Integrating urinary and plasma omics to identify markers and therapeutic targets for cardiac disease

Authors and affiliations

Sophie C. de Ruiter, MSc [1], Marion van Vugt, MSc [2, 3, 4], Chris Finan, PhD [3, 4, 5, 6], Rui Provindencia, MD PhD [6, 7], Sandesh Chopade, MSc [3, 5], Diederick E. Grobbee, PhD [1], Hester M. den Ruijter, PhD [8], Sanne A.E. Peters, PhD [1, 9, *], A. Floriaan Schmidt, PhD [2, 3, 4, 5, *]

1. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
2. Department of Cardiology, Amsterdam Cardiovascular Sciences, Amsterdam University Medical Centres, University of Amsterdam, Amsterdam, the Netherlands
3. Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, United Kingdom
4. Division Heart and Lungs, Department of Cardiology, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
5. UCL British Heart Foundation Research Accelerator, London, United Kingdom
6. Health Data Research UK and Institute of Health Informatics, University College London, London, United Kingdom
7. Barts Heart Centre, St Bartholomew's Hospital, Barts Health NHS Trust, London, United Kingdom
8. Laboratory of Experimental Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
9. The George Institute for Global Health, School of Public Health, Imperial College London, London, UK

* joint senior author

Running title: Therapeutic targets and biomarkers for cardiac disease

Word count text: 3670

Word count abstract: 245

Number of references: 54

Number of tables and figures: 10

Number of (web) appendices: 2

Corresponding author: Sophie de Ruiter (S.C.deRuiter-4@umcutrecht.nl)

Abstract

Introduction: Urinary metabolites, representing kidney regulated filtration of metabolism end products, contain cardiac disease biomarkers such as NT-proBNP. We set out to integrate plasma proteins with urinary metabolites to identify potentially druggable metabolic pathways for cardiac disease.

Methods: Data was leveraged from a genome-wide association study (GWAS) on 954 urinary metabolites. Mendelian randomisation was used to identify urinary metabolites associating with atrial fibrillation (AF), heart failure (HF), dilated cardiomyopathy (DCM), or hypertrophic cardiomyopathy (HCM). By interrogating eight independent plasma protein GWAS, jointly including 92,277 participants and 1,562 unique proteins, we identified druggable plasma proteins with a directionally concordant effect on urinary metabolites and cardiac outcomes.

Results: In total, 38 unique urinary metabolites associated with cardiac disease, predominantly covering breakdown products from amino acid metabolism (n=12), xenobiotic metabolism (n=5), and unclassified metabolism origins (n=16). Subsequently, we identified 32 druggable proteins expressed in cardiac tissue, which had a directionally concordant association with the identified urinary metabolites and cardiac outcomes. This included positive control findings, for example higher values of AT1B2 (targeted by digoxin) decreased the risk of HCM, which we were able to link to a novel unclassified urinary metabolite (X-15497). Additionally, we showed that increased plasma RET values, a mediator of GDF-15 signalling, reduced the risk of HF, and linked this to the novel unclassified urinary breakdown product X-23776.

Conclusion: We were able to identify 32 druggable proteins affecting cardiac disease, and link these to urinary measurements of metabolite breakdown processes identifying potentially novel disease pathways.

Introduction

Heart failure (HF) and atrial fibrillation (AF) are cardiac diseases which manifest as ventricular and atrial dysfunction, respectively. While AF frequently affects people without established HF, in people with HF an AF diagnosis often signals early progressive HF, where ventricular disfunction has an early knock-on effect on atrial rhythm and output. People with hypertrophic cardiomyopathies (HCM), a subtype of HF characterized by an abnormally large and stiff cardiac septum, which may limit cardiac output and affect diastolic function, are particularly susceptible to develop AF. Aside from HCM, dilated cardiomyopathy (DCM) represents a second major non-ischemic HF subtype, where the left ventricle becomes enlarged and thin, limiting cardiac contractility. As such, AF and HF are closely interrelated diseases, with specific underlying pathologies like HCM and DCM playing distinct roles in their manifestation.

Aside from N-terminal prohormone of brain natriuretic peptide (NT-proBNP), which associates with HF and AF progression(1-3), there is paucity of biomarkers to assist medical practitioners in tailored treatment management, diagnosis, and prognosis. NT-proBNP, which is secreted by cardiomyocytes (cardiac cells responsible for contraction), can be measured in both blood plasma and urine, illustrating the interface between cardiac tissue processes, blood, and urine. Urine biomarkers provide a non-invasive measurement of potential causal metabolic changes reflecting disease onset and progression. While urinary metabolites represent by-products cleared from plasma, the source of these metabolites may reflect pathophysiological changes in tissues such as the cardiac muscle. Cardiac function and kidney function are intertwined, through for example the renin-angiotensin-aldosterone system, and the concept of cardiorenal syndrome has been implemented to describe disorders involving both organs, with dysfunction in one organ causing acute or chronic dysfunction in the other(4, 5). Urinary metabolites are therefore a potentially important source of non-invasive disease markers, which may unveil upstream involvement of biological pathways relevant for cardiac progression and/or novel cardiac therapeutics.

Recently, availability of affordable, high-throughput assays capable of measuring hundreds of proteins and metabolites has facilitated their analysis in large sample size genome wide association studies (GWAS). Given that proteins form the main constituents of drug targets, GWAS on protein values provide an invaluable resource to identify potential drug targets through genomics tools such as Mendelian randomisation (MR). Utilizing MR, one can infer the effect of an exposure (e.g. a metabolite or protein) on the risk of disease, by identifying genetic variants strongly associated with the exposure, and determining whether there is a dose-response relationship between the genetic associations with the exposure and the genetic associations with disease(6, 7). Given that genetic variants are fixed at gametogenesis, genetic evidence is robust against bias due to confounding or reverse causation(8, 9). These principles have been extensively validated in the context of cardiovascular disease(10-13), including in a recent study on the effects of plasma proteins on cardiac function and structure(14).

In the current manuscript we used MR to analyse urinary metabolism breakdown products for possible involvement with four cardiac disease traits: AF, HF, DCM, and HCM. Subsequently, we sought to establish potential plasma protein intermediates by identifying proteins associating with urinary metabolite values, as well as with cardiac disease risk. The Human Protein Atlas (HPA) was consulted to identify proteins expressed in cardiac tissue, providing further support for cardiac involvement. The subset of proteins expressed in cardiac tissue were subsequently mapped to biological pathways, with druggability determined through linkage with ChEMBL.

Methods

Available data

We used a GWAS on 954 urinary metabolites, which were identified and quantified among 1,627 individuals using a non-targeted mass spectrometry Metabolon assay(15). Metabolites were categorised based on the metabolic origin of the urinary breakdown products: amino acids, carbohydrates, cofactors

and vitamins, energy, lipids, nucleotides, peptides, xenobiotics, and unclassified. Xenobiotic metabolites - substances not naturally produced by the human body, include breakdown product from drugs, environmental pollutants, and various synthetic compounds. Unclassified metabolites are potentially novel metabolites which have not yet been related to an established metabolic origin. Genetic associations with plasma protein values were available from the eight GWAS, detailed in Supplementary Note 1. Genetic associations with cardiac disease onset were available from Nielsen *et al*. (60,620 AF cases)(16), from Shah *et al*. (47,309 HF cases)(17), from Garnier *et al*. (2,719 DCM cases)(18), and from Zhou *et al.* (2,993 HCM cases)(19). Study specific participants characteristics and case/control definitions are described in Supplementary Note 1.

Mendelian randomization analyses

Three MR analyses were conducted: 1) genome-wide MR identifying urinary metabolites which, through upstream effects in plasma, associate with onset of cardiac disease (step 1 **Figure 1, Appendix Tables S1-2**), 2) *cis*-MR identifying plasma proteins affecting values of urinary metabolites (step 2 **Figure 1**), and 3) *cis*-MR identifying plasma proteins which associate with the onset of cardiac disease (step 3 **Figure 1**). In the genome-wide MR we selected genetic variants from across the genome, while in the *cis*-MR genetic variants were selected from a 200 kilobase pair window around and within the protein encoding gene. In both cases, variants were selected based on an exposure F-statistic of at least 24 and a minor allele frequency of 0.01 or larger. The remaining variants were clumped to a linkage disequilibrium (LD) r-squared of 0.30, based on a random 5,000 sample of unrelated UK Biobank participants. The *cis*-MR analyses (steps 2-3 **Figure 1)** were conducted per individual protein GWAS, where the largest sample size GWAS was used in the case of overlap in measured proteins.

MR analyses were conducted using generalized least squares (GLS) implementations of the inverse variance weighted (IVW) estimator, as well as the pleiotropy robust MR-Egger estimator(20). The GLS implementation was used to additionally correct for residual LD, which optimized estimator precision(21).

To reduce the potential for horizontal pleiotropy, we excluded variants with large leverage statistics (larger than three times the mean leverage) and/or outlier statistics (Chi-square larger than 10.83). Analyses with fewer than 6 variants as instrument were discarded to ensure we had sufficient data to accurately model the exposure effects. The potential influence of horizontal pleiotropy was further minimized by applying a model selection framework identifying the MR model (IVW or MR-Egger) most supported by the available data.

Effect estimates and multiple testing

MR results are presented as mean differences (MD) for continuous outcomes, or odds ratios (ORs) for binary outcomes, representing the effect of an SD increase of the exposure (i.e., either urinary metabolite or plasma protein value). Effect estimates are accompanied by 95% confidence intervals (CIs) and p-values. Metabolite effects on disease were filtered for a multiplicity corrected p-value threshold of 1.37×10^{-4} based on the 365 principal components that were needed to explain 90% of the urinary metabolite variance (**Appendix Figure S1**). Protein MR effect estimates were evaluated against a multiplicity corrected p-value of 8.42×10^{-7} for the metabolites, based on the number of tested proteins (n=1,562) and metabolites (n=38). The multiplicity corrected p-value was 1.76×10^{-5} for cardiac outcomes, based on the number of tested proteins (n=709) and outcomes (n=4). Proteins are referred to using their Uniprot label(22), while genes are referred to using their Ensembl label in italicised font.

