nature portfolio

Peer Review File

Circulating monocyte gene expression profiles associated with cardiac remodeling and incident heart failure in the Multi-Ethnic Study of Atherosclerosis

Corresponding Author: Dr Tess Peterson

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Peterson et al evaluated the association between blood monocyte gene expression profiles (Illumina microarrays), cardiac remodeling by CMR imaging, and incident heart failure in 1,195 subjects without ischemic heart disease enrolled in the MESA study. The authors report that of seven target traditional monocyte activation markers, CD163 was associated with LV dysfunction and CD14 was weakly associated with time to incident HF. Agnostic analysis identified 1,136 monocyte genes associated with cardiac remodeling by CMR. Of 96 genes identified associated with more than one cardiac remodeling parameters, three significantly associated with time to incident HF – PCCB, MCTP1, and VIM, PCCB and MCTP1 with a lower hazard for HF, and VIM with a higher hazard for HF. Association of these genes with specific cardiac characteristics was variable. Cis-eQTL Mendelian randomization using a separate GWAS cohort validated the association between PCCB and HF. The authors propose that monocyte PCCB is a strong candidate for future mechanistic investigation in non-ischemic HF.

The study topic is interesting and uses well curated patient cohorts with careful analyses. While associative, these data provide a rich resource for further study and future mechanistic investigations, and also highlight a heretofore unappreciated association between the monocyte metabolic gene PCCP and HF. I would ask the authors to address the following:

1. Methods – while there were 813 patient MRIs in the analysis sample, there was variability in the specific parameters measured (only 329 subjects had ECV analysis for instance). The implications of this for the presented findings should be discussed and may be mentioned as a study limitation.
2. Methods – the monocyte isolation process used anti-CD14 magnetic beads. This may have resulted in low enrichment of non-classical CD16+ monocytes that generally have low CD14 expression. The analysis is most accurately described as a study of CD14+ monocyte gene expression.
3. Figure 2 – monocyte CCR2 expression appears to significantly correlate with ECV fraction. Please discuss.
4. Table 2 – CD14 expression associated significantly with the time to incident HF. However, the text states this was not significant upon adjustment for multiple testing. Please describe this in further detail and show the loss of significance.
5. Figure 4A – the authors found no enrichment for Gene Ontology biological processes using the 96 genes that associated with more than one cardiac MRI characteristic. There were no common genes for all MRI characteristics. What are the implications then for biological plausibility? Please discuss.
6. Results, page 9. PCCP was associated with a lower hazard for incident HF, positive association with larger LA volume and fibrosis, an but better GCS. Why the dichotomous association between strain and fibrosis for PCCB?
7. PCCB, MTCP1, and VIM expression exhibited low pairwise correlations – can the authors discuss the implications of this weak association.
8. On page 13, the authors state: " poorer left ventricular systolic functon(LV-GCS, LVEF) and greater concentricity (LVM/V) was negatively associated with genes related to mitochondrial ATP synthesis, oxidative phosphorylation, and the TCA cycle, possibly suggestive of more pro-inflammatory activity." This wording is confusing – please revise.
9. Please include the gene set enrichment analyses related to oxidative stress pathways in the main figures rather than the supplement.

Reviewer #2

(Remarks to the Author)
This is an interesting manuscript about the role of monocyte in heart failure. The authors used transcriptomic as well as cardiac mri images and used those data to identify transcriptomic profiles of patients that associate with heart failure incidence, cardiac function and fibrosis.
The data are sound, the topic is of interest and the manuscript well written. I have only one suggestion to improve the quality of the manuscript.
1) The authors should validate the expression of their top targets in isolated cells- ideally from a separate cohort- or at least in the isolated monocytes used for the study and compare the expression also to isolated cells from heart failure patients.

Reviewer #3

(Remarks to the Author)
The authors explored monocyte gene-expression data from atherosclerosis-free participants in the MESA study and conducted both targeted and untargeted exploration of associations between expression data and CMR measurements. The authors identified several overlaps between CMR features and gene expression in monocytes, and some of these genes were associated with clinical heart-failure (HF) outcomes. The manuscript is well written, and the data presentation is sensible. My comments below.

[Major]

The main question in this paper is whether a transcript associated with more than two cardiac MRI features is more important than transcripts that are more strongly associated with a single measure or with HF outcomes. Using “>1 features” might have been a good filter to select a tractable number of genes, but I could not find a reasonable biological rationale for focusing on this criterion.

The associations among PCCB,expression QTLs, CMR measures, and heart-failure (HF) outcomes are not explicitly shown and are not readily traceable for readers. A clear connection is needed across genetic variation (eQTL), expression–endophenotype (CMR), and expression–outcome (HF). On a related note, were the CMR measures actually associated with HF incidence?

[Minor]

Abstract Line 9
"PCCB was associated with >1 cardiac characteristic" This is a bit difficult to follow. Ideally, the authors should provide the actual cardiac MRI measures (e.g., “XXXX was associated with LAVi, LV-GCS, and ECV,” etc.).

Line 79
"The analysis sample for incident HF—which did not require CMR—was n=1195 (Supplemental Figure S1B)." How many participants overlap between the CMR/transcriptomic analysis and the incident-HF analysis? Does this overlap (or lack thereof) affect the results and their interpretation?

Line 133 - 143
The authors list 54 genes among 96 genes associated with CMR phenotypes, yet none of the pathways are enriched in these 96 genes. It is a bit difficult to follow the feasibility and interpretability of this analysis. Such descriptive, post-hoc reasoning does not clarify what is newly learned in this study. I was wondering whether, under either a more stringent threshold (e.g., FDR < 0.01 or Bonferroni) or a more lenient one (e.g., FDR < 0.20), any pathways became enriched.

Line 153 - 165
The authors provided directional explanations of these measures (adverse/protective), which helps readers who are not familiar with cardiac MRI measurements. However, in the sentence for XXX, the stated direction is opposite to that of the other genes and to the effect on clinical outcomes. This should be aligned with the rest. For example: “Higher monocyte expression of MTCP1 was associated with better LV function,” etc.

Line 186
Given the modest association with HF (P = 0.01), the claim about “no evidence of significant pleiotropy was identified” reads as an overstatement, particularly when other outcomes exhibit stronger evidence (P < 0.01). I recommend tempering the language or contextualizing it relative to the other results

Line 188 - 191
The authors utilize cis-eQTLs to evaluate associations between XXX and UUU expression and heart failure. The UUU eQTL was associated with heart failure (P = 0.01). I could not find information about the SNP’s effect direction (e.g., increased/decreased UUU expression and increased/decreased HF risk), nor could I find the SNP’s identifier. Please consider providing the rsID and chromosome/position and reference/alternate alleles. In addition, reporting allele frequency and the effect size/direction for both the eQTL and the heart-failure GWAS would aid interpretation.

