twins are a nonrandom selection of the population that are likely to be genetically distinct, and SNP–trait associations may not carry over between GWASs. On this basis, we have confidence in our positive results, but we are less certain of our negative results; that is, we are not certain that there is not an immune basis for the protective effect of serine on PAH risk. Second, while we have addressed a number of possible confounders of the proposed causal relationship between metabolites and PAH risk, it is possible that additional confounders were missed. We have attempted to minimize this possibility through the application of multiple orthogonal analyses and the use of several independent cohorts. Third, the sample size of PAH cases in both the FinnGen GWAS and the UK clinical study is relatively small; moreover, the 2 FinnGen GWASs are not independent. The agreement between the MR and clinical analyses is significant, but it is possible that our study included false negatives, which will be revealed by future studies with larger sample sizes. Fourth, ATF4 has a number of roles independent of serine, including in the integrated stress response.[42] Further experimental study is necessary to determine whether the effect of ATF4 LOF on PAH can be corrected by exogenous addition of serine. Finally, in the initial MR study of serum metabolites as they relate to risk of PAH, homostachydrine passed the BH multiple testing correction but not the Benjamini–Yekutieli or the Bonferroni corrections, which make fewer assumptions about interdependence between statistical tests. In the use of the BH correction for multiple testing when using the IVW test to study the entire metabolome as it relates to risk of disease, we are following

a precedence in the literature,[43–46] but as a result we can be less confident that the homostachydrine result is a true positive. However, our follow-up analyses, including robust MR tests and study of PAH severity in an independent clinical cohort all support the idea that serum homostachydrine is linked to risk and severity of PAH.

In conclusion, we have demonstrated the utility of MR for the discovery of disease-associated metabolites. Our follow-up analysis in a clinical cohort adds significant orthogonal evidence for our conclusions by adding a direct measurement of serum metabolites in patients with PAH. We conclude that there is consistent evidence for a protective effect of serum serine on PAH and a harmful effect of homostachydrine. Both of these metabolites may be acting via modulation of energy metabolism. Further study is necessary and could lead to guidance cautioning consumption of certain food products containing homostachydrine or increasing serine consumption for patients with PAH.

Disclosures

None.

Supplemental Material

Tables S1–S3
Figures S1–S2

REFERENCES

1. Rhodes CJ, Sweatt AJ, Maron BA. Harnessing big data to advance treatment and understanding of pulmonary hypertension. *Circ Res*. 2022;130:1423–1444. doi: 10.1161/CIRCRESAHA.121.319969

2. Lemay S-E, Awada C, Shimauchi T, Wu W-H, Bonnet S, Provencher S, Boucherat O. Fetal gene reactivation in pulmonary arterial hypertension: GOOD, BAD, or BOTH? *Cells*. 2021;10:1473. doi: 10.3390/cells10061473

3. Rhodes CJ, Ghataorhe P, Wharton J, Rue-Albrecht KC, Hadinnapola C, Watson G, Bleda M, Haimel M, Coghlan G, Corris PA, et al. Plasma metabolomics implicates modified transfer RNAs and altered bioenergetics in the outcomes of pulmonary arterial hypertension. *Circulation*. 2017;135:460–475. doi: 10.1161/CIRCULATIONAHA.116.024602

4. Swietlik EM, Ghataorhe P, Zalewska KI, Wharton J, Howard LS, Taboada D, Cannon JE; UK National Cohort Study of PAH; Morrell NW, Wilkins MR, et al. Plasma metabolomics exhibit response to therapy in chronic thromboembolic pulmonary hypertension. *Eur Respir J*. 2021;57:2003201. doi: 10.1183/13993003.03201-2020

5. Julian TH, Boddy S, Islam M, Kurz J, Whittaker KJ, Moll T, Harvey C, Zhang S, Snyder MP, McDermott C, et al. A review of Mendelian randomization in amyotrophic lateral sclerosis. *Brain*. 2022;145:832–842. doi: 10.1093/brain/awab420

6. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res*. 2019;4:186. doi: 10.12688/wellcomeopenres.15555.1

7. Davey Smith G, Ebrahim S. "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32:1–22. doi: 10.1093/ije/dyg070

8. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, Reeve MP, Laivuori H, Aavikko M, Kaunisto MA, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature*. 2023;613:508–518. doi: 10.1038/s41586-022-05473-8

9. Shin S-Y, Fauman EB, Petersen A-K, Krumsiek J, Santos R, Huang J, Arnold M, Erte I, Forgetta V, Yang T-P, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014;46:543–550. doi: 10.1038/ng.2982

10. Kettunen J, Demirkan A, Würtz P, Draisma HHM, Haller T, Rawal R, Vaarhorst A, Kangas AJ, Lyytikäinen L-P, Pirinen M, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. *Nat Commun*. 2016;7:11122. doi: 10.1038/ncomms11122

11. Marshall DD, Powers R. Beyond the paradigm: combining mass spectrometry and nuclear magnetic resonance for metabolomics. *Prog Nucl Magn Reson Spectrosc*. 2017;100:1–16. doi: 10.1016/j.pnmrs.2017.01.001