Triangulation of evidence through consideration of concordant cardiac effects

The three distinct MR analyses were integrated by I) identifying urinary metabolites affecting cardiac disease, II) identifying whether there were proteins associated with these urinary metabolites, and III) focussing on the subset of proteins with a directionally concordant effect on the same cardiac disease trait the urinary metabolites associated with. For example, if a higher value of a urinary metabolite decreased the risk of HF, and the protein increased the urinary metabolite value, the protein effect was considered directionally concordant if in turn higher protein values decreased HF risk. When all these conditions are

met – directional concordance between the protein, metabolite, and cardiac outcome – we describe this as a triangulated association (**Figure 2**). We note that because each performed MR analysis makes a distinct, and independent, assumption on the absence of horizontal pleiotropy, focusing on the directionally concordant results is anticipated to identify a robust subset of analyses with limited influence of potential horizontal pleiotropy.

Prioritization of proteins expressed in cardiac tissue

The set of proteins with a triangulated effect on metabolite and cardiac outcome were subsequently prioritised for the presence of cardiac tissue mRNA expression, differentiating between proteins with general cardiac expression, and proteins which were over-expressed (Supplementary Note 1).

Annotation of prioritized proteins

We identified plasma proteins targeted by approved drugs (drugged proteins), as well as potentially druggable proteins based on ChEMBL-v33. For the drugged and druggable proteins, we additionally extracted compound indications and side-effects, sourcing information from ChEMBL and the British National Formulary (BNF, access date September 4, 2021). Finally, all proteins available in the considered GWAS were mapped to biological pathways using the Reactome knowledgebase-v85(23) and prioritised proteins were tested for enrichment against the full set of 1,567 proteins.

Partial replication of protein associations with cardiac traits

Due to the availability of 8 proteomics GWAS, some studies assayed the same proteins, which we leveraged to replicate our findings, specifically focussing on associations with cardiac traits. Replication was sought using a nominally significant p-value of 0.05 or smaller and considering the effect direction of the triangulated analysis. In addition, a more stringent p-value threshold of 8.93×10^{-4} (0.05 divided by the number of proteins available in more than one study) was used for more conservative replication.

Results

Urinary metabolites associating with the onset of cardiac disease

Out of 954 evaluated urinary metabolites (**Appendix Table S1**), 38 metabolites were found to be associated with at least one type of cardiac disease: 25 for AF, 11 for HF, 5 for DCM, and 14 for HCM. (**Appendix Tables S2-3, Figure 3**). Urinary metabolites were frequently associated with AF, where particularly amino acid and xenobiotic breakdown products associated with cardiac diseases. For example, higher values of 1-methylhistamine increased the risk of HF (OR 1.07, 95%CI 1.05; 1.09), while decreasing the risk of AF (OR 0.83, 95%CI 0.76; 0.90). Higher values of 1-methylxanthine, a xenobiotic breakdown product of caffeine, decreased the risk of AF, DCM, and HCM (**Appendix Table S2, Figure 3)**. Finally, we were able to associate 16 urinary metabolites of unknown metabolic origin (unclassified metabolites) with cardiac disease.

Triangulated proteins with concordant effects on urinary metabolites and cardiac diseases

We identified 709 proteins associating with at least one of the 38 urinary metabolites, which included 235 proteins affecting the considered cardiac diseases (steps 2-3 **Figure 1, Appendix Figures S2-3)**. By considering directional concordance and cardiac tissue expression we identified 91 prioritised proteins (step 4 **Figure 1, Tables 1-4, Appendix Tables S4-5)**.

This prioritised set included 6 proteins which were overexpressed in cardiac tissues (**Figure 4**): TIG1 (affecting 1-methylxanthine), DNJA4 (affecting urinary metabolite X-19141), NAR3 (affecting urinary metabolites n-acetylcitrulline, 4-guanidinobutanoate, and X-23776), SPT20 (affecting metabolites 1-methylhistamine and X-19141), GLO2 (affecting metabolites n-methylpipecolate and 3-hydroxyindolin-2-one sulfate), and PLXA1 (affecting metabolites 1-methylhistamine, glycocholenate sulfate, and 1-methylxanthine) (**Appendix Table S6**).

Drugged and druggable triangulated proteins

The set of 91 prioritised proteins included 20 drugged proteins, and 12 druggable proteins (step 4 **Figure 1)**. Fourteen of the drugged proteins (ACES, AT1B2, DPEP1, EGLN1, ICAM1, MET, NQO2, NTRK3, OLA1, RET, TEC, TRY3, TRYB1, and UD11) were targeted by drugs with a cardiac indication and/or side effect (**Figure 5, Appendix Tables S7-8**). Among these, higher values of AT1B2 decreased the risk of HCM (OR 0.79, 95% CI 0.72; 0.87), which we were able to link to metabolite X-15497. AT1B2 is inhibited, as part of a protein complex group, through the inhibitor digoxin which is contraindicated for HCM. Higher values of ACES (targeted by acetylcholinesterase inhibitors indicated for myasthenia gravis) were associated with an increased risk of HF (OR 1.06, 95% CI 1.04; 1.08) and AF (OR 1.09, 95% CI 1.05; 1.13), supporting ongoing HF repurposing efforts of acetylcholinesterase inhibitors(24) (**Figure 5, Appendix Table S7**). Higher values of ICAM1 (affected by statins indicated for coronary heart disease(25), by BI-505 for myeloma, by alicaforsen for ulcerative colitis(26)) increased the risk of HF (OR 1.02, 95% CI 1.01; 1.03) and HCM (OR 1.10, 95% CI 1.06; 1.14). TRYB1 (targeted by arginine, in phase 2 evaluation for hypertension) was found to increase the risk of AF (OR 1.02, 95% CI 1.01; 1.03), and UD11 (targeted by various inhibiting drugs including the diuretics furosemide and bumetanide used for treatment of hypertension) increased the risk of DCM (OR 2.87, 95% CI 1.97; 4.17). Additionally, NQO2 (targeted by the alpha-1 blocker prazosin, indicated for hypertension) and RET (targeted by various drugs indicated for cancer) which mediates GDF15 signalling(27), were associated with a decreased risk of HF.

Irrespective of the known side effects or indications, the identified proteins frequently associated with AF, with many of these proteins (partially) acting through xenobiotic (n=8) and amino acid (n=6) metabolism (**Figure 5**). Specifically, GLO2, ACES, EGLN1 (OR 0.88, 95% CI 0.84; 0.92), LYAM2, IBP3, and ENTP5 affected AF and associated with xenobiotic metabolite values. Similarly, DPEP1 (OR 0.97, 95% CI 0.96; 0.98), TRYB1, HEXI1, and TEC (OR 0.87, 95% CI 0.83; 0.90) affected AF and associated with

amino acid metabolite values. Higher values of TRY3 (OR 0.92, 95% CI 0.90; 0.94) and NAGK decreased AF risk and associated with both amino acid and xenobiotic metabolites values.

Most proteins affecting HF associated with unclassified metabolism pathways (n=7). Specifically, increased values of ACES, RET (OR 0.97, 95% CI 0.96; 0.98), ADH1B, ERAP1 (OR 0.97, 95% CI 0.96; 0.98), FCG2A, NQO1, and CTGF affected HF risk and also associated with unclassified urinary metabolites.

Similarly, drugged and druggable proteins affecting HCM predominantly associated with unclassified metabolism pathways (n=5). Specifically, higher values of ICAM1, NTRK3, CAN2, SIGL7, and AT1B2 affected HCM risk.

For DCM we identified three drugged or druggable proteins with higher values of UD11 and CATA increasing DCM risk and associating with peptide metabolite prolylglycine, and higher values of PLA2R (OR 0.92, 95% CI 0.91; 0.94) decreasing DCM risk and associating with xenobiotic metabolite N-methylpipecolate.

Triangulated proteins with an effect on multiple cardiac traits

We additionally identified eleven pleiotropic proteins associated with multiple cardiac outcomes. Most of these were associated with unclassified and xenobiotic urinary metabolism breakdown products. Specifically, ACES, PLXB2, PLXA1, CYTF, and GNPTG affected AF and HF (**Figure 6**). ICAM1, MICB, and EVA1C affected both HCM and HF, while MATN3 and ACYP2 affected HCM and AF, while PLA2R affected HCM and DCM (**Figure 6, Appendix Table S6**).

Proteins enriched in biological pathways

Based on the Reactome knowledgebase, we found the "antigen presentation: folding, assembly and peptide loading of class I MHC" pathway to be enriched among the triangulated proteins. This pathway involved proteins TPSN (associated with HCM) and ERAP1 (associated with HF) (**Appendix Tables S6, S9**).

Replication of protein effects on cardiac outcomes

Of the 91 triangulated proteins associating with one or more cardiac traits 56 were available in more than a single study, applying a nominal 0.05 p-value threshold this resulted in a replication of 47 (84%) protein associations with cardiac disease. Applying a multiplicity corrected threshold resulted in 37 (66%) replicated proteins (**Appendix Table S10**). Replication results for protein effects on metabolite values are described in Supplementary Note 1 (**Appendix Table S11**).

Discussion

In this study, we used Mendelian randomisation to assess relationships between 954 urinary metabolites, 1,562 proteins, and four cardiac diseases AF, HF, DCM, and HCM with the aim of identifying potential biomarkers and therapeutic targets. We found directionally concordant associations between 38 metabolites, 91 cardiac expressed proteins, and the considered cardiac diseases. The identified urinary breakdown products associating with cardiac disease predominantly included metabolites with an origin in amino acids metabolism, xenobiotics metabolism, and of unclassified metabolism. Of the 91 proteins, six were overexpressed in cardiac tissue. Furthermore, 20 proteins were targeted by an approved drug, which included seven proteins that were targeted by a compound indicated for cardiovascular diseases: AT1B2 (digoxin), EGLN1 (roxadustat, currently in phase 2 clinical trials for ST-elevation myocardial infarction), ICAM1 (lovastatin), NQO2 (prazosin), TRYB1 (arginine), TRY3 (amiloride), and UD11 (furosemide), as well as seven targets for drugs with a known cardiovascular side-effect: ACES (donepezil), DPEP1 (cilastatin), MET (sirolimus, everolimus), NTRK3 (entrectenib), OLA1 (sirolimus, everolimus), RET

(regorafenib) and TEC (sirolimus, everolimus). We additionally identified eleven pleiotropic proteins with an effect on more than one cardiac outcome.