Line 191 needs line break

Figure 3C
The results need additional clarification. The density plot is difficult to interpret because the underlying elements are not defined. Please specify: (i) whether the plotted units are individual genes within each pathway, (ii) the number of genes per ontology, (iii) whether the gene set was pre-filtered by association P values with the relevant CMR measures (and the threshold used), and (iv) whether the reported FDR corresponds to pathway-level enrichment significance.

Figure 4B
This figure is difficult to interpret in its current form. Given the number of feature intersections, an UpSet plot would communicate the set overlaps more clearly and compactly (see https://en.wikipedia.org/wiki/UpSet_plot).

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Thank you for addressing all of my concerns. The revised manuscript is much improved - I have no further comments. Congratulations to the authors on a nice study.

Reviewer #2

(Remarks to the Author)
The authors addressed all my comments.

Reviewer #3

(Remarks to the Author)
The authors responded to all requests clearly. Thank you for your additional effort.

Circulating monocyte gene expression profiles associated with cardiac remodeling and incident heart failure: the Multi-Ethnic Study of Atherosclerosis

REVIEW RESPONSE

We thank each Reviewer for their thoughtful comments and suggestions, which we address point-by-point below, alongside corresponding analysis and text edits. Page and line numbers are in reference to the clean version of the revised submission. We are confident these revisions have greatly improved the messaging and our overall manuscript.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Peterson et al evaluated the association between blood monocyte gene expression profiles (Illumina microarrays), cardiac remodeling by CMR imaging, and incident heart failure in 1,195 subjects without ischemic heart disease enrolled in the MESA study. The authors report that of seven target traditional monocyte activation markers, CD163 was associated with LV dysfunction and CD14 was weakly associated with time to incident HF. Agnostic analysis identified 1,136 monocyte genes associated with cardiac remodeling by CMR. Of 96 genes identified associated with more than one cardiac remodeling parameters, three significantly associated with time to incident HF – PCCB, MCTP1, and VIM, PCCB and MCTP1 with a lower hazard for HF, and VIM with a higher hazard for HF. Association of these genes with specific cardiac characteristics was variable. Cis-eQTL Mendelian randomization using a separate GWAS cohort validated the association between PCCB and HF. The authors propose that monocyte PCCB is a strong candidate for future mechanistic investigation in non-ischemic HF.

The study topic is interesting and uses well curated patient cohorts with careful analyses. While associative, these data provide a rich resource for further study and future mechanistic investigations, and also highlight a heretofore unappreciated association between the monocyte metabolic gene PCCB and HF. I would ask the authors to address the following:

1. Methods – while there were 813 patient MRIs in the analysis sample, there was variability in the specific parameters measured (only 329 subjects had ECV analysis for instance). The implications of this for the presented findings should be discussed and may be mentioned as a study limitation.

   We agree this is an important detail. Variations in sample size across imaging metrics (detailed in Supplemental Figure S1A) was largely driven by the fact contrast-enhanced MRIs were only obtained among a subset of MESA participants as part of a CMR sub-study. To more clearly address these differences, we have added a table to the supplement (**Supplementary Data S2B**) comparing cardiac MRI participants with vs. without ECV measured—the one contrast-enhanced MRI metric we analyzed—which we also include below. It suggested participants included vs. excluded from ECV-specific cross-sectional analyses had differing distributions of some key risk factors, including race and ethnicity, total blood cholesterol, and lipid-lowering therapy use. Groups statistically differed by age and BMI, as well, though absolute differences were small.

CHARACTERISTIC	MEDIAN [IQR] or N (%)		p-value
	With ECV (n=329)	Without ECV (n=484)	
Age, years	67 [60, 75]	69 [62, 76]	0.04
Male sex	53% (173)	47% (230)	0.18
Race and ethnicity	--	--	
White	67% (220)	37% (178)	<0.001
Black	18% (58)	24% (118)	
Hispanic	15% (51)	39% (188)	
Smoking status	--	--	
Current	10% (34)	8% (40)	0.08
Former	51% (169)	46% (220)	
Pack-years of smoking	1.5 [0.0, 20.5]	0.0 [0.0, 13.4]	0.05
Body mass index, kg/m^2	28.8 [25.7, 32.2]	27.9 [25.2, 31.3]	0.04
Systolic blood pressure, mmHg	118 [109, 133]	121 [111, 137]	0.06
Anti-hypertensive medication	53% (173)	52% (251)	0.90
Total cholesterol, mg/dL	174 [155, 201]	186 [161, 208]	0.002
High-density lipoprotein cholesterol, mg/dL	51 [43, 61]	52 [43, 64]	0.10
Lipid-lowering medication	45% (148)	33% (160)	0.001
Diabetes [a]	18% (58)	15% (75)	0.48
Estimated glomerular filtration rate, mL/min/1.73 m^2	83.7 [70.8, 92.6]	80.9 [68.1, 92.6]	0.11

We have also added a note to the Limitations section of the manuscript acknowledging the potential for selection bias between these cross-sectional analyses due to these sample differences (page 20, lines 444-447):

"ECV was available only among a subset of participants who differed from the broader CMR sample on distributions of some risk factors; this may have introduced selection bias, limiting comparability between ECV and other cross-sectional imaging analyses."

2. Methods – the monocyte isolation process used anti-CD14 magnetic beads. This may have resulted in low enrichment of non-classical CD16+ monocytes that generally have low CD14 expression. The analysis is most accurately described as a study of CD14+ monocyte gene expression.

   We have added "CD14+" as a descriptor in the Abstract (page 2, line 6), Introduction (page 5, line 69), Results (page 11, line 205) and now clearly note this as a potential limitation in the Discussion (page 20, lines 438-441):

   "First, monocytes were isolated using anti-CD14 magnetic beads; given traditionally classified non-classical/Mon3 subsets generally express low levels of CD14, this likely limited our ability to fully assess relationships of gene expression in monocytes with low membrane-bound CD14."

3. Figure 2 – monocyte CCR2 expression appears to significantly correlate with ECV fraction. Please discuss.

   While there looks to be an association between higher expression of *CCR2* and myocardial ECV, we are performing several statistical tests in these analyses, which we must account for in our inferences. These estimates are indeed significant without multiple testing correction but are no longer significant with correction. Our current main body text does not adequately describe this, so we have clarified in multiple places. First, we have added the following text to the Methods on these targeted analyses to clarify how they were adjusted for multiple testing and distinguishing them from agnostic analyses (page 8, lines 150-151):

   "Targeted analyses were corrected for family-wise error using Bonferroni adjustment."

   We also added a direct description of this particular CCR2/ECV result to corresponding main body Results text (page 11, lines 214-218):

   Old text:

   "No other target monocyte phenotype markers (CD14, CCR2, CCL2, HLA-DRA, CX3CR1, or CD16) were associated with cardiac phenotypes following multiple testing correction."