12. Xu Y, Ritchie SC, Liang Y, Timmers PRHJ, Pietzner M, Lannelongue L, Lambert SA, Tahir UA, May-Wilson S, Foguet C, et al. An atlas of genetic scores to predict multi-omic traits. *Nature*. 2023;616:123–131. doi: 10.1038/s41586-023-05844-9

13. Moore C, Sambrook J, Walker M, Tolkien Z, Kaptoge S, Allen D, Mehenny S, Mant J, Di Angelantonio E, Thompson SG, et al. The INTERVAL trial to determine whether intervals between blood donations can be safely and acceptably decreased to optimise blood supply: study protocol for a randomised controlled trial. *Trials*. 2014;15:363. doi: 10.1186/1745-6215-15-363

14. Roederer M, Quaye L, Mangino M, Beddall MH, Mahnke Y, Chattopadhyay P, Tosi I, Napolitano L, Terranova Barberio M, Menni C, et al. The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis. *Cell*. 2015;161:387–403. doi: 10.1016/j.cell.2015.02.046

15. Boddy S, Islam M, Moll T, Kurz J, Burrows D, McGown A, Bhargava A, Julian TH, Harvey C, Marshall JN, et al. Unbiased metabolome screen leads to personalized medicine strategy for amyotrophic lateral sclerosis. *Brain Commun*. 2022;4:fcac069. doi: 10.1093/braincomms/fcac069

16. 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;40:304–314. doi: 10.1002/gepi.21965

17. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc*. 1995;57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

18. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001;29:1165–1188. doi: 10.1214/aos/1013699998

19. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37:658–665. doi: 10.1002/gepi.21758

20. Bowden J, Spiller W, Del Greco MF, Sheehan N, Thompson J, Minelli C, Davey SG. Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the radial plot and radial regression. *Int J Epidemiol*. 2018;47:1264–1278. doi: 10.1093/ije/dyy101

21. Bowden J, Del Greco MF, Minelli C, Davey Smith G, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: the role of the I^2 statistic. *Int J Epidemiol*. 2016;45:1961–1974. doi: 10.1093/ije/dyw220

22. Benza RL, Gomberg-Maitland M, Elliott CG, Farber HW, Foreman AJ, Frost AE, McGoon MD, Pasta DJ, Selej M, Burger CD, et al. Predicting survival in patients with pulmonary arterial hypertension: the REVEAL risk score calculator 2.0 and comparison with ESC/ERS-based risk assessment strategies. *Chest*. 2019;156:323–337. doi: 10.1016/j.chest.2019.02.004

23. Burgess S, Thompson SG. Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. *Am J Epidemiol*. 2015;181:251–260. doi: 10.1093/aje/kwu283

24. Wang Q, Dhindsa RS, Carss K, Harper AR, Nag A, Tachmazidou I, Vitsios D, Deevi SVV, Mackay A, Muthas D, et al. Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature*. 2021;597:527–532. doi: 10.1038/s41586-021-03855-y

25. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, et al. REVEL: an ensemble method for predicting the pathogenicity of rare missense variants. *Am J Hum Genet*. 2016;99:877–885. doi: 10.1016/j.ajhg.2016.08.016

26. Chen S, Francioli LC, Goodrich JK, Collins RL, Kanai M, Wang Q, Alföldi J, Watts NA, Vittal C, Gauthier LD, et al. A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. *bioRxiv*. 2022:2022.03.20.485034. Preprint posted online October 10, 2022.

27. Servillo L, Giovane A, Casale R, Cautela D, D'Onofrio N, Balestrieri ML, Castaldo D. Homostachydrine (pipecolic acid betaine) as authentication marker of roasted blends of *Coffea arabica* and *Coffea canephora* (Robusta) beans. *Food Chem*. 2016;205:52–57. doi: 10.1016/j.foodchem.2016.02.154

28. He WJ, Chen J, Razavi AC, Hu EA, Grams ME, Yu B, Parikh CR, Boerwinkle E, Bazzano L, Qi L, et al. Metabolites associated with coffee consumption and incident chronic kidney disease. *Clin J Am Soc Nephrol*. 2021;16:1620–1629. doi: 10.2215/CJN.05520421

29. Elsworth B, Lyon M, Alexander T, Liu Y, Matthews P, Hallett J, Bates P, Palmer T, Haberland V, Smith GD, et al. The MRC IEU OpenGWAS data infrastructure. *bioRxiv*. 2020.08.10.244293. Preprint posted online August 10, 2020.