We identified seven amino acids metabolites associated with AF of which six (cysteinylglycine disulfide, 4-guanidinobutanoate, N-acetylcitrulline, N-acetyl-3-methylhistidine, N-acetylarginine, and N-acetylasparagine) increased the risk of AF, corroborating previous observations of altered amino acid metabolism in AF patients(28). We found 27 proteins associated with these amino acids and AF, including 7 drugged or druggable proteins. For example, we showed that higher values of MANBA (mannosidase beta) increased AF risk, which is supported by recent GWAS associating genetic variants near *MANBA* with AF(29) and by two recent MR studies identifying MANBA as a potential drug target for AF with corresponding effect directions(30, 31). MANBA is targeted by migalastat, a drug used to treat Fabry disease(32), which closely resembles HCM(33), with patients often suffering from cardiac arrhythmias and conduction tissue infiltration, leading to electrical instabilities. An observational study on Fabry disease patients revealed that atrial function was better preserved in those treated with migalastat compared to untreated individuals(34). Through our analyses of urinary metabolites, we were able to link MANBA to fatty acid metabolism through nonenedioate, with faulty fat metabolism a canonical characteristic of Fabry disease(35). Jointly, this suggest MANBA may offer therapeutic potential for treatment of AF and related arrythmias through altered lipid metabolism.

We identified eleven urinary metabolites associating with HF which included five metabolites with an origin in amino acids metabolism and five with an unclassified origin. We were able to map these to 27 plasma proteins with a directionally concordant HF effect, which included MICB (MHC class I polypeptide-related sequence B). MICB increased the risk of both HF and HCM, and was linked to seven yet unclassified metabolite breakdown products. This closely aligns with our pathway enrichment analysis identifying MHC class I related pathways implicated in HCM (based on protein TPSN) and HF (based on protein ERAP1), which are linked to CD8+ T-cell mediated cardiac remodelling(36). While these proteins

are not yet druggable, our analysis did identify ICAM1 (associating with HCM and HF), RET and NQO2 (both associating with HF), which have already been drugged. Through consideration of urinary metabolites we were able to associate ICAM1 with 1-methylxanthine and X-18889, where the former is a breakdown product of xanthine, and the oxidation of xanthine has been linked to ICAM1 neutrophil binding capacity(37). Notably, immune cell activation, reflecting systemic but low-grade inflammation, is a likely key pathophysiologic process progressing towards HCM, other cardiomyopathies, and end-stage HF(38). Regorafenib, a RET inhibiting drug that is used for cancer treatment, has registered side effects including myocardial ischemia and hypertension. These side effects support our findings that lower values of RET are associated with a higher risk of HF onset, where the unclassified urinary metabolite X-23776 was involved in the triangulated association. The GDNF family receptor α–like (GFRAL) acts a receptor for GDF15, which signals through the RET co-receptor. GDF15 was found to be relevant for energy homeostasis and has been identified as a biomarker of cardiovascular and metabolic diseases. There is potential for a new class of drugs based on GDF15-GFRAL-RET, where the signalling pathway is modified(27). Additionally, GDF15 has been recognized as an indicator for HF risk and prediction of hospitalisation for HF or all-cause mortality before(39). The protein NQO2 was associated with a decreased risk of HF, through the amino acid breakdown product indoleacetylglutamine. Prazosin, which acts as an adrenergic receptor alpha-1 antagonist and targets NQO2, has shown potential benefits for HF patients in previous observational studies and trials(40, 41), suggesting prazosin might be a relevant repurposing candidate for HF.

We found five urinary metabolites and eight proteins – of which three drugged or druggable – associated with DCM risk. For example, we showed that increased values of UD11 increased the risk for DCM and reduced the values of the peptide metabolite prolylglycine. UD11 is targeted by beta-blockers carvedilol and labetalol (used for HF, hypertension, coronary artery disease and myocardial ischemia), by the diuretics furosemide and bumetanide (used for HF and hypertension) and by acetazolamide, which is not

used in practice but has shown benefits in acute decompensated heart failure in a previous clinical trial(42).

Despite the robustness of combining multiple MR analyses sourcing various GWAS on both proteins and metabolites, there are some potential limitations that we wish to discuss. First of all, the genetic data we used was sourced from multiple GWAS which used various high throughput assays, varying in accuracy and analytic scope(43). Instead of directly measuring concentrations, these assays evaluate relative values of proteins or metabolites. Hence, the magnitude of association is difficult to interpret, and potential follow-up studies are needed to explore compound specific pharmacokinetic and pharmacodynamic properties. Our current analysis predominantly provides information on the putative drug mechanism and whether compounds should be inhibitory or activating. In addition, associations between urinary metabolites and cardiac disease should not be misinterpreted as implying that urinary metabolites themselves have a causal effect on disease. Rather we use urinary metabolites as a proxy for more distal metabolic processes, such as those occurring in or impacting cardiac tissue. Furthermore, GWAS data were mainly obtained from European cohorts, which may limit generalizability to non-European populations. Finally, while we had access to a large sample size GWAS on HF (47,309 cases), we did not have information on HF subtypes, and future studies considering HF subtypes may yield additional insights.

Examining the associations between urinary metabolites, cardiac disease and proteins in a triangulated way provides a comprehensive and robust analysis. Concordance in results from all three analyses – which are based on distinct sources of data – indicate a more reliable effect. Discordant results however do not necessarily indicate invalid MRs but are potentially indicative of a limited influence of the protein on the metabolism process influencing cardiac disease, or may reflect more intricate biological processes.

In conclusion, 38 urinary metabolites and 91 proteins were identified for involvement with cardiac disease, providing important evidence to support drug development of novel and repurposing targets. Additionally, our manuscript illustrates the value of integrating multi-modal sources of evidence to identify information potentially relevant for disease diagnosis, prognosis and aetiology.

Acknowledgments: A preprint version of this manuscript has been deposited at: <URL available upon publication> . This research has been conducted using the UK Biobank Resource under Application Number 12113. The authors are grateful to UK Biobank participants.

Funding: SD and SAEP are supported by a VIDI Fellowship (project number 09150172010050) from the Dutch Organisation for Health Research and Development (ZonMW) awarded to SAEP. AFS is supported by BHF grant PG/22/10989, the UCL BHF Research Accelerator AA/18/6/34223, MR/V033867/1, the National Institute for Health and Care Research University College London Hospitals Biomedical Research Centre, and the EU Horizon scheme (AI4HF 101080430 and DataTools4Heart 101057849). This work was supported by the NWO Snellius supercomputer project (application 2023.022).

Author contributions:

SD, AFS, and SP designed the study. SD and MV performed the analyses. SD drafted the manuscript. All authors provided critical input on the analysis, as well as the drafted manuscript.

Declaration of competing interests: AFS and CF have received funding from New Amsterdam Pharma for unrelated projects. The other authors declare that they have no conflict of interest.

Code availability: Analyses were conducted using Python v3.7.13 (for GNU Linux), Pandas v1.3.5, Numpy v1.21.6, bio-misc v0.1.4, and matplotlib v3.4.3.

Data availability: The individual GWAS data on cardiac outcomes leveraged in this study can be accessed as follows: atrial fibrillation (cases=60,620, total n=1,030,836; https://www.ebi.ac.uk/gwas/publications/30061737), heart failure (cases=47,309, total n=977,323; https://www.ebi.ac.uk/gwas/publications/31919418), dilated cardiomyopathy (cases=2,719, total n=6,980; https://www.ebi.ac.uk/gwas/publications/33677556) and hypertrophic cardiomyopathy (cases=2,993, total

n=1,197,200; http://results.globalbiobankmeta.org/pheno/HCM). The GWAS data on urinary metabolite values can be accessed via https://www.ebi.ac.uk/gwas/publications/31959995 (n=1,627). The individual GWAS data on plasma protein values can be accessed as follows: deCODE (n=35,559, https://www.decode.com/summarydata/), SCALLOP (n=30,931, https://www.ebi.ac.uk/gwas/publications/33067605), Ahola-Olli et al. (n=8,293, https://www.ebi.ac.uk/gwas/publications/27989323), Framingham (n=6,861, https://www.ebi.ac.uk/gwas/publications/21909115), AGES-Reykjavik (n=5,368, https://www.ebi.ac.uk/gwas/publications/35078996), INTERVAL (n=3,301, https://www.ebi.ac.uk/gwas/publications/29875488), Gilly et al. (n=1,328, https://www.ebi.ac.uk/gwas/publications/33303764), and Yang et al. (n=636, https://www.ebi.ac.uk/gwas/publications/34239129).

References

1. Cortés R, Portolés M, Salvador A, Bertomeu V, de Burgos FG, Martínez-Dolz L, et al. Diagnostic and prognostic value of urine NT-proBNP levels in heart failure patients. European journal of heart failure. 2006;8(6):621-7.
2. Panagopoulou V, Deftereos S, Kossyvakis C, Raisakis K, Giannopoulos G, Bouras G, et al. NTproBNP: an important biomarker in cardiac diseases. Current topics in medicinal chemistry. 2013;13(2):82-94.
3. Cao Z, Jia Y, Zhu B. BNP and NT-proBNP as diagnostic biomarkers for cardiac dysfunction in both clinical and forensic medicine. International journal of molecular sciences. 2019;20(8):1820.
4. Atlas SA. The renin-angiotensin aldosterone system: pathophysiological role and pharmacologic inhibition. Journal of managed care pharmacy. 2007;13(8 Supp B):9-20.
5. Rangaswami J, Bhalla V, Blair JE, Chang TI, Costa S, Lentine KL, et al. Cardiorenal syndrome: classification, pathophysiology, diagnosis, and treatment strategies: a scientific statement from the American Heart Association. Circulation. 2019;139(16):e840-e78.
6. Schmidt AF, Finan C, Gordillo-Marañón M, Asselbergs FW, Freitag DF, Patel RS, et al. Genetic drug target validation using Mendelian randomisation. Nature communications. 2020;11(1):1-12.
7. Schmidt AF, Hingorani AD, Finan C. Human Genomics and Drug Development. Cold Spring Harb Perspect Med. 2022;12(2).
8. Hingorani A, Humphries S. Nature's randomised trials. The Lancet. 2005;366(9501):1906-8.
9. de Ruiter SC, Schmidt AF, Grobbee DE, den Ruijter HM, Peters SA. Sex-specific Mendelian randomisation to assess the causality of sex differences in the effects of risk factors and treatment: Spotlight on hypertension. Journal of Human Hypertension. 2023:1-7.
10. Cupido AJ, Asselbergs FW, Natarajan P, Group CIW, Ridker PM, Hovingh GK, et al. Dissecting the IL-6 pathway in cardiometabolic disease: A Mendelian randomization study on both IL6 and IL6R. British Journal of Clinical Pharmacology. 2022;88(6):2875-84.
11. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nature genetics. 2020;52(10):1122-31.
12. Schmidt AF, Hunt NB, Gordillo-Marañón M, Charoen P, Drenos F, Kivimaki M, et al. Cholesteryl ester transfer protein (CETP) as a drug target for cardiovascular disease. Nature communications. 2021;12(1):5640.
13. Consortium I-RMRA. The interleukin-6 receptor as a target for prevention of coronary heart disease: a mendelian randomisation analysis. The Lancet. 2012;379(9822):1214-24.
14. Schmidt AF, Bourfiss M, Alasiri A, Puyol-Anton E, Chopade S, van Vugt M, et al. Druggable proteins influencing cardiac structure and function: Implications for heart failure therapies and cancer cardiotoxicity. Science Advances. 2023;9(17):eadd4984.
15. Schlosser P, Li Y, Sekula P, Raffler J, Grundner-Culemann F, Pietzner M, et al. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. Nature genetics. 2020;52(2):167-76.
16. Nielsen JB, Thorolfsdottir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. Nature genetics. 2018;50(9):1234-9.
17. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nature communications. 2020;11(1):1-12.
18. Garnier S, Harakalova M, Weiss S, Mokry M, Regitz-Zagrosek V, Hengstenberg C, et al. Genome-wide association analysis in dilated cardiomyopathy reveals two new players in systolic heart failure on chromosomes 3p25. 1 and 22q11. 23. European Heart Journal. 2021;42(20):2000-11.
19. Zhou W, Kanai M, Wu K-HH, Rasheed H, Tsuo K, Hirbo JB, et al. Global Biobank Meta-analysis Initiative: Powering genetic discovery across human disease. Cell Genomics. 2022;2(10).

20. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. International journal of epidemiology. 2015;44(2):512-25.
21. Burgess S, Zuber V, Valdes-Marquez E, Sun BB, Hopewell JC. Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. Genetic epidemiology. 2017;41(8):714-25.
22. Consortium U. The universal protein resource (UniProt) in 2010. Nucleic acids research. 2010;38(suppl_1):D142-D8.
23. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, et al. The reactome pathway knowledgebase. Nucleic acids research. 2020;48(D1):D498-D503.
24. Androne A, Hryniewicz K, Goldsmith R, Arwady A, Katz S. Acetylcholinesterase inhibition with pyridostigmine improves heart rate recovery after maximal exercise in patients with chronic heart failure. Heart. 2003;89(8):854-8.
25. Rezaie-Majd A, Prager GW, Bucek RA, Schernthaner GH, Maca T, Kress H-G, et al. Simvastatin reduces the expression of adhesion molecules in circulating monocytes from hypercholesterolemic patients. Arteriosclerosis, thrombosis, and vascular biology. 2003;23(3):397-403.
26. Hosten TA, Zhao K, Han HQ, Liu G, He XH. Alicaforsen: an emerging therapeutic agent for ulcerative colitis and refractory pouchitis. Gastroenterology Research. 2014;7(2):51.
27. Rochette L, Dogon G, Zeller M, Cottin Y, Vergely C. GDF15 and cardiac cells: current concepts and new insights. International Journal of Molecular Sciences. 2021;22(16):8889.
28. She J, Guo M, Li H, Liu J, Liang X, Liu P, et al. Targeting amino acids metabolic profile to identify novel metabolic characteristics in atrial fibrillation. Clinical science. 2018;132(19):2135-46.
29. Miyazawa K, Ito K, Ito M, Zou Z, Kubota M, Nomura S, et al. Cross-ancestry genome-wide analysis of atrial fibrillation unveils disease biology and enables cardioembolic risk prediction. Nature genetics. 2023;55(2):187-97.
30. Wang X, Huang T, Jia J. Proteome-Wide Mendelian Randomization Analysis Identified Potential Drug Targets for Atrial Fibrillation. Journal of the American Heart Association. 2023;12(16):e029003.
31. Ning Z, Huang Y, Lu H, Zhou Y, Tu T, Ouyang F, et al. Novel Drug Targets for Atrial Fibrillation Identified Through Mendelian Randomization Analysis of the Blood Proteome. Cardiovascular Drugs and Therapy. 2023:1-8.
32. McCafferty EH, Scott LJ. Migalastat: a review in Fabry disease. Drugs. 2019;79:543-54.
33. Moroni A, Tondi L, Milani V, Pieroni M, Pieruzzi F, Bevilacqua F, et al. Left atrial remodeling in hypertrophic cardiomyopathy and Fabry disease: A CMR-based head-to-head comparison and outcome analysis. International Journal of Cardiology. 2023;393:131357.
34. Pogoda C, Brand S-M, Duning T, Schmidt-Pogoda A, Sindermann J, Lenders M, et al. Impact of enzyme replacement therapy and migalastat on left atrial strain and cardiomyopathy in patients with Fabry disease. Frontiers in Cardiovascular Medicine. 2023;10.
35. Li X, Ren X, Zhang Y, Ding L, Huo M, Li Q. Fabry disease: Mechanism and therapeutics strategies. Frontiers in Pharmacology. 2022;13:1025740.
36. Komai K, Ito M, Nomura S, Shichino S, Katoh M, Yamada S, et al. Single-cell analysis revealed the role of CD8+ effector T cells in preventing cardioprotective macrophage differentiation in the early phase of heart failure. Frontiers in Immunology. 2021;12:763647.
37. Sellak H, Franzini E, Hakim J, Pasquier C. Reactive oxygen species rapidly increase endothelial ICAM-1 ability to bind neutrophils without detectable upregulation. 1994.
38. Lillo R, Graziani F, Franceschi F, Iannaccone G, Massetti M, Olivotto I, et al. Inflammation across the spectrum of hypertrophic cardiac phenotypes. Heart Failure Reviews. 2023:1-11.
39. Bradley J, Schelbert EB, Bonnett LJ, Lewis GA, Lagan J, Orsborne C, et al. Growth differentiation factor-15 in patients with or at risk of heart failure but before first hospitalisation. Heart. 2023.
40. Antani J, Antani NJ, Nanivadekar A. Prazosin in chronic congestive heart failure due to ischemic heart disease. Clinical cardiology. 1991;14(6):495-500.

41. Colucci WS, Wynne J, Holman BL, Braunwald E. Long-term therapy of heart failure with prazosin: a randomized double blind trial. The American Journal of Cardiology. 1980;45(2):337-44.
42. Mullens W, Dauw J, Martens P, Verbrugge FH, Nijst P, Meekers E, et al. Acetazolamide in acute decompensated heart failure with volume overload. New England Journal of Medicine. 2022;387(13):1185-95.
43. Katz DH, Robbins JM, Deng S, Tahir UA, Bick AG, Pampana A, et al. Proteomic profiling platforms head to head: leveraging genetics and clinical traits to compare aptamer-and antibody-based methods. Science Advances. 2022;8(33):eabm5164.
44. Ferkingstad E, Sulem P, Atlason BA, Sveinbjornsson G, Magnusson MI, Styrmisdottir EL, et al. Large-scale integration of the plasma proteome with genetics and disease. Nature Genetics. 2021;53(12):1712-21.
45. Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman ÅK, Schork A, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. Nature metabolism. 2020;2(10):1135-48.
46. Ahola-Olli AV, Würtz P, Havulinna AS, Aalto K, Pitkänen N, Lehtimäki T, et al. Genome-wide association study identifies 27 loci influencing concentrations of circulating cytokines and growth factors. The American Journal of Human Genetics. 2017;100(1):40-50.
47. Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, Chasman DI, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478(7367):103-9.
48. Gudjonsson A, Gudmundsdottir V, Axelsson GT, Gudmundsson EF, Jonsson BG, Launer LJ, et al. A genome-wide association study of serum proteins reveals shared loci with common diseases. Nature communications. 2022;13(1):1-13.
49. 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.
50. Gilly A, Park Y-C, Png G, Barysenka A, Fischer I, Bjørnland T, et al. Whole-genome sequencing analysis of the cardiometabolic proteome. Nature communications. 2020;11(1):1-9.
51. Yang C, Farias FH, Ibanez L, Suhy A, Sadler B, Fernandez MV, et al. Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. Nature neuroscience. 2021;24(9):1302-12.
52. Rücker G, Schwarzer G, Carpenter JR, Binder H, Schumacher M. Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. Biostatistics. 2011;12(1):122-42.
53. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, et al. The druggable genome and support for target identification and validation in drug development. Science translational medicine. 2017;9(383):eaag1166.

Tables

Table 1. Characteristics of concordant proteins and urinary metabolites associated with AF

UNIPROT ID	PROTEIN	DRUGGABILITY	METABOLITE CLASS(ES)	NO. METABOLITES	METABOLITE-CARDIAC EFFECT	PROTEIN-CARDIAC EFFECT	NO. DRUGS	CARDIAC DRUG EFFECT
Q8TB22	SPT20	Not yet druggable	Unclassified, Amino Acid	2	Mixed	1.05 (1.03; 1.07)	0	
P16581	LYAM2	Drugged	Xenobiotics	1	Decreasing	1.25 (1.17; 1.33)	4	
P35030	TRY3	Drugged	Xenobiotics, Amino Acid	2	Mixed	0.92 (0.90; 0.94)	15	Yes
P16444	DPEP1	Drugged	Amino Acid, Unclassified, Peptide	5	Mixed	0.97 (0.96; 0.98)	2	Yes
O95479	G6PE	Not yet druggable	Unclassified	1	Increasing	0.96 (0.94; 0.98)	0	
P19827	ITIH1	Not yet druggable	Unclassified	1	Increasing	0.94 (0.92; 0.96)	0	
Q16775	GLO2	Druggable	Xenobiotics	2	Mixed	0.90 (0.85; 0.94)	0	
Q14554	PDIA5	Not yet druggable	Amino Acid, Unclassified	2	Increasing	1.03 (1.02; 1.05)	0	
O76096	CYTF	Not yet druggable	Unclassified	2	Mixed	0.97 (0.95; 0.98)	0	
P22303	ACES	Drugged	Xenobiotics, Lipid	2	Mixed	1.09 (1.05; 1.13)	23	Yes
Q9UJJ9	GNPTG	Not yet druggable	Xenobiotics	1	Increasing	0.94 (0.92; 0.96)	0	
P35625	TIMP3	Not yet druggable	Unclassified, Peptide	2	Increasing	1.03 (1.02; 1.04)	0	
P16930	FAAA	Not yet druggable	Amino Acid	1	Increasing	0.90 (0.88; 0.92)	0	
P08581	MET	Drugged	Peptide	1	Increasing	0.92 (0.88; 0.95)	98	Yes
P02760	AMBP	Not yet druggable	Unclassified	1	Increasing	1.21 (1.13; 1.31)	0	
P32971	TNFL8	Not yet druggable	Amino Acid	2	Increasing	0.87 (0.82; 0.93)	0	
P80075	CCL8	Not yet druggable	Xenobiotics, Amino Acid, Unclassified, Carbohydrate	5	Mixed	1.02 (1.01; 1.03)	0	
O00462	MANBA	Drugged	Lipid	1	Increasing	1.07 (1.06; 1.08)	1	