   Updated text:

   "While there may have been a positive association between monocyte CCR2 expression and ECV (Model 2 mean difference: 0.40; 95% CI: 0.07–0.73, p=0.02), this estimate did not meet our threshold for significance, following Bonferroni adjustment for multiple testing (p<0.007). Therefore, no other target monocyte phenotype markers (CD14, CCR2, CCL2, HLA-DRA, CX3CR1, or CD16) were associated with cardiac phenotypes following multiple testing correction."

   We have also added the following text to the Figure 2 legend to clarify the Bonferroni threshold for significance and which estimates met those thresholds:

   "The threshold for two-sided test significance adjusting for family-wise error using Bonferroni was p<0.007, met only by estimates of association between monocyte CD163 expression and both LV-GCS and LVEF."

4. Table 2 – CD14 expression associated significantly with the time to incident HF. However, the text states this was not significant upon adjustment for multiple testing. Please describe this in further detail and show the loss of significance.

   Thank you for the opportunity to clarify this, which relates to the prior comment, as well. In addition to the Methods text addition described above, which clarifies the use of Bonferroni (page 8, lines 150-151: "*Targeted analyses were corrected for family-wise error using Bonferroni adjustment*."), we also added a threshold descriptor to the Table 2 legend text:

   "The threshold for significance adjusting for family-wise error using Bonferroni was p<0.007."

5. Figure 4A – the authors found no enrichment for Gene Ontology biological processes using the 96 genes that associated with more than one cardiac MRI characteristic. There were no common genes for all MRI characteristics. What are the implications then for biological plausibility? Please discuss.

   We certainly agree the lack of statistical enrichment among the 96 individual genes associated with more than one CMR metric may make interpretation of biological overlap more difficult. We see two key points of consideration here.

   First, we do believe this lack of statistical enrichment is in part related to low statistical power for this particular enrichment test (as opposed to true lack of biological relatedness) and henceforth why we chose to still describe some patterns in Results text (pages 12-13, lines 245-256). For example, the GSEA applied to the results of our incident HF analysis yielded several GO terms with enrichment at an FDR<0.10, shown below (instead of FDR<0.05, which we report in accordance with our *a priori* threshold). Also shown here is that, because the list comprises only 96 evaluated genes (instead of the full 9,430 genes used in GSEA tests for individual CMR metrics), the power for these GSEA tests (evidenced by small 'setSizes') was very low for most processes.

ID	Description	setSize	NES	p-value	FDR
GO:0051641	cellular localization	14	2.078888	0.001105	0.059775
GO:0007010	cytoskeleton organization	7	1.928939	0.002031	0.059775
GO:0010628	positive regulation of gene expression	6	1.953126	0.002756	0.060881
GO:0007178	transmembrane receptor protein serine/ threonine kinase signaling pathway	3	1.743671	0.004727	0.087544
GO:0007155	cell adhesion	6	1.879532	0.005539	0.087544
GO:0048856	anatomical structure development	19	1.876543	0.005566	0.087544
GO:0045087	innate immune response	7	1.726583	0.010779	0.098185
GO:0006935	chemotaxis	3	1.631898	0.013258	0.101342

   Related, because GSEA applied to results of CMR analyses had much higher power, patterns in enrichment across those model sets are especially important. We indeed identified several patterns of enrichment across CMR metrics, which we describe in the Results (page 12, lines 233-243), throughout the Discussion, and present partially in **Figure 3B** and fully in **Supplementary Data S10-S14**.

   Second, we intentionally chose to evaluate CMR metrics that capture complementary (instead of directly overlapping) biology, e.g., structural remodeling (LVMI), contractile function (LV-GCS), chronically elevated LV filling pressure (LAVI), and expansion of myocardial extracellular space (ECV). Our data support this distinction, with evaluated CMR metrics having moderate to low correlation in our study population (all |*rho*|<0.36, shown below).

	LVEF	|LV-GCS|	LVM/V	LAVI	ECV
LVEF	--	0.24	-0.03	0.04	0.07
|LV-GCS|	--	--	-0.36	0.23	0.29
LVM/V	--	--	--	-0.33	-0.25
LAVI	--	--	--	--	0.19
ECV	--	--	--	--	--

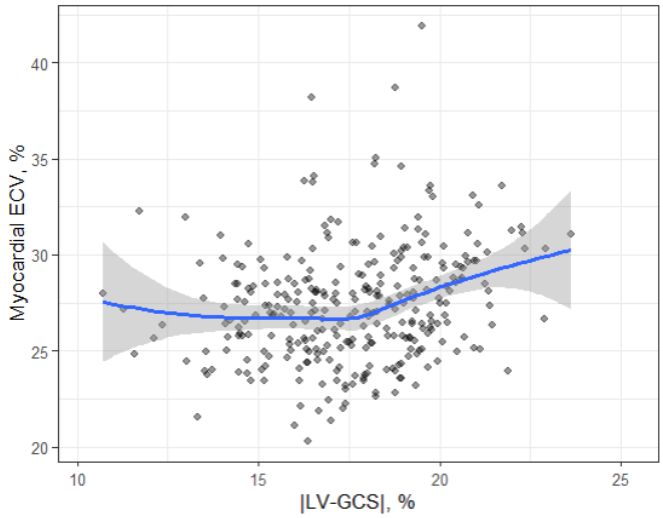
Accordingly, while there are no individual genes associated with *all* CMR metrics, this is not entirely unexpected or biologically inconsistent, given each metric truly reflects a different axis of cardiac remodeling, each likely having varying temporal and mechanistic contributors.

6. Results, page 9. PCCB was associated with a lower hazard for incident HF, positive association with larger LA volume and fibrosis, but better GCS. Why the dichotomous association between strain and fibrosis for PCCB?