30. Yang Y, Xu J, Zhou J, Xue J, Gao J, Li X, Sun B, Yang B, Liu Z, Zhao Z, et al. High betaine and dynamic increase of betaine levels are both associated with poor prognosis of patients with pulmonary hypertension. *Front Cardiovasc Med*. 2022;9:852009. doi: 10.3389/fcvm.2022.852009

31. Wu Q, Chen X, Li J, Sun S. Serine and metabolism regulation: a novel mechanism in antitumor immunity and senescence. *Aging Dis*. 2020;11:1640–1653. doi: 10.14336/AD.2020.0314

32. Tomaszewski M, Bębnowska D, Hrynkiewicz R, Dworzyński J, Niedźwiedzka-Rystwej P, Kopeć G, Grywalska E. Role of the immune system elements in pulmonary arterial hypertension. *J Clin Med Res*. 2021;10:3757. doi: 10.3390/jcm10163757

33. Ni S, Ji T, Dong J, Chen F, Feng H, Zhao H, Chen D, Ma W. Immune cells in pulmonary arterial hypertension. *Heart Lung Circ*. 2022;31:934–943. doi: 10.1016/j.hlc.2022.02.007

34. Jones RJ, De Bie EMDD, Groves E, Zalewska KI, Swietlik EM, Treacy CM, Martin JM, Polwarth G, Li W, Guo J, et al; UK National Cohort Study of Idiopathic and Heritable PAH Consortium. Autoimmunity is a significant feature of idiopathic pulmonary arterial hypertension. *Am J Respir Crit Care Med*. 2022;206:81–93. doi: 10.1164/rccm.202108-1919OC

35. Tajan M, Hennequart M, Cheung EC, Zani F, Hock AK, Legrave N, Maddocks ODK, Ridgway RA, Athineos D, Suárez-Bonnet A, et al. Serine synthesis pathway inhibition cooperates with dietary serine and glycine limitation for cancer therapy. *Nat Commun*. 2021;12:366. doi: 10.1038/s41467-020-20223-y

36. Perea-Gil I, Seeger T, Bruyneel AAN, Termglinchan V, Monte E, Lim EW, Vadgama N, Furihata T, Gavidia AA, Arthur Ataam J, et al. Serine biosynthesis as a novel therapeutic target for dilated cardiomyopathy. *Eur Heart J*. 2022;43:3477–3489. doi: 10.1093/eurheartj/ehac305

37. Rosner B. *Fundamentals of Biostatistics*. Cengage Learning; 2015:1–888.

38. Şanlı BA, Whittaker KJ, Motsi GK, Shen E, Julian TH, Cooper-Knock J. Unbiased metabolome screen links serum urate to risk of Alzheimer's disease. *Neurobiol Aging*. 2022;120:167–176. doi: 10.1016/j.neurobiolaging.2022.09.004

39. Julian TH, Cooper-Knock J, MacGregor S, Guo H, Aslam T, Sanderson E, Black GCM, Sergouniotis PI. Phenome-wide Mendelian randomisation analysis identifies causal factors for age-related macular degeneration. *Elife*. 2023;12:e82546. doi: 10.7554/eLife.82546

40. Zhao Y-Y, Liu Y, Stan R-V, Fan L, Gu Y, Dalton N, Chu P-H, Peterson K, Ross J Jr, Chien KR. Defects in caveolin-1 cause dilated cardiomyopathy and pulmonary hypertension in knockout mice. *Proc Natl Acad Sci U S A*. 2002;99:11375–11380. doi: 10.1073/pnas.172360799

41. Lin T, Gu J, Huang C, Zheng S, Lin X, Xie L, Lin D. (1)H NMR-based analysis of serum metabolites in monocrotaline-induced pulmonary arterial hypertensive rats. *Dis Markers*. 2016;2016:5803031. doi: 10.1155/2016/5803031

42. Pakos-Zebrucka K, Koryga I, Mnich K, Ljujic M, Samali A, Gorman AM. The integrated stress response. *EMBO Rep*. 2016;17:1374–1395. doi: 10.15252/embr.201642195

43. Gu Y, Jin Q, Hu J, Wang X, Yu W, Wang Z, Wang C, Liu Y, Chen Y, Yuan W. Causality of genetically determined metabolites and metabolic pathways on osteoarthritis: a two-sample Mendelian randomization study. *J Transl Med*. 2023;21:357. doi: 10.1186/s12967-023-04165-9

44. Tanha HM, Sathyanarayanan A; International Headache Genetics Consortium; Nyholt DR. Genetic overlap and causality between blood metabolites and migraine. *Am J Hum Genet*. 2021;108:2086–2098. doi: 10.1016/j.ajhg.2021.09.011

45. Amin N, Liu J, Bonnechere B, MahmoudianDehkordi S, Arnold M, Batra R, Chiou Y-J, Fernandes M, Ikram MA, Kraaij R, et al. Interplay of metabolome and gut microbiome in individuals with major depressive disorder vs control individuals. *JAMA Psychiatry*. 2023;80:597–609. doi: 10.1001/jamapsychiatry.2023.0685

46. Chen H, Qiao J, Wang T, Shao Z, Huang S, Zeng P. Assessing causal relationship between human blood metabolites and five neurodegenerative diseases with GWAS summary statistics. *Front Neurosci*. 2021;15:680104. doi: 10.3389/fnins.2021.680104