

Diet-responsive proteogenomic effects following short-term restriction of animal products in humans

Corresponding Author: Dr Antigone Dimas

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)

Simistiras et al reports a longitudinal study exploring how periodic restriction of animal products influences the genetic regulated plasma protein expressions. The longitudinal design, combined with the comparison to a continuously omnivorous group, is a clear strength. However, several critical issues limit confidence in the reported findings and should be addressed to improve the robustness and interpretability of the manuscript:

Major comments:

1. The baseline (T1) plasma proteome profiling as well as other clinical measurements should be compared between PR and TR group. Since the dietary pattern of the PR group was initiated during childhood, it is very likely that their long-term dietary habits have already shaped their plasma proteome before the study started. Without baseline comparisons, it is difficult to distinguish acute diet-induced changes from chronic, pre-existing group differences.
2. The identification of cis-pQTLs could be much improved by directly modeling genotype × diet interactions within a mixed-effects framework. This would provide a more formal test of whether genetic effects on protein levels differ by diet, rather than relying on separate analyses.
3. Since this is a longitudinal study, the analysis should prioritize within-individual changes across time (T1 vs. T2) rather than group-based comparisons. Also, the baseline heterogeneity should be considered as confounding factors when performing this kind of analyses.
4. The comparisons of BMI and other clinical markers (e.g., GGT, eosinophil counts, cholesterol) between T1 and T2 should be presented for both PR and TR groups. Many plasma protein changes are known to be associated with BMI and other physiological traits. If BMI also increased in the PR group at T2, this could cause the changes of plasma proteins that are unrelated to diet. At present, it is unclear which confounders were included in the statistical models.
5. Significant seasonal variation was observed in the results. Season should therefore be explicitly modeled as a covariate to avoid confounding. Without this adjustment, it is not possible to separate diet-related effects from seasonal fluctuations in plasma protein levels.
5. The health outcome measures are not clearly defined. The manuscript often refers to potential links with obesity, liver function, or immune-related diseases, but no direct clinical endpoints were reported in this cohort. The authors should clearly state the observational variables in this study.
6. It is unclear whether dietary intake was monitored during the study period. Were dietary records, recalls, or surveys collected to confirm adherence to the fasting practice? A more detailed description of the dietary protocol is needed for interpreting the observed changes.
7. The discussion around SERPINB1 and neutrophil counts is overstated. While SERPINB1 is known to be involved in the regulation of neutrophil survival, the manuscript does not demonstrate a direct association between SERPINB1 abundance and neutrophil counts in this study as well as the influence on the immune system.

8. Given the small sample size of this study, any power analysis has been done before performing the MR and colocalization analyses? For example, I notice that only a few cis-pQTLs can explain substantial variance, most are weak instruments.

9. In general, there are a lot of speculative statements in the results section. For example, “genotype-dependent seasonal variation may contribute to individual differences in seasonal susceptibility to these diseases.”, “these signalling cascades may aid the identification of downstream effectors of FGF21 that could constitute potential therapeutics.” They should be moved to discussion. Also, the authors should be more cautious about the clinical interpretation and overstatements.

10. The authors should also explicitly state whether genetic ancestry PCs were included in all models to avoid population stratification bias

Minor comments:

1. More details in missing data handling (expression below LOD) in proteome profiling should be provided
2. Any multiple testing correction methods have been included in the genetic analysis to control the false discovery?
3. The report of sample sizes after QC across text, tables, and figures should be consistent

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

Simistiras et al. evaluate the effects of diet on genetic regulation of proteins using proteomic and genetic data on two samples from the FastBio study: individuals who alternate between omnivore and vegetarian diets, and individuals who are consistently omnivore. The study design is clever with samples from two time points for each sample to evaluate both differences in animal products in diet as well as seasonal differences. Overall, I think this is a fascinating and largely new area of study with a lot of translational potential given that diet is modifiable while the genome is not. My main concerns with the manuscript are a lack of clarity in how differential mapping was assessed as I outline further below.