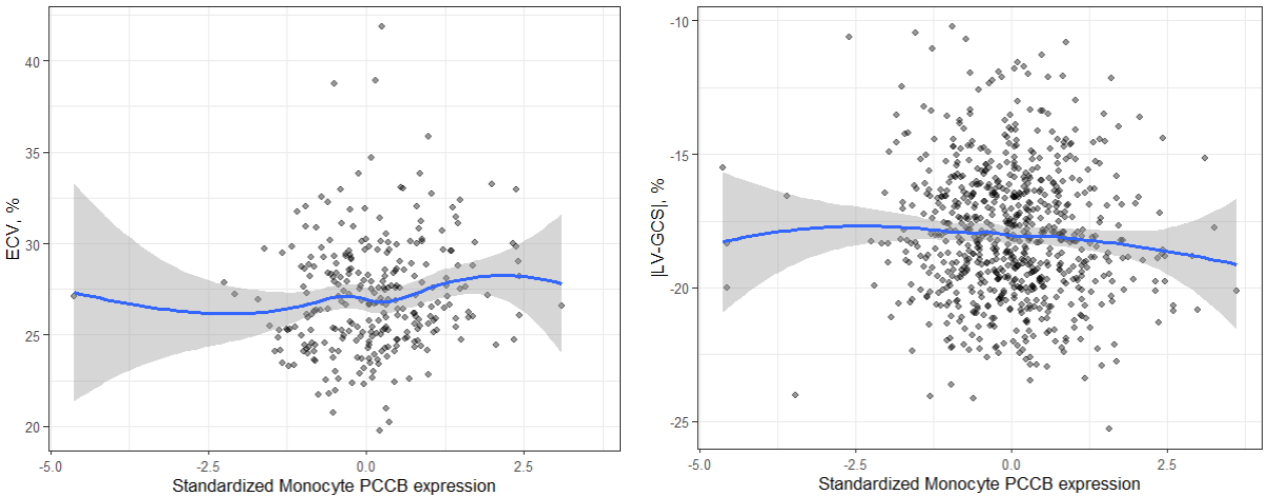
   This was a surprising result that piqued our interest, as well. This is related to the correlations of CMR metrics in our data we present above—specifically, in this study population that underwent highly standardized CMR imaging and interpretation, higher ECV was unexpectedly correlated with higher (*less adverse*) |LV-GCS| (rho=0.29). We initially posited that there may be confounding by LV geometry. We evaluated this using a linear regression model of the relationship between ECV and |LV-GCS|, adjusting for LVM/V, and we found no attenuation or directional change of the unexpected association.

Covariate Set	**Mean Difference in |LV-GCS| per 1% increment in ECV**	**p-value**
ECV, study site, age, sex, race, ethnicity	0.13 (95% CI: 0.04 to 0.21)	0.004
ECV, LVM/V, study site, age, sex, race, ethnicity	0.11 (95% CI: 0.04 to 0.19)	0.004

   To further explore this relationship, we then fit a Loess curve of the relationship between ECV and |LV-GCS| and observed somewhat of a J-shape, suggesting it is participants in the higher end of |LV-GCS| and ECV distributions driving this observation.



We also further explored the associations between *PCCB* expression and ECV compared to *PCCB* and GCS (below), and we found the ECV relationship was a bit less stable (i.e., perhaps non-linear), while the association between *PCCB* and LV-GCS (which had directional agreement with incident HF analyses) was more stable and linear. However, in evaluating whether the relationship between *PCCB* expression and ECV was perhaps better modeled using a spline, we found no evidence of improved model fit over our linear models (where spline DF=3, ANOVA testing model fit improvement over linear model: p=0.34; where spline DF=4, ANOVA p=0.49).



Taken together, one possible interpretation of these data is that GCS may not always monotonically worsen with higher ECV among this asymptomatic community-based cohort with largely preserved LVEF (median: 61%, IQR: 56–66%), particularly at this single time point. Among patients with non-ischemic dilated cardiomyopathy, the correlation between ECV and GCS is weak at r=0.38 (Azuma, et al. *Magn Reson Imaging* (2020). PMID: 32898651). There may be compensatory increases in regional strain in segments less affected by interstitial fibrosis, which can frequently be patchy in distribution (Glashan, et al. *Eur Heart J* (2018). PMID: 29617764), thereby preserving global GCS. These factors may contribute to differential association shapes and directionality with monocyte PCCB expression. Therefore, the remodeling axes captured by GCS (and LVEF, which had an estimate with the same directionality, p=0.01, FDR=0.14) vs. LAVI and ECV are, though related, may be biologically distinct at this time point in this population. We discuss this in the Discussion (page 17, lines 357-373):

> *"Such discordance in monocyte transcript associations across cardiac phenotypes may reflect differences in pathophysiology governing anatomic interstitial fibrosis which may generally precede cardiomyocyte contractile dysfunction—thus highlighting that, though related, the processes are distinct with different time courses.[33-35] For example, …"*

And now further in Discussion of all PCCB results, specifically (page 20, lines 434-437):

> *"Taken together, these results suggest monocyte PCCB expression likely does not represent a directly protective or harmful causal factor but instead may reflect a complex immunometabolic response that tracks with both adaptive functional preservation and underlying structural remodeling."*

7. PCCB, MTCP1, and VIM expression exhibited low pairwise correlations – can the authors discuss the implications of this weak association.

   This is an interesting point which we did not previously highlight. We believe this suggests each transcript may reflect complementary, instead of redundant, associations. We took this opportunity to explore this further by assessing whether these transcripts predicted incident

HF independent of each other (i.e., in one combined Cox model), and the results further supported this hypothesis. Shown below are the original presented estimates when transcripts are modeled individually in separate models, compared to these new estimates when modeled together in one model. While there was mild attenuation of point estimates and lower precision, each remain significantly associated with time to incident HF, independent of each other **and** important confounders and risk factors (Model 2 covariates).

Transcript	***Modeled Separately***		***Modeled Together***	
	HR (95% CI)	**p-value**	**HR (95% CI)**	**p-value**
PCCB	0.59 (0.44 to 0.79)	5.5E-04	0.70 (0.51 to 0.97)	0.034
MTCP1	0.59 (0.42 to 0.81)	1.1E-03	0.65 (0.45 to 0.94)	0.023
VIM	1.68 (1.11 to 2.53)	0.014	1.67 (1.07 to 2.61)	0.025

We have added this finding to the Results section (page 13-14, lines 276-279):

> *"Of note, circulating monocyte transcript levels of PCCB, MTCP1, and VIM transcript levels each predicted incident HF in a single combined model—i.e., independent of each other—with only mildly attenuated point estimates (PCCB aHR: 0.70, 95% CI: 0.51–0.97; MTCP: 0.65, 0.45–0.94; VIM: 1.67, 1.07–2.61)."*

We have also added a note on these results to the Discussion (page 19, lines 409-411):

> *"In our data, levels of these transcripts exhibited low pairwise correlations and independently predicted incident HF, suggesting they may reflect complementary rather than redundant associations."*

8. On page 13, the authors state: " poorer left ventricular systolic function (LV-GCS, LVEF) and greater concentricity (LVM/V) was negatively associated with genes related to mitochondrial ATP synthesis, oxidative phosphorylation, and the TCA cycle, possibly suggestive of more pro-inflammatory activity." This wording is confusing – please revise.

   We have now separated this sentence into two statements, which we hope and believe are much clearer (page 16, lines 347-351):

   > *"Here we found that poorer left ventricular systolic function (LV-GCS, LVEF) and worse concentric remodeling (LVM/V) was inversely associated with genes related to mitochondrial ATP synthesis, oxidative phosphorylation, and the TCA cycle. This metabolic pattern is suggestive of reduced oxidative capacity, which has been associated with a more pro-inflammatory monocyte phenotype.[9,31]"*

9. Please include the gene set enrichment analyses related to oxidative stress pathways in the main figures rather than the supplement.

   We agree that the GSEA results related to oxidative stress are quite interesting. The *specific* oxidative stress-related GO terms enriched across CMR parameters varied. Our primary GSEA figure (Figure 3C) specifically filtered for GO terms associated with *both* of our functional CMR metrics (LV-GCS and LVEF), as noted in the Results and figure footnote.

There were unfortunately no specific oxidative stress terms enriched for both LV-GCS and LVEF, though they each had individual related terms. Comparisons are included below.

CMR Outcome	Enriched GO Term(s) (FDR<0.05) Related to Oxidative Stress
LV-GCS	direct: response to oxidative stress, cell death in response to oxidative stress; other related terms: nitric oxide metabolic process, response to hypoxia
LVEF	direct: cell redox homeostasis; other related terms: nitric oxide transport, cellular iron ion homeostasis (e.g., via Fenton reaction)

For consistency, we hope you agree it is preferable to leave these terms out of Figure 3C. We have, however, edited the Discussion section on this topic to more directly highlight what processes were enriched, and address this variation in statistically enriched terms (pages 17-18, lines 375-379):

> *"Specifically, GO terms including "response to oxidative stress" were statistically enriched among results of transcripts associated with LV-GCS (FDR=0.01), LVM/V (FDR=0.01), and LAVI (FDR=0.048)—and other related terms were enriched among results for LVEF (e.g., "cell redox homeostasis" FDR=0.048) and ECV (e.g., "cellular oxidant detoxification" FDR=0.02)."*

We also then go on to describe another closely related process enriched among results for LAVI and ECV—glutathione metabolism—in the same Discussion section (lines 385-390).

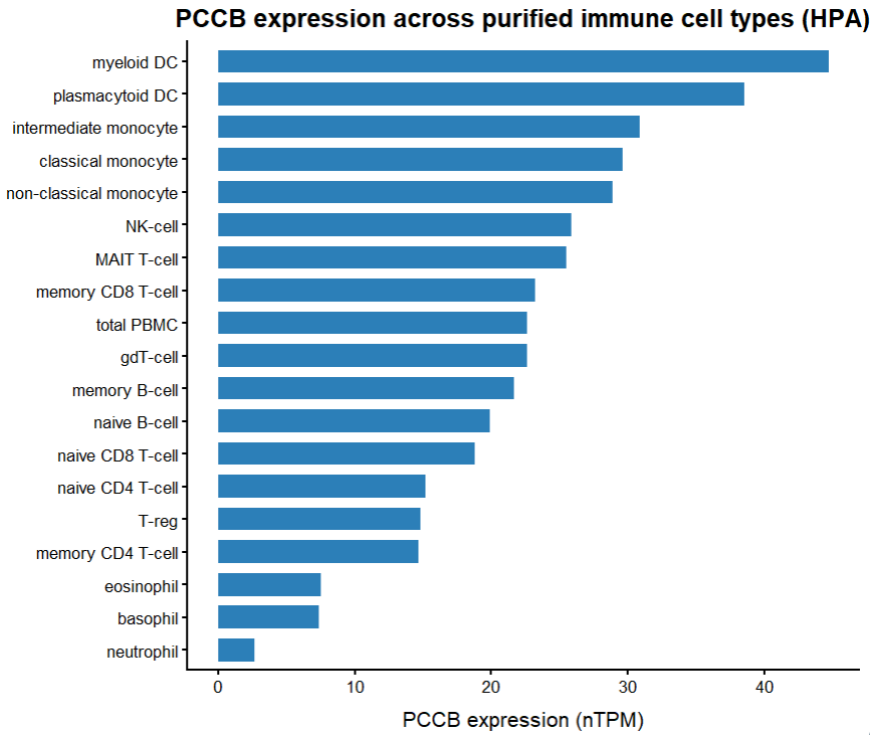
Reviewer #2 (Remarks to the Author):

This is an interesting manuscript about the role of monocyte in heart failure. The authors used transcriptomic as well as cardiac mri images and used those data to identify transcriptomic profiles of patients that associate with heart failure incidence, cardiac function and fibrosis. The data are sound, the topic is of interest and the manuscript well written. I have only one suggestion to improve the quality of the manuscript.

1. The authors should validate the expression of their top targets in isolated cells- ideally from a separate cohort- or at least in the isolated monocytes used for the study and compare the expression also to isolated cells from heart failure patients.

   We certainly agree additional layers of evidence would be valuable in supporting the validity of our observations. While it is currently not feasible for our group to pursue experimental validation on MESA samples or new samples from HF patients, we took this opportunity to identify additional approaches to support our observations. We have completed MR analyses using isolated peripheral monocyte-specific *cis*-eQTLs, as reported in the original version of the manuscript, with findings presented in the Results (page 14, lines 281-294) and ***Supplemental Data S16-S18***. This *cis*-MR analysis highlighted monocyte *PCCB* expression as a particularly strong candidate for further investigation and was thus the focus of our additional assessments.

First, we confirmed detectable expression of *PCCB* in isolated human monocytes across multiple independent studies using publicly available immune cell RNAseq data available via the Human Protein Atlas (HPA). Relative to other isolated human immune cell types, we found that PCCB expression was high in monocytes as well as myeloid dendritic cells, some of which can be derived from monocytes. We have added the figure below to the ***Supplementary Information*** file (***Figure S6***):



With accompanying legend text:

> "*Gene expression from the Human Protein Atlas immune cell atlas is reported as normalized transcripts per million (nTPM) and represents the average expression of PCCB across depicted isolated immune cell types. DC=dendritic cell; NK=natural killer; MAIT=mucosal-associated invariant. Data are publicly available from https://www.proteinatlas.org/humanproteome/single+cell/immune+cell/data*."

As well as corresponding text additions to the Results (page 14, lines 295-303):

> *"PCCB Expression Across Immune Cell Types*
>
> The Human Protein Atlas immune cell resource integrates RNAseq data on n=109 samples of sorted peripheral immune cells from predominantly healthy donors in three independent studies.[27] Using these previously published data, we confirmed detectable expression of *PCCB* in isolated human peripheral monocytes. In comparing relative distributions of normalized transcripts per million (nTPM) across 18 immune cell types, we found *PCCB* expression was relatively high in monocytes compared to other isolated human immune cell types, as well as in myeloid dendritic cells (DC), some of which can be derived from monocytes (***Supplementary Information Figure S6***).

To our knowledge, there are no previously published gene expression data (microarray or RNAseq) on isolated peripheral monocytes from heart failure patients and controls. We therefore next sought out previously published single-cell RNAseq data on blood from heart failure patients and controls. We identified 4 publications:

- Abplanalp, et al. *Cardiovasc Res* (2021); PMID: 32311026. Peripheral blood CD31+ cell scRNAseq data from 8 HF patients and 8 controls were unfortunately not made publicly available by the authors.
- Kneuer, et al. *Circ Res* (2025); PMID: 40709473. Peripheral blood mononuclear scRNAseq data from patients with HFpEF (n=6), patients with HFrEF (n=8), and healthy controls (n=7) were unfortunately not made publicly available by the authors. They provide some data from their murine models and experiments involving human PBMCs and treatment with nitro-oleic acid (NO_2-OA), which were not relevant for our investigation.
- Rao, et al. *Basic Res Cardiol* (2021); PMID: 34601654. Deposited peripheral blood scRNAseq data from this study included a single control sample, preventing us from meaningfully comparing peripheral cell *PCCB* expression between disease states.
- Wang, et al. *Explor Cardiol* (2025); https://doi.org/10.37349/ec.2025.101247. Microarray expression data on isolated PBMCs from patients with (i) stage 2–3 chronic kidney disease, (ii) end-stage renal disease (ESRD), (iii) ESRD with HFpEF, and (iv) ESRD with HFrEF. These data were made available by the authors. Patients with HFpEF or HFrEF (groups iii/iv, n=10) compared to patients without HF (groups i/ii, n=10) had significantly lower peripheral mononuclear cell expression of *PCCB* (*p*=1.63E-04). This is consistent with our observations in the MESA cohort, showing lower monocyte expression of *PCCB* was associated with poorer systolic function (by LV-GCS) and higher hazard of incident heart failure. However, the data does not represent monocyte-specific gene expression (instead total PBMC expression). Furthermore, the controls represented a mix of milder kidney disease (stage 2-3) and end-stage renal disease, while all of the HF patient samples had end-stage renal disease, likely confounding the comparison. So while this does provide some support of PCCB-related peripheral immune perturbation in this more severe cardiorenal clinical context, we do not believe it represents a clear external validation of our findings in the community-based MESA cohort. We hope you agree with our decision to therefore leave these data out of our revised manuscript.

Reviewer #3 (Remarks to the Author):

The authors explored monocyte gene-expression data from atherosclerosis-free participants in the MESA study and conducted both targeted and untargeted exploration of associations between expression data and CMR measurements. The authors identified several overlaps between CMR features and gene expression in monocytes, and some of these genes were associated with clinical heart-failure (HF) outcomes. The manuscript is well written, and the data presentation is sensible. My comments below.

[Major]

1. The main question in this paper is whether a transcript associated with more than two cardiac MRI features is more important than transcripts that are more strongly associated with a single measure or with HF outcomes. Using “>1 features” might have been a good filter to select a tractable number of genes, but I could not find a reasonable biological rationale for focusing on this criterion.

   We thank the reviewer for this valid critique and the opportunity to clarify our motivation for this criterion. We intentionally chose to evaluate CMR metrics that capture complementary (rather than directly overlapping) biological axes of cardiac remodeling, i.e., structural remodeling (LVMI), contractile function (LV-GCS), chronically elevated LV filling pressure (LAVI), and expansion of myocardial extracellular space (ECV). Our data support this distinction, with evaluated CMR metrics showing moderate to low correlation in our study (all |*rho*|<0.36, shown below). For this reason, we believe individual cardiac MRI metric results are important.

	LVEF	|LV-GCS|	LVM/V	LAVI	ECV
LVEF	--	0.24	-0.03	0.04	0.07
|LV-GCS|	--	--	-0.36	0.23	0.29
LVM/V	--	--	--	-0.33	-0.25
LAVI	--	--	--	--	0.19
ECV	--	--	--	--	--

   However, there is a substantial number of population-based imaging studies that show incremental risk as abnormalities accumulate across different domains of cardiac structural and functional remodeling, i.e., a multi-domain burden is more prognostically informative than any single abnormal imaging metric in isolation (Shah, et al. *Circulation* (2017). PMCID: PMC5241178 ; Mordi, et al. *J Cardiovasc Magn Reson* (2015). PMCID: PMC4328968 ; Kalogeropoulos, et al. *JACC Cardiovasc Imaging* (2012). PMCID: PMC3293933). For this reason, we believed focusing incident HF analyses on transcripts associated across multiple CMR metrics would prioritize signals that may reflect more coordinated subclinical remodeling. Our manuscript does not currently articulate this well, so we have added the following text (including references cited above) to our Methods and Results.

   Methods (addition underlined, page 9, lines 161-163):

   > *“Evidence from population studies suggests the accumulation of abnormalities across cardiac structural and functional domains is more prognostically informative than singular metrics.[20-23] Given this, only transcripts cross-sectionally associated with more than one CMR parameter with a two-sided false discovery rate (FDR)<0.05 and concordant directionality were further assessed for associations with time to incident adjudicated heart failure.”*

   Results (addition underlined, page 13, lines 258-260):

   > *“To prioritize signals that may reflect more coordinated and prognostically informative subclinical remodeling, these 96 transcripts associated with more than one cardiac characteristic at FDR<0.05 were then evaluated for associations with time to incident HF.”*

2. The associations among PCCB, expression QTLs, CMR measures, and heart-failure (HF) outcomes are not explicitly shown and are not readily traceable for readers. A clear connection is needed across genetic variation (eQTL), expression–endophenotype (CMR), and expression–outcome (HF). On a related note, were the CMR measures actually associated with HF incidence?

   We agree a summary table for PCCB associations would be very helpful for readers. We have added the following as a new table, **Supplementary Data 18**, which is now referenced on page 14, line 294 of the manuscript.

Outcome	Estimate	*p*-value	FDR	Signal
Cross-Sectional (Mean Difference in CMR per SD Increment in PCCB)				
|LV-GCS|	0.31 (0.13 to 0.48)	5.72E-04	0.002	protective
LVEF	0.69 (0.16 to 1.22)	0.010	0.14	protective
LVMV	-0.020 (-0.037 to -0.003)	0.024	0.37	protective
LAVI	1.53 (0.64 to 2.42)	7.84E-04	0.002	harmful
ECV	0.52 (0.17 to 0.86)	3.17E-03	0.004	harmful
Time to Incident Heart Failure (Hazard Ratio per SD Increment in PCCB)				
Incident HF	0.59 (0.44 to 0.79)	5.46E-04	0.002	protective
Two-Sample MR Association of Lifetime Genetically Higher PCCB with Prevalent HF (Odds Ratio)				
Prevalent HF	1.17 (1.03 to 1.35)	0.014	NA	harmful

   We also now included the MR odds ratio estimate directly in Results section text (line 292) and now better discuss this interesting directionality in the Discussion (pages 19-20, lines 422-437)—which further qualifies considerations around causality and includes an additional reference describing an example of a similar directional pattern observed with the immune mediator IL-6 and coronary heart disease (Swerdlow, et al. *Lancet.* 2012. PMCID: PMC3316968).

   Old text:

   > *"Monocyte expression of PCCB had an inverse association with hazard of incident HF in this study, consistent with its inverse association with poorer LV function measured by LV-GCS. We also found evidence of an association between the leading monocyte-specific PCCB cis-eQTL and incident HF in a large multi-ancestry HF GWAS, strengthening its candidacy for future investigation in HF."*

   Updated text:

   > *"In this study, higher monocyte expression of PCCB was cross-sectionally associated with a mixed cardiac phenotype characterized by better systolic function measured by LV-GCS (and suggestively LVEF at p=0.01, FDR=0.14), alongside more adverse structural remodeling measured by LAVI and ECV. Consistent with the functional metrics, higher PCCB expression was associated lower risk of incident HF over ~8 years. In contrast, cis-MR suggested that higher genetically predicted monocyte PCCB*

expression was associated with greater odds of HF. Our MR analysis aimed to capture genetically predicted lifetime monocyte gene expression levels, whereas monocyte gene expression measured in the periphery at a single time point reflects much more complex biology affected by both genetic and environmental factors. Similar patterns of discordant directionality have been described across the literature and been linked to compensatory feedback biology, including for the classic inflammatory marker IL-6 and its relationship with coronary heart disease.[51] Taken together, these results suggest monocyte PCCB expression likely does not represent a directly protective or harmful causal factor, but instead may reflect a complex immunometabolic response that tracks with both adaptive functional preservation and underlying structural remodeling."

To your question regarding the relationship between CMR parameters and time to incident HF, they are indeed significantly predictive in MESA. While this is outside the scope of the approved MESA concept for this manuscript, we present these estimates here, using the same exam 5 CMR measurements and incident HF following exam 5. This analysis sample is not restricted to participants with monocyte transcriptomics to maximize statistical power.

CMR Input	HR for Incident HF per SD Increment in CMR (95% CI)	*p*-value
|LV-GCS|	0.61 (0.49 to 0.77)	1.5E-05
LVEF	0.65 (0.53 to 0.79)	3.1E-05
LVMV	1.21 (1.01 to 1.46)	0.043
LAVI	1.55 (1.32 to 1.83)	1.9E-07
ECV	1.69 (1.11 to 2.57)	0.014

[Minor]

1. Abstract Line 9
"PCCB was associated with >1 cardiac characteristic" This is a bit difficult to follow. Ideally, the authors should provide the actual cardiac MRI measures (e.g., "XXXX was associated with LAVi, LV-GCS, and ECV," etc.).

   We have edited the abstract as suggested. This sentence now reads (page 2, lines 16-18):

   "Monocyte expression of PCCB was associated with left ventricular strain, left atrial size, myocardial fibrosis, and time to incident HF, and was validated using Mendelian randomization with peripheral monocyte-specific cis-eQTLs."

2. Line 79
"The analysis sample for incident HF—which did not require CMR—was n=1195 (Supplemental Figure S1B)." How many participants overlap between the CMR/transcriptomic analysis and the incident-HF analysis? Does this overlap (or lack thereof) affect the results and their interpretation?

Participants included in the CMR analysis are a direct subset (a MESA sub-study) of participants included in the incident HF analysis. We have clarified this with a brief text addition to the beginning of the Results (page 10, lines 188-189), as well as the footnote of Table 1:

> *"Participants included in cross-sectional CMR analyses are a direct subset of participants included in incident event analyses."*

3. Line 133 - 143
   The authors list 54 genes among 96 genes associated with CMR phenotypes, yet none of the pathways are enriched in these 96 genes. It is a bit difficult to follow the feasibility and interpretability of this analysis. Such descriptive, post-hoc reasoning does not clarify what is newly learned in this study. I was wondering whether, under either a more stringent threshold (e.g., FDR < 0.01 or Bonferroni) or a more lenient one (e.g., FDR < 0.20), any pathways became enriched.

   We certainly agree the lack of statistical enrichment among the 96 individual genes associated with more than one CMR metric may make interpretation of biological overlap more difficult. We see two key points of consideration here.

   First, we do believe this lack of statistical enrichment is in part related to low statistical power for this particular enrichment test (as opposed to true lack of biological relatedness) and henceforth why we chose to still describe some patterns in Results text (pages 12-13, lines 245-256). For example, the GSEA applied to the results of our incident HF analysis yielded several GO terms with enrichment at an FDR<0.10, shown below (instead of FDR<0.05, which we report in accordance with our *a priori* threshold). Also shown here is that, because the list comprises only 96 evaluated genes (instead of the full 9,430 genes used in GSEA tests for individual CMR metrics), the power for GSEA tests (evidenced by small 'setSizes') was very low for most processes.

ID	Description	setSize	NES	p-value	FDR
GO:0051641	cellular localization	14	2.078888	0.001105	0.059775
GO:0007010	cytoskeleton organization	7	1.928939	0.002031	0.059775
GO:0010628	positive regulation of gene expression	6	1.953126	0.002756	0.060881
GO:0007178	transmembrane receptor protein serine/ threonine kinase signaling pathway	3	1.743671	0.004727	0.087544
GO:0007155	cell adhesion	6	1.879532	0.005539	0.087544
GO:0048856	anatomical structure development	19	1.876543	0.005566	0.087544
GO:0045087	innate immune response	7	1.726583	0.010779	0.098185
GO:0006935	chemotaxis	3	1.631898	0.013258	0.101342

   Related, because GSEA applied to results of CMR analyses had much higher power, patterns in enrichment across those model sets are especially important. We indeed identified several patterns of enrichment across CMR metrics, which we describe in the Results (page 12, lines 233-243), throughout the Discussion, and present partially in **Figure 3B** and fully in **Supplementary Data S10-S14**.

Second, we intentionally chose to evaluate CMR metrics that capture complementary (instead of directly overlapping) biology, e.g., structural remodeling (LVMI), contractile function (LV-GCS), chronically elevated LV filling pressure (LAVI), and expansion of myocardial extracellular space (ECV). Our data support this distinction to some degree, with evaluated CMR metrics showing moderate to low correlation in our study population (all |*rho*|<0.36).

	LVEF	|LV-GCS|	LVM/V	LAVI	ECV
LVEF	--	0.24	-0.03	0.04	0.07
|LV-GCS|	--	--	-0.36	0.23	0.29
LVM/V	--	--	--	-0.33	-0.25
LAVI	--	--	--	--	0.19
ECV	--	--	--	--	--

Lastly, we did evaluate how results of signal overlap changed when applying different FDR thresholds to our CMR regression analyses, as you suggest, and we found that relaxing the threshold in our CMR filtering step allowed for detection of several similar enriched GO processes to those we report in GSEAs of individual CMR models.

CMR Model Threshold	**N Genes Associated with ≥2 CMR Metrics, Concordant Direction**	**GO Processes Enriched in Incident HF Model 2 Results, GSEA (Pathway FDR <0.05)**
FDR ≤0.20	1311	**433 GO terms**, including: tissue development, MAPK cascade, myofibril assembly, cell-cell adhesion, cell migration, oxidative phosphorylation, mitochondrial gene expression, T cell activation, mononuclear cell differentiation
FDR ≤0.10	215	**44 GO terms**, including: tissue development, MAPK cascade, myofibril assembly, and cell migration
FDR ≤0.01	0	NA

While we did not feel able to justify setting a more lenient threshold for the manuscript, we believe these observations further support the argument that statistical power and thresholding throughout the analytic pipeline impacted our ability to detect statistically enriched patterns in these overlapping transcripts. We do, however, believe keeping our threshold to the *a priori* level, though conservative, is more appropriate for reporting. In the Discussion (page 21, line 459), we highlight our power limitations, but we have now added a specific note as it relates to GSEA (underlined below):

> *"Type II error should also be considered in the interpretation of our results, as our statistical power was limited by sample size and the method used for multiple testing correction, which assumes independence of tests (i.e., biological independence of transcripts and biological process annotations used in GSEA)."*

4. Line 153 - 165
The authors provided directional explanations of these measures (adverse/protective), which helps readers who are not familiar with cardiac MRI measurements. However, in the sentence for XXX, the stated direction is opposite to that of the other genes and to the effect

on clinical outcomes. This should be aligned with the rest. For example: “Higher monocyte expression of MTCP1 was associated with better LV function,” etc.

We agree this would improve clarity, so we have edited the text to align our reporting (page 13, lines 271-273):

Old text:

“Lower monocyte expression of MTCP1 was associated with poorer LV function measured by lower LVEF (0.87% lower per SD decrement; 95% CI: 0.39 to 1.35; FDR=0.0004) and lower |LV-GCS| (0.24% lower per SD decrement; 95% CI: 0.40 to 0.07; FDR=0.006).”

Updated text:

“Higher monocyte expression of MTCP1 was associated with better LV function measured by higher LVEF (0.87% higher per SD increment; 95% CI: 0.39 to 1.35; FDR=0.0004) and higher |LV-GCS| (0.24% higher per SD increment; 95% CI: 0.40 to 0.07; FDR=0.006).”

5. Line 186
Given the modest association with HF (P = 0.01), the claim about “no evidence of significant pleiotropy was identified” reads as an overstatement, particularly when other outcomes exhibit stronger evidence (P < 0.01). I recommend tempering the language or contextualizing it relative to the other results

We agree and have edited the text to describe our observation with less assumptions and also directly acknowledge the limitation of the assessment (page 14, lines 287-290):

*“We did not identify any previously reported associations with traits that may suggest horizontal pleiotropy (**Supplementary Data 16**), though the possibility of unrecognized pleiotropic effects cannot be excluded.”*

6. Line 188 - 191
The authors utilize cis-eQTLs to evaluate associations between XXX and UUU expression and heart failure. The UUU eQTL was associated with heart failure (P = 0.01). I could not find information about the SNP’s effect direction (e.g., increased/decreased UUU expression and increased/decreased HF risk), nor could I find the SNP’s identifier. Please consider providing the rsID and chromosome/position and reference/alternate alleles. In addition, reporting allele frequency and the effect size/direction for both the eQTL and the heart-failure GWAS would aid interpretation.

We have added rsID and chromosome/position more clearly to the Online Methods (page 6): *“Selected IVs were leading eQTLs for monocyte expression of candidate genes (PCCB: rs9830653, (hg19)3:136443219; and VIM: rs1707279, (hg19)10:17009650).”*

We have also added details on eQTL and HF GWAS effect estimates with effect allele frequencies as a new table, **Supplementary Data S16**:

Analysis	Effect Allele	Other Allele	EA Frequency	Beta	SE	*p*-value
PCCB (ENSG00000114054): rs9830653 leading cis-eQTL						
eQTL	T	C	0.5665	-0.0796	0.00569	2.76E-41
HF GWAS	T	C	0.5592	-0.0131	0.00538	0.0146
VIM (ENSG00000026025): rs1707279 leading cis-eQTL						
eQTL	A	G	0.3151	0.0402	0.00627	2.23E-10
HF GWAS	A	G	0.3241	0.0062	0.00472	0.1877

7. Line 191 needs line break

   Added.

8. Figure 3C
   The results need additional clarification. The density plot is difficult to interpret because the underlying elements are not defined. Please specify: (i) whether the plotted units are individual genes within each pathway, (ii) the number of genes per ontology, (iii) whether the gene set was pre-filtered by association P values with the relevant CMR measures (and the threshold used), and (iv) whether the reported FDR corresponds to pathway-level enrichment significance.

   Thank you for this very helpful feedback and the chance to improve Figure 3C clarity. Each of your questions are answered below and have corresponding additions to the legend text.

   i. This is a nice opportunity to clarify, thank you. The plots are kernel densities of the ranked β estimates from our models (x-axes), so the y-axes represent the density of mapped genes at a given β estimate, and the total area under each curve is ~1. The y-axes are therefore unitless and relative within gene sets mapping to the indicated GO: Biological Process annotation on the left. This is the classic representation of GSEA in ridge plots, and we now clarify this in the figure legend text (addition underlined): *"Ridge plots depicting the kernel density distributions of association estimates…"*

   ii. The number of genes mapping to each process are now listed in parentheses following the annotation descriptor—very helpful.

   iii. The gene set was not pre-filtered by association p-values, as this would side-step one of the core advantages of GSEA. GSEA evaluates enrichment using the full ranked list of evaluated genes, allowing for more sensitive detection of modest but coordinated effects. Pre-filtering can therefore actually bias GSEA results and lower statistical power by removing those genes with modest but coordinated associations. This being said, we did intentionally verify all annotations had at least one mapped gene with an indicated CMR association FDR<0.05; most annotations had many mapped genes meeting that model significance threshold.

   iv. The reported FDR does indeed correspond to pathway-level significance, adjusting for multiple testing of ~8600 mapped GO annotations. Legend text now reads: *"All presented annotations are statistically enriched by gene set enrichment analysis with a pathway-level FDR<0.05."*

9. Figure 4B
   This figure is difficult to interpret in its current form. Given the number of feature intersections, an UpSet plot would communicate the set overlaps more clearly and compactly (see https://en.wikipedia.org/wiki/UpSet_plot).

   Based on the Reviewer's reference to feature intersections, we believe they may be referring to Figure 4A (instead of Figure 4B), in which case we very much like this suggestion. We have replaced the Venn diagram in Figure 4A with an UpSet plot (shown here), which we agree more cleanly depicts feature overlap (both associations have FDR<0.05 and concordant directionality).