Q9BUP3	HTAI2	Not yet druggable	Xenobiotics, Unclassified, Amino Acid	4	Increasing	0.95 (0.93; 0.97)	0	
Q8NFP4	MDGA1	Not yet druggable	Amino Acid	1	Increasing	0.98 (0.98; 0.99)	0	
Q9UIW2	PLXA1	Not yet druggable	Lipid, Amino Acid, Xenobiotics	3	Decreasing	1.09 (1.06; 1.12)	0	
Q9UJ70	NAGK	Druggable	Amino Acid, Xenobiotics	2	Increasing	0.91 (0.88; 0.94)	0	
P61244	MAX	Druggable	Unclassified	1	Increasing	1.09 (1.05; 1.14)	0	
Q9UHX3	EMR2	Not yet druggable	Lipid	1	Decreasing	0.92 (0.89; 0.95)	0	
O75973	C1QRF	Not yet druggable	Unclassified, Carbohydrate	3	Mixed	1.05 (1.04; 1.07)	0	
O15232	MATN3	Not yet druggable	Amino Acid, Unclassified, Xenobiotics	3	Mixed	0.97 (0.96; 0.98)	0	
P42680	TEC	Drugged	Amino Acid	1	Increasing	0.87 (0.83; 0.90)	61	Yes
Q9GZT9	EGLN1	Drugged	Xenobiotics	1	Increasing	0.88 (0.84; 0.92)	9	Yes
Q9NTK5	OLA1	Drugged	Unclassified	1	Increasing	0.81 (0.77; 0.85)	47	Yes
Q96EY8	MMAB	Not yet druggable	Unclassified	1	Increasing	1.04 (1.03; 1.06)	0	
Q8WW22	DNJA4	Not yet druggable	Unclassified	1	Increasing	1.34 (1.26; 1.44)	0	
Q6UW15	REG3G	Not yet druggable	Unclassified	1	Increasing	0.97 (0.96; 0.98)	0	
Q96JA1	LRIG1	Not yet druggable	Amino Acid	2	Increasing	0.94 (0.91; 0.96)	0	
P17936	IBP3	Druggable	Xenobiotics	1	Increasing	0.93 (0.91; 0.96)	0	
Q13508	NAR3	Not yet druggable	Unclassified, Amino Acid	4	Increasing	1.07 (1.04; 1.11)	0	
Q16270	IBP7	Not yet druggable	Amino Acid, Unclassified	4	Increasing	0.94 (0.92; 0.95)	0	
P14621	ACYP2	Not yet druggable	Unclassified	1	Increasing	0.84 (0.80; 0.88)	0	
Q15661	TRYB1	Drugged	Amino Acid, Unclassified	2	Increasing	1.02 (1.01; 1.03)	1	Yes
O94992	HEXI1	Druggable	Amino Acid	1	Increasing	0.83 (0.78; 0.88)	0	
O75356	ENTP5	Drugged	Xenobiotics	1	Increasing	0.89 (0.86; 0.93)	1	

Q9H3S1	SEM4A	Not yet druggable	Unclassified	1	Decreasing	1.23 (1.17; 1.29)	0
O15031	PLXB2	Not yet druggable	Unclassified	4	Increasing	1.05 (1.04; 1.06)	0
P01009	A1AT	Not yet druggable	Unclassified	1	Increasing	1.03 (1.02; 1.04)	0
Q29983	MICA	Not yet druggable	Xenobiotics, Lipid, Carbohydrate, Amino Acid, Unclassified	7	Mixed	1.02 (1.02; 1.03)	0

N.B. protein names are based on their Uniprot synonym. Druggability was based on whether the proteins are targeted by a developmental compound in clinical trials (druggable) or by a compound which has received marketing approval (drugged). The metabolite classes represent the identified metabolism origin of the urinary breakdown products associated with AF. The protein-cardiac effect column represents MR results as the odds ratios and 95% confidence intervals, representing the effect of an SD increase of the protein value on risk of AF. The No. drug column records the number of approved drugs with affinity for the protein as available in ChEMBL. For a full list of the compounds and their indication and side-effects please refer to **Appendix Table S7**.

Table 2. Characteristics of concordant proteins and urinary metabolites associated with HF

UNIPROT ID	PROTEIN	DRUGGABILITY	METABOLITE CLASS(ES)	NO. METABOLITES	METABOLITE-CARDIAC EFFECT	PROTEIN-CARDIAC EFFECT	NO. DRUGS	CARDIAC DRUG EFFECT
O76096	CYTF	Not yet druggable	Unclassified	1	Increasing	0.93 (0.91; 0.95)	0	
P22303	ACES	Drugged	Xenobiotics, Unclassified	2	Mixed	1.06 (1.04; 1.08)	23	Yes
P05362	ICAM1	Drugged	Xenobiotics	1	Decreasing	1.02 (1.01; 1.03)	6	Yes
Q9UJJ9	GNPTG	Not yet druggable	Unclassified	1	Decreasing	1.08 (1.05; 1.10)	0	
P13929	ENOB	Not yet druggable	Amino Acid	1	Increasing	1.05 (1.03; 1.08)	0	
P02656	APOC3	Druggable	Amino Acid	1	Increasing	1.29 (1.18; 1.41)	0	
Q9NWX6	THG1	Not yet druggable	Unclassified	1	Decreasing	1.07 (1.05; 1.10)	0	
Q9UIW2	PLXA1	Not yet druggable	Amino Acid	1	Increasing	0.90 (0.87; 0.92)	0	
P29279	CTGF	Druggable	Unclassified	1	Increasing	1.08 (1.05; 1.11)	1	
P49788	TIG1	Not yet druggable	Xenobiotics	1	Decreasing	1.03 (1.02; 1.03)	0	
P16083	NQO2	Drugged	Amino Acid	1	Increasing	0.95 (0.93; 0.97)	64	Yes
Q07866	KLC1	Not yet druggable	Amino Acid	1	Increasing	1.28 (1.20; 1.38)	0	
Q9NT99	LRC4B	Not yet druggable	Amino Acid	1	Increasing	1.27 (1.16; 1.38)	0	
P49914	MTHFS	Not yet druggable	Unclassified	1	Decreasing	1.06 (1.04; 1.08)	0	
P12318	FCG2A	Druggable	Unclassified	1	Increasing	0.99 (0.98; 0.99)	0	
P25311	ZA2G	Not yet druggable	Unclassified	1	Increasing	0.92 (0.90; 0.95)	0	
Q9NZ08	ERAP1	Druggable	Unclassified	1	Increasing	0.97 (0.96; 0.98)	1	
P07949	RET	Drugged	Unclassified	1	Decreasing	0.97 (0.96; 0.98)	75	Yes
P58658	EVA1C	Not yet druggable	Xenobiotics	1	Decreasing	1.51 (1.35; 1.70)	0	
Q96EE4	CC126	Not yet druggable	Amino Acid, Unclassified	3	Increasing	1.07 (1.04; 1.09)	0	
Q5SRI9	MANEA	Not yet druggable	Amino Acid	1	Increasing	0.97 (0.97; 0.98)	0	
P15559	NQO1	Drugged	Unclassified	1	Increasing	0.97 (0.96; 0.98)	4	
O15031	PLXB2	Not yet druggable	Unclassified	1	Increasing	1.03 (1.02; 1.04)	0	
P00326	ADH1B	Druggable	Unclassified	1	Decreasing	1.14 (1.08; 1.20)	2	
O75521	ECI2	Not yet druggable	Amino Acid	1	Increasing	0.93 (0.90; 0.96)	0	
O14931	NCTR3	Not yet druggable	Unclassified	1	Increasing	1.03 (1.02; 1.05)	0	
Q29980	MICB	Not yet druggable	Unclassified	1	Increasing	1.05 (1.04; 1.07)	0	

N.B. protein names are based on their Uniprot synonym. Druggability was based on whether the proteins are targeted by a developmental compound in clinical trials (druggable) or by a compound which has received marketing approval (drugged). The metabolite classes represent the identified metabolism origin of the urinary breakdown products associated with AF. The protein-cardiac effect column represents MR results as the odds ratios and 95% confidence intervals, representing the effect of an SD increase of the protein value on risk of AF. The No. drug column records the number of approved drugs with affinity for the protein as available in ChEMBL. For a full list of the compounds and their indication and side-effects please refer to **Appendix Table S7**.

Table 3. Characteristics of concordant proteins and urinary metabolites associated with HCM

UNIPROT ID	PROTEIN	DRUGGABILITY	METABOLITE CLASS(ES)	NO. METABOLITES	METABOLITE-CARDIAC EFFECT	PROTEIN-CARDIAC EFFECT	NO. DRUGS	CARDIAC DRUG EFFECT
P05362	ICAM1	Drugged	Xenobiotics, Unclassified	2	Mixed	1.10 (1.06; 1.14)	6	Yes
O76003	GLRX3	Not yet druggable	Unclassified, Peptide	2	Increasing	2.97 (2.21; 4.00)	0	
P01042	KNG1	Not yet druggable	Amino Acid, Peptide	2	Increasing	0.80 (0.74; 0.87)	0	
P14415	AT1B2	Drugged	Unclassified	1	Increasing	0.79 (0.72; 0.87)	9	Yes
O15444	CCL25	Not yet druggable	Unclassified	1	Increasing	1.16 (1.09; 1.23)	0	
O15232	MATN3	Not yet druggable	Unclassified	1	Increasing	1.20 (1.14; 1.27)	0	
Q9UH65	SWP70	Not yet druggable	Unclassified	1	Increasing	0.87 (0.81; 0.92)	0	
Q16288	NTRK3	Drugged	Unclassified	1	Increasing	0.73 (0.65; 0.83)	38	Yes
Q9NR46	SHLB2	Not yet druggable	Unclassified	1	Increasing	1.27 (1.16; 1.40)	0	
P05155	IC1	Not yet druggable	Unclassified	2	Increasing	1.16 (1.09; 1.24)	0	
Q13018	PLA2R	Druggable	Xenobiotics	1	Increasing	1.06 (1.03; 1.08)	0	
Q9BUD6	SPON2	Not yet druggable	Amino Acid	1	Decreasing	1.56 (1.32; 1.83)	0	
O95861	BPNT1	Not yet druggable	Unclassified	1	Increasing	0.29 (0.17; 0.49)	0	
P17655	CAN2	Drugged	Unclassified	1	Decreasing	0.57 (0.46; 0.70)	2	
Q96PJ5	FCRL4	Not yet druggable	Unclassified	1	Increasing	0.89 (0.84; 0.93)	0	
P58658	EVA1C	Not yet druggable	Unclassified	1	Increasing	0.27 (0.19; 0.38)	0	
Q9Y286	SIGL7	Druggable	Unclassified	1	Increasing	0.51 (0.39; 0.68)	0	
P14621	ACYP2	Not yet druggable	Unclassified	1	Increasing	1.36 (1.23; 1.51)	0	
Q9NT22	EMIL3	Not yet druggable	Amino Acid	1	Decreasing	0.64 (0.53; 0.78)	0	
O43155	FLRT2	Not yet druggable	Amino Acid	1	Decreasing	0.63 (0.52; 0.77)	0	
Q29980	MICB	Not yet druggable	Unclassified	6	Increasing	1.09 (1.07; 1.11)	0	
P0C0L4	CO4A	Not yet druggable	Unclassified	1	Increasing	0.88 (0.85; 0.92)	0	
O15533	TPSN	Not yet druggable	Unclassified	1	Increasing	0.75 (0.65; 0.85)	0	

N.B protein names are based on their Uniprot synonym. Druggability was based on whether the proteins are targeted by a developmental compound in clinical trials (druggable) or by a compound which has received marketing approval (drugged). The metabolite classes represent the identified metabolism origin of the urinary breakdown products associated with AF. The protein-cardiac effect column represents MR results as the odds ratios and 95% confidence intervals, representing the effect of an SD increase of the protein value on risk of AF. The No. drug column records the number of approved drugs with affinity for the protein as available in ChEMBL. For a full list of the compounds and their indication and side-effects please refer to **Appendix Table S7**.

Table 4. Characteristics of concordant proteins and urinary metabolites associated with DCM

UNIPROT ID	PROTEIN	DRUGGABILITY	METABOLITE CLASS(ES)	NO. METABOLITES	METABOLITE-CARDIAC EFFECT	PROTEIN-CARDIAC EFFECT	NO. DRUGS	CARDIAC DRUG EFFECT
Q6B8I1	DS13A	Not yet druggable	Lipid	1	Decreasing	0.37 (0.27; 0.51)	0	
O14879	IFIT3	Not yet druggable	Lipid	2	Decreasing	3.65 (2.34; 5.69)	0	
P04040	CATA	Druggable	Peptide	1	Decreasing	1.99 (1.61; 2.46)	0	
Q14515	SPRL1	Not yet druggable	Xenobiotics, Unclassified	2	Decreasing	1.14 (1.08; 1.20)	0	
Q13018	PLA2R	Druggable	Xenobiotics	1	Decreasing	0.92 (0.91; 0.94)	0	
Q9UHC6	CNTP2	Not yet druggable	Unclassified	1	Decreasing	0.76 (0.67; 0.86)	0	
Q9UI08	EVL	Not yet druggable	Lipid	1	Decreasing	2.08 (1.71; 2.53)	0	
P22309	UD11	Drugged	Peptide	1	Decreasing	2.87 (1.97; 4.17)	76	Yes

N.B protein names are based on their Uniprot synonym. Druggability was based on whether the proteins are targeted by a developmental compound in clinical trials (druggable) or by a compound which has received marketing approval (drugged). The metabolite classes represent the identified metabolism origin of the urinary breakdown products associated with AF. The protein-cardiac effect column represents MR results as the odds ratios and 95% confidence intervals, representing the effect of an SD increase of the protein value on risk of AF. The No. drug column records the number of approved drugs with affinity for the protein as available in ChEMBL. For a full list of the compounds and their indication and side-effects please refer to **Appendix Table S7**.

Figure legends (main figures)

Figure 1. Flowchart of the main steps of this study

Abbreviations: AF = atrial fibrillation, DCM = dilated cardiomyopathy, HCM = hypertrophic cardiomyopathy, HF = heart failure.

Figure 2. A triangulation diagram illustrating an example of a triangulated association – with directional concordance – between a protein and its effects on urinary metabolite values and cardiac disease.

N.B. The plus symbol (in red) represents a positive association. The minus symbol (in blue) represents a negative association (a risk-decreasing effect).

Figure 3. Urinary metabolites associated with at least one cardiac outcome

N.B. Significant associations after multiplicity correction are indicated by a star, non-significant associations by a dot. The top row provides the number of drugged and druggable proteins, sourced from BNF and ChEMBL. Genetic associations with the cardiac outcomes were obtained from Nielsen *et al*. (60,620 AF cases)(16), from Shah *et al*. (47,309 HF cases)(17), from Garnier *et al*. (2,719 DCM cases)(18), and from Zhou *et al.* (2,993 HCM cases)(19). For more detailed information, please refer to the Methods section and **Appendix Table S2**.

Abbreviations: AF = atrial fibrillation, DCM = dilated cardiomyopathy, HCM = hypertrophic cardiomyopathy, HF = heart failure, MR = mendelian randomisation.

Figure 4. Cardiac mRNA expression of directionally concordant proteins affecting cardiac disease

N.B. Proteins are identified by their Uniprot label. Proteins displayed in bold font indicate that they are overexpressed in cardiac tissue compared to non-cardiac tissue. Data were sourced from the Human Protein Atlas; please refer to the Methods section for more details.

Figure 5. Forest plots of cardiac effects of drugged or druggable prioritised proteins

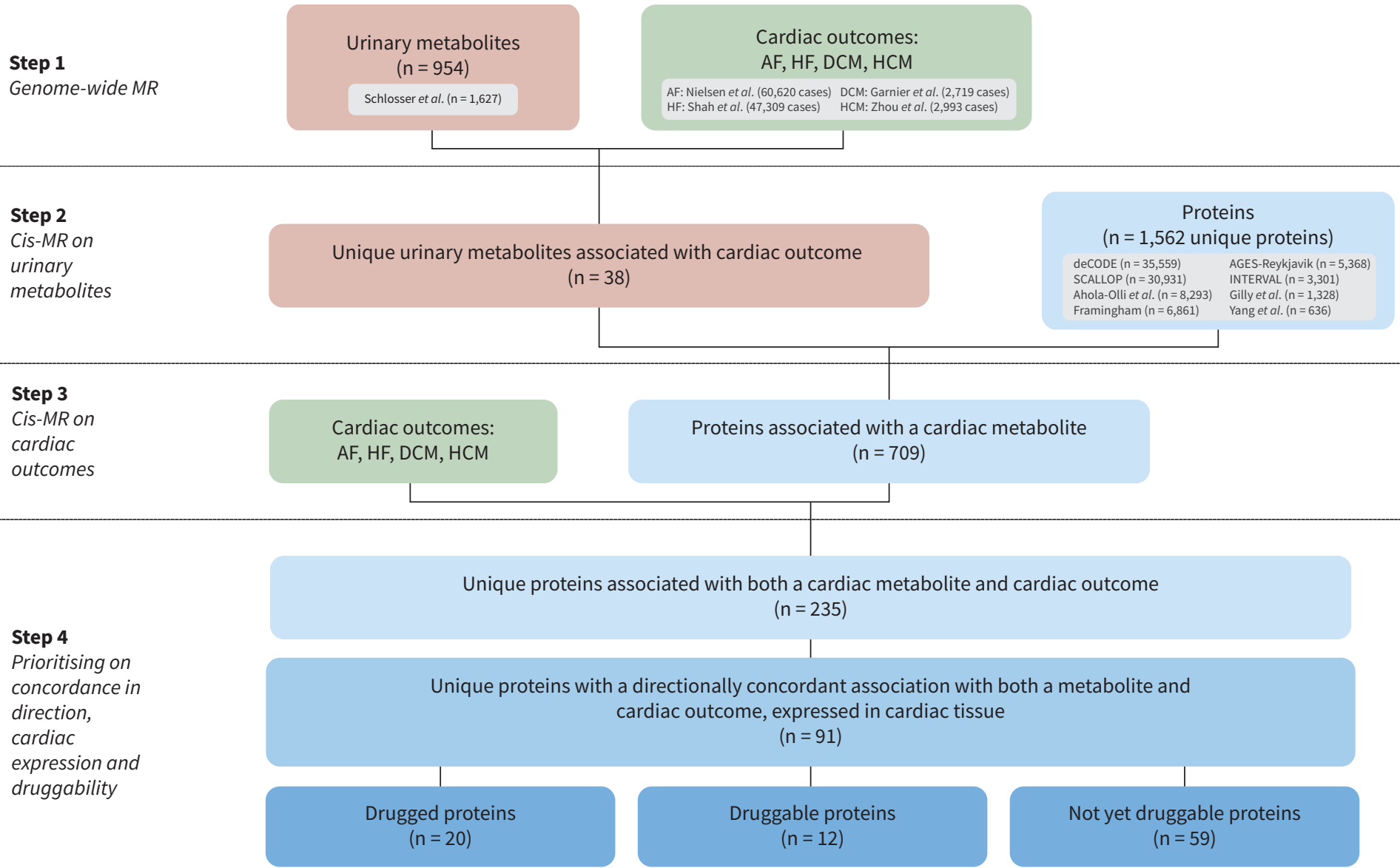
N.B. The metabolite classes of the metabolites affected by the protein are indicated on the right y-axis. Genetic associations with the cardiac outcomes were sourced from Nielsen *et al*. (60,620 AF cases)(16), from Shah *et al*. (47,309 HF cases)(17), from Garnier *et al*. (2,719 DCM cases)(18), and from Zhou *et al.* (2,993 HCM cases)(19). For a more detailed description, please refer to the Methods section. Full numerical results can be found in **Appendix Table S6-7**.

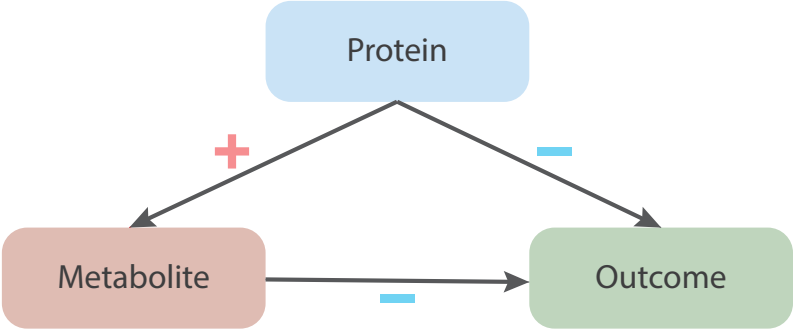
Abbreviations: AF = atrial fibrillation, CI = confidence interval, DCM = dilated cardiomyopathy, HCM = hypertrophic cardiomyopathy, HF = heart failure, OR = odds ratio.

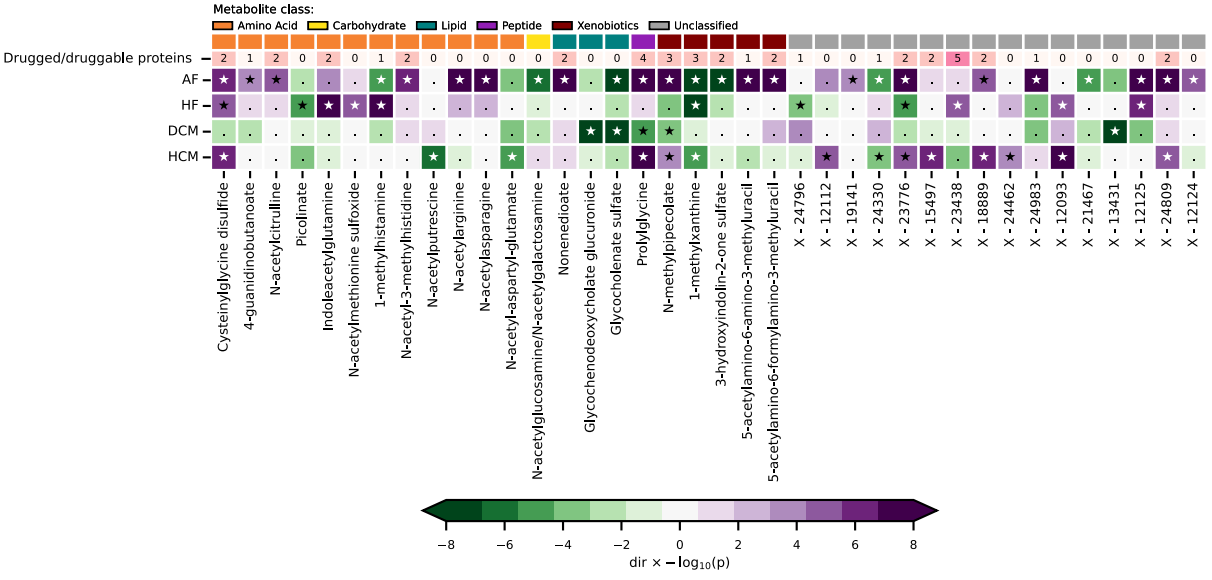
Figure 6. Forest plot of the cardiac effects of prioritised proteins associating with more than one cardiac outcome

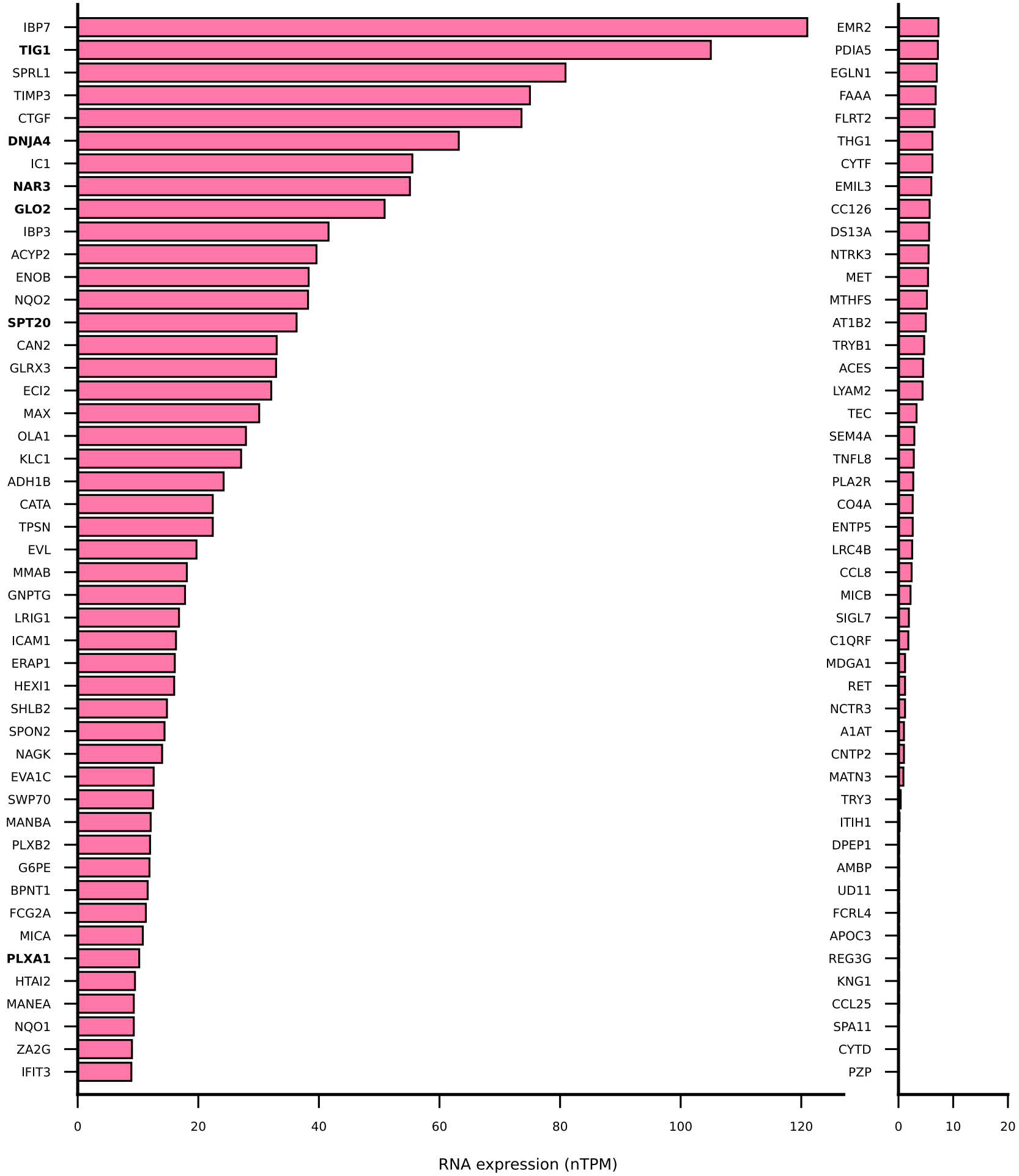
N.B. The metabolite classes of the metabolites affected by the protein are indicated on the right y-axis. Genetic associations with the cardiac outcomes were sourced from Nielsen *et al*. (60,620 AF cases)(16), from Shah *et al*. (47,309 HF cases)(17), from Garnier *et al*. (2,719 DCM cases)(18), and from Zhou *et al.* (2,993 HCM cases)(19). For a more detailed description, please refer to the Methods section. Full numerical results can be found in **Appendix Table S6**.

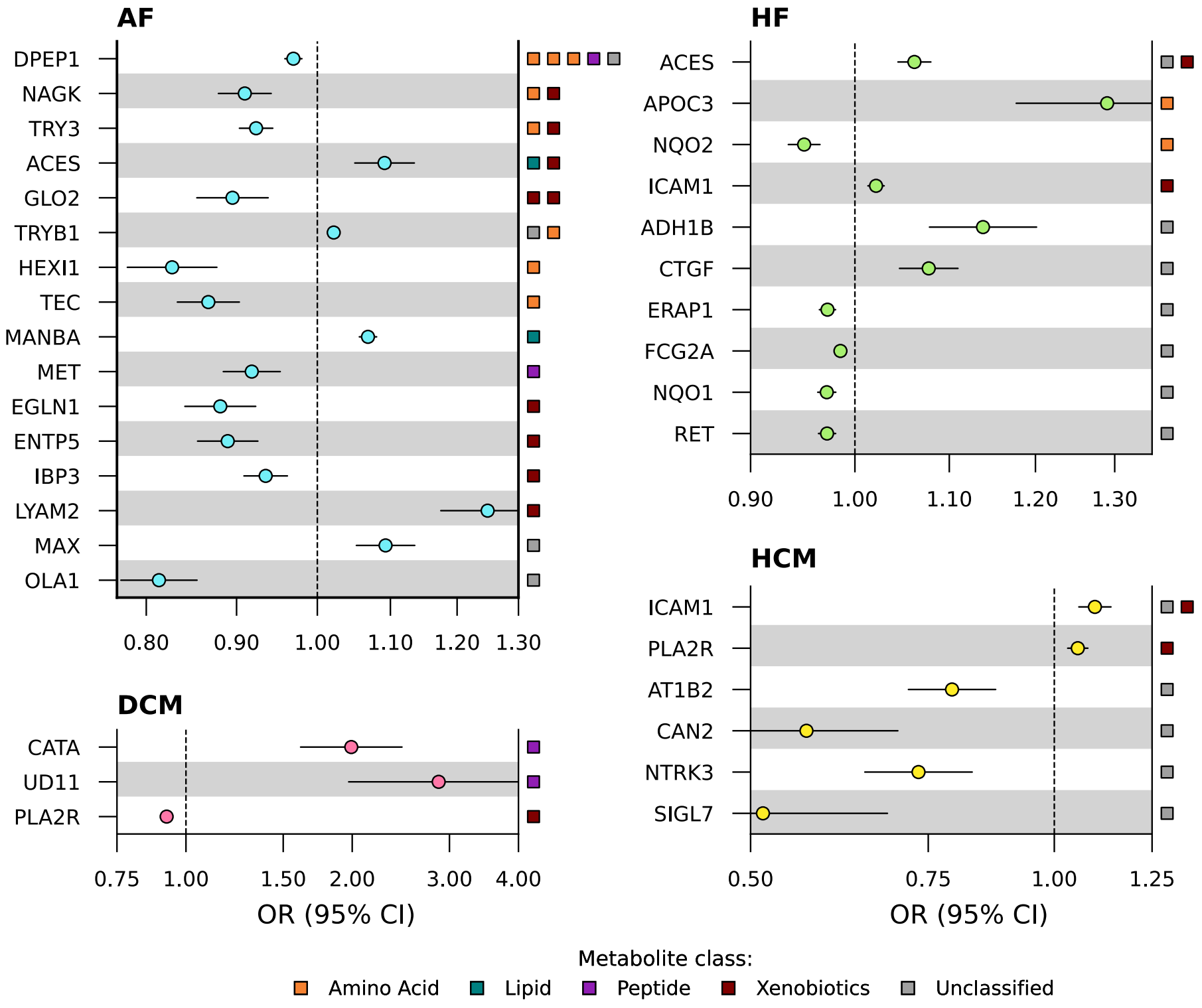
Abbreviations: AF = atrial fibrillation, CI = confidence interval, DCM = dilated cardiomyopathy, HCM = hypertrophic cardiomyopathy, HF = heart failure, OR = odds ratio.

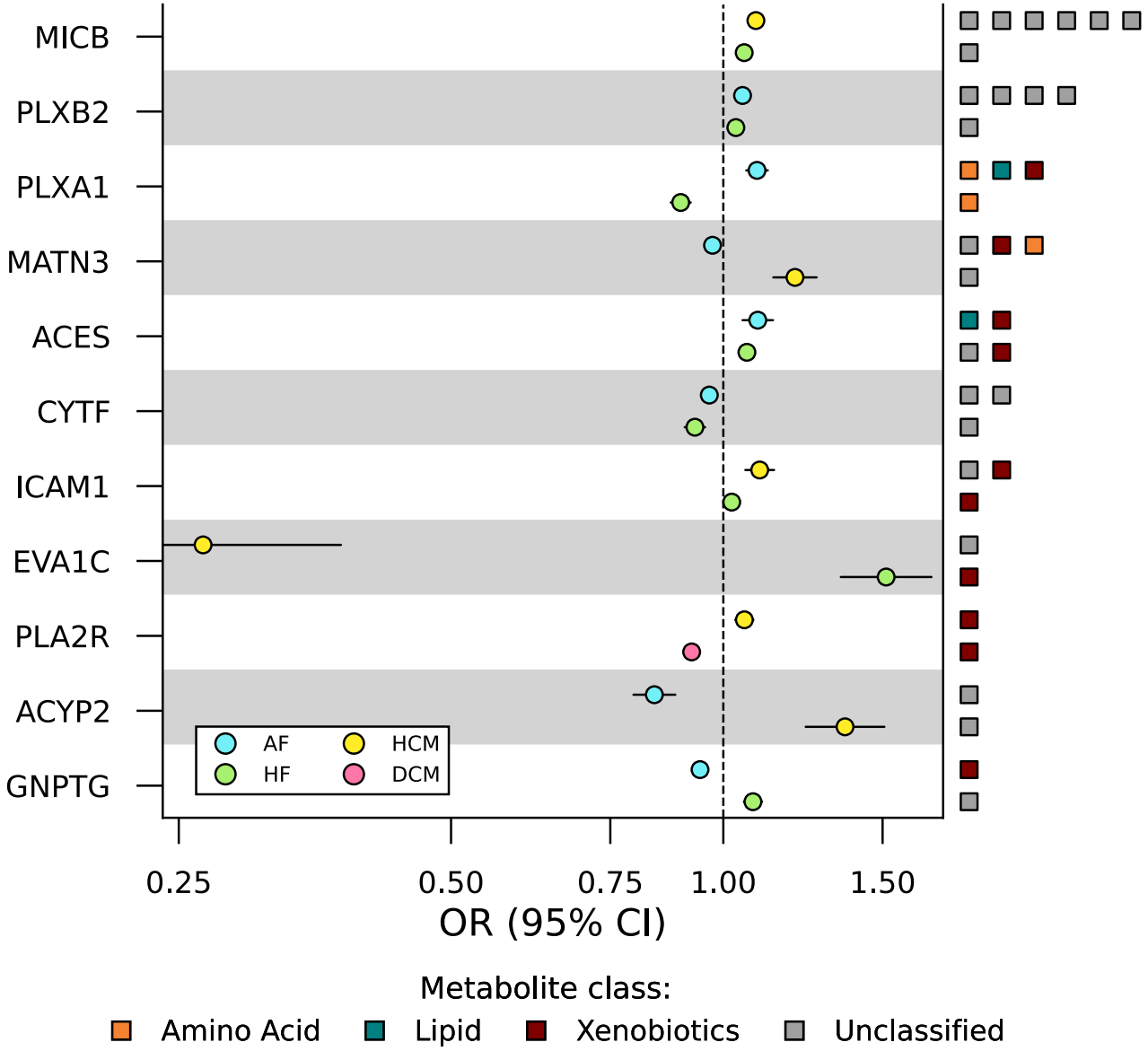

Step 1
Genome-wide MR
Urinary metabolites
(n = 954)
Schlosser et al. (n = 1,627)
Cardiac outcomes:
AF, HF, DCM, HCM
AF: Nielsen et al. (60,620 cases)
HF: Shah et al. (47,309 cases)
DCM: Garnier et al. (2,719 cases)
HCM: Zhou et al. (2,993 cases)
Step 2
Cis-MR on urinary metabolites
Unique urinary metabolites associated with cardiac outcome
(n = 38)
Proteins
(n = 1,562 unique proteins)
deCODE (n = 35,559)
SCALLOP (n = 30,931)
Ahola-Olli et al. (n = 8,293)
Framingham (n = 6,861)
AGES-Reykjavik (n = 5,368)
INTERVAL (n = 3,301)
Gilly et al. (n = 1,328)
Yang et al. (n = 636)
Step 3
Cis-MR on cardiac outcomes
Cardiac outcomes:
AF, HF, DCM, HCM
Proteins associated with a cardiac metabolite
(n = 709)
Step 4
Prioritising on concordance in direction, cardiac expression and druggability
Unique proteins associated with both a cardiac metabolite and cardiac outcome
(n = 235)
Unique proteins with a directionally concordant association with both a metabolite and cardiac outcome, expressed in cardiac tissue
(n = 91)
Drugged proteins
(n = 20)
Druggable proteins
(n = 12)
Not yet druggable proteins
(n = 59)


Protein
+
−
Metabolite
Outcome
−


Metabolite class:
Amino Acid
Carbohydrate
Lipid
Peptide
Xenobiotics
Unclassified
Drugged/druggable proteins
AF
HF
DCM
HCM
Cysteinylglycine disulfide
4-guanidinobutanoate
N-acetylcitrulline
Picolinate
Indoleacetylglutamine
N-acetylmethionine sulfoxide
1-methylhistamine
N-acetyl-3-methylhistidine
N-acetylputrescine
N-acetylarginine
N-acetylasparagine
N-acetyl-aspartyl-glutamate
N-acetylglucosamine/N-acetylgalactosamine
Nonenedioate
Glycochenodeoxycholate glucuronide
Glycocholenate sulfate
Prolylglycine
N-methylpipecolate
1-methylxanthine
3-hydroxyindolin-2-one sulfate
5-acetylamino-6-amino-3-methyluracil
5-acetylamino-6-formylamino-3-methyluracil
X - 24796
X - 12112
X - 19141
X - 24330
X - 23776
X - 15497
X - 23438
X - 18889
X - 24462
X - 24983
X - 12093
X - 21467
X - 13431
X - 12125
X - 24809
X - 12124
−8
−6
−4
−2
0
2
4
6
8
dir × −log10(p)


IBP7
TIG1
SPRL1
TIMP3
CTGF
DNJA4
IC1
NAR3
GLO2
IBP3
ACYP2
ENOB
NQO2
SPT20
CAN2
GLRX3
ECI2
MAX
OLA1
KLC1
ADH1B
CATA
TPSN
EVL
MMAB
GNPTG
LRIG1
ICAM1
ERAP1
HEXI1
SHLB2
SPON2
NAGK
EVA1C
SWP70
MANBA
PLXB2
G6PE
BPNT1
FCG2A
MICA
PLXA1
HTAI2
MANEA
NQO1
ZA2G
IFIT3
EMR2
PDIA5
EGLN1
FAAA
FLRT2
THG1
CYTF
EMIL3
CC126
DS13A
NTRK3
MET
MTHFS
AT1B2
TRYB1
ACES
LYAM2
TEC
SEM4A
TNFL8
PLA2R
CO4A
ENTP5
LRC4B
CCL8
MICB
SIGL7
C1QRF
MDGA1
RET
NCTR3
A1AT
CNTP2
MATN3
TRY3
ITIH1
DPEP1
AMBP
UD11
FCRL4
APOC3
REG3G
KNG1
CCL25
SPA11
CYTD
PZP
0
20
40
60
80
100
120
0
10
20
RNA expression (nTPM)


AF
DPEP1
NAGK
TRY3
ACES
GLO2
TRYB1
HEXI1
TEC
MANBA
MET
EGLN1
ENTP5
IBP3
LYAM2
MAX
OLA1
0.80
0.90
1.00
1.10
1.20
1.30
HF
ACES
APOC3
NQO2
ICAM1
ADH1B
CTGF
ERAP1
FCG2A
NQO1
RET
0.90
1.00
1.10
1.20
1.30
DCM
CATA
UD11
PLA2R
0.75
1.00
1.50
2.00
3.00
4.00
OR (95% CI)
HCM
ICAM1
PLA2R
AT1B2
CAN2
NTRK3
SIGL7
0.50
0.75
1.00
1.25
OR (95% CI)
Metabolite class:
Amino Acid
Lipid
Peptide
Xenobiotics
Unclassified


MICB
PLXB2
PLXA1
MATN3
ACES
CYTF
ICAM1
EVA1C
PLA2R
ACYP2
GNPTG
AF
HF
HCM
DCM
0.25
0.50
0.75
1.00
1.50
OR (95% CI)
Metabolite class:
Amino Acid
Lipid
Xenobiotics
Unclassified