Major Comments

- Please provide more information about differential cis-pQTL mapping using mashr and the other tests performed in the methods section. Even though the workflow is cited, since this is a or indeed, the, central analysis in the paper, I think all necessary information to understand the analysis should be included. In addition to more detail in general, some specific questions include
 - o What does a training set mean in this context? In general, I think of a training dataset as a subset of individuals, but it appears that this training set is the total number of summary statistics. For the training set, what information is provided? Summary statistics? If so, which summary statistics? More clarity here will benefit the reader.
 - o What groups are the mixture weights for? The four groups (PR1, PR2, NR1, NR2)? Are the mixture weights estimated on average across all tests and then cis-pQTLs that differ from those mixture are identified? Something else? I don't think a lot of additional language is needed here, just half a sentence to clarify.
 - o How is local false sign rate determined? What does it mean? Again, a one sentence explainer should suffice with the citation.
 - o What does “manually tested” mean?
 - o How is the statistical test performed to assess the differential mapping?
 - o Were the comparisons adjusted for multiple testing?
- I really like the pirate plots! However, I think displaying the plots for small groups is misleading (e.g., ≤ 6) – perhaps just, or in addition, plot the individual points as well?
- Please always report p-values in text and figures even if not significant (e.g., text for rs74148404) and all figures. This helps provide an idea of the strength, or lack thereof, of the association.
- How were adjusted p-values calculated for figures?
- Line 357-358 – more information would be helpful about the imputation. What does “dietary group” refer to here? PR and NR only? It is not quite clear to me why we couldn't impute everything together and why that might result in bias. Alternatively, if imputing separately is important why not impute separately for the different seasons too?

Minor Comments

- How do the non-restricted and restricted PCs (genetic ancestry) compare? Can these be highlighted differently on the PC plot? Or if completed in another paper, please reference the results
- (line 98) It is not clear to me what TSS is – please define. Transcription start site?

Audrey Hendricks

(Remarks on code availability)

The github site appears to be thorough including detailed readme file and commented code. It is easy to navigate. I did not easily find the pirate plots, but that is a minor point. I did not attempt to run the code. While there are no explicit directions for installing and running the code, it is clear what is needed through the code dependencies and packages including version numbers and the software needed is commonly available.

Reviewer #3

(Remarks to the Author)

Simistiras and et al. studies the proteogenomic effects of short-term restriction of animal products in humans. it provides some important insights on the dietary restriction in 200 healthy humans. The authors considered this as "the first study to demonstrate that short-term, real-world dietary patterns can modify genetically driven protein abundance in humans and link these diet-responsive regulatory effects to broader health-related phenotypes." The same group has previously published another work based on the same human studies in NPJ Metabolic Health and Disease (2025), explores how short-term, real-world dietary patterns—specifically, periodic restriction of animal products—can modulate metabolic pathways and influence genetically driven protein abundance. in addition, a preprint at medRxiv is also largely on the same topic. The novelty becomes questionable, and the authors should address and highlight the distinct focus of these studies if there is any and how all these data can advance the science.

Since the study focused on diet restriction. short-term fasting while sleeping should be one important factor. are the samples analyzed from humans under fed or fasting states? would feed and fasting state affect the outcome of the analysis?

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

The authors have addressed my concerns, and I agree with publication in its current form.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

The authors have addressed all my comments and concerns.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)

Novelty remains to be a concern.

(Remarks on code availability)

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Diet-responsive proteogenomic effects following short-term restriction of animal products in humans

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

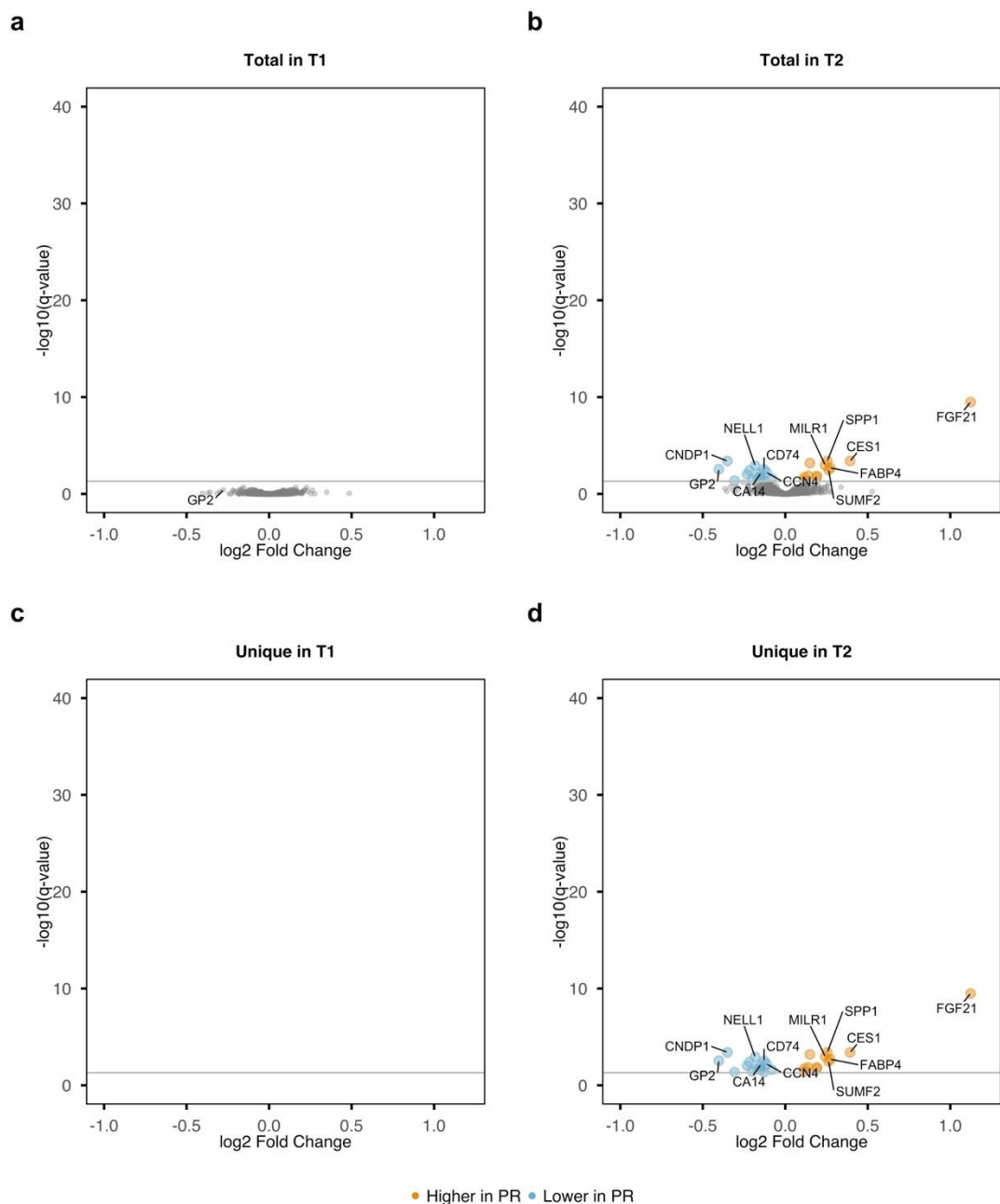
Simistiras et al reports a longitudinal study exploring how periodic restriction of animal products influences the genetic regulated plasma protein expressions. The longitudinal design, combined with the comparison to a continuously omnivorous group, is a clear strength. However, several critical issues limit confidence in the reported findings and should be addressed to improve the robustness and interpretability of the manuscript:

Major comments:

1.The baseline (T1) plasma proteome profiling as well as other clinical measurements should be compared between PR and TR group. Since the dietary pattern of the PR group was initiated during childhood, it is very likely that their long-term dietary habits have already shaped their plasma proteome before the study started. Without baseline comparisons, it is difficult to distinguish acute diet-induced changes from chronic, pre-existing group differences.

We thank the Reviewer for raising this key point and agree that potential pre-existing group differences stemming from long-term dietary habits may influence the results of a study like the present one. Indeed, this comparison has been described in Rouskas et al¹ (reference no. 9 in the manuscript <https://www.nature.com/articles/s44324-025-00057-2>), where we explored the impact of animal product restriction on plasma levels of metabolites and proteins, irrespective of genetics. In this study, we compared protein profiles between dietary groups at T1 (baseline) and found that there were no significant differences, suggesting that at baseline there are no long-term differences (Supplementary Text S3 and Supplementary Fig. S9 from Rouskas et al, also included below). At T2 however, we demonstrated extensive, diet-associated differences between the two groups.

To address this important comment and to make this finding more obvious in the present manuscript, we have added the following phrase to the population sample description in the Results (lines 102-105): “At the proteomic level, no significant differences were detected between groups when both followed an omnivorous diet (T1), however pronounced differences emerged at T2, when the PR group had practiced short-term animal product restriction.” See corresponding Figure from Rouskas et al¹ below.



Supplementary Figure S9 from Rouskas et al¹ showing that at T1 (baseline) there are no significant differences in protein abundance between the two dietary groups. Differences in protein abundance are evident only at T2, whereupon the group practicing dietary restriction had abstained from animal products for three-to-four weeks.

2. The identification of cis-pQTLs could be much improved by directly modeling genotype $\times$ diet interactions within a mixed-effects framework. This would provide a more formal test of whether genetic effects on protein levels differ by diet, rather than relying on separate analyses.

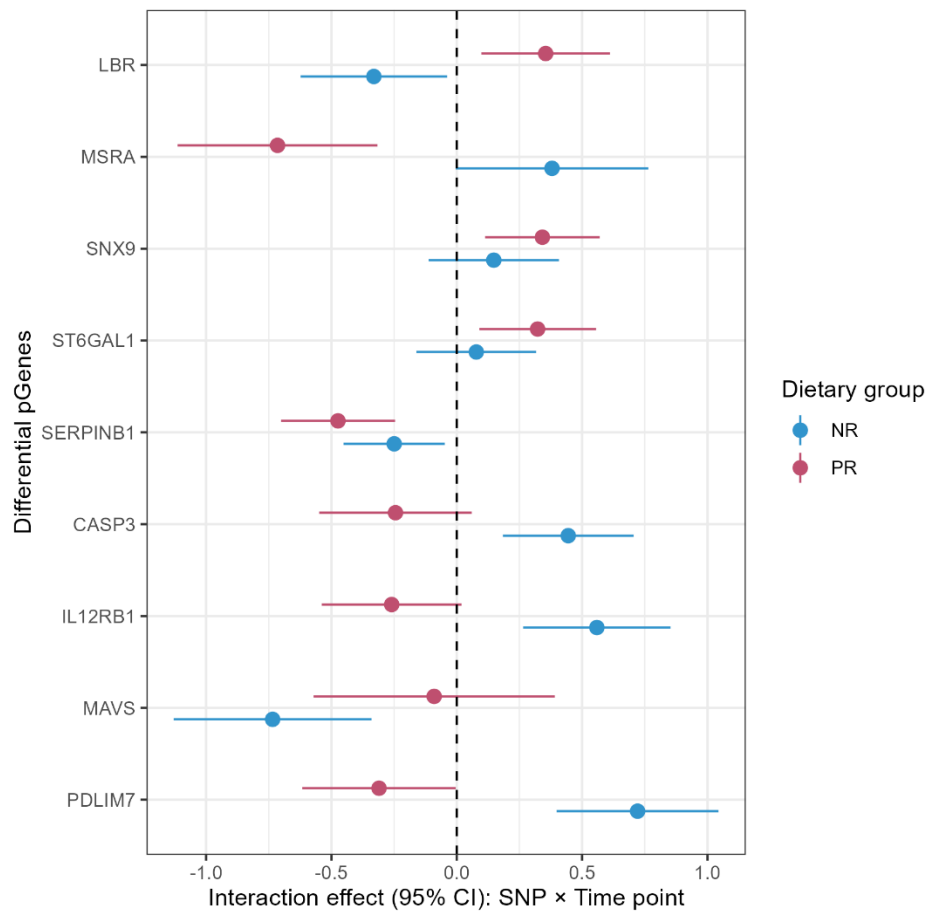
We thank the Reviewer for this comment and agree that genotype $\times$ diet interaction testing within linear mixed model frameworks is an established methodological

approach. However, the implementation of such models at scale, including thousands of proteins along with millions of SNPs in multiple contexts, presents computational challenges and may result in significant loss of statistical power particularly when interaction terms are modelled independently across highly correlated variants and molecular traits². To overcome these challenges, post hoc joint modelling approaches have been developed. These approaches, especially within a Bayesian framework, can leverage information across multiple phenotypes, SNPs and contexts improving effect estimates³ and enhance the statistical power to detect context-specific effects, instead of relying on independent interaction modelling^{3,4}.

The methodological approach employed in the present study (mashr) is a well-established analytical framework that relies on an empirical Bayes hierarchical model, to identify cis-pQTLs that respond to different environmental conditions⁵. In this case cis-pQTL respond to changes in dietary intake, proxying SNP x diet interactions. This approach, originally developed by Urrut et al⁵, has been applied in multiple recent studies examining context-dependent genetic regulation similarly to the present work⁶⁻⁹. For example, Lin et al⁸ applied mashr in a similar setting and uncovered hundreds of cis-eQTLs responding to a high cholesterol high fat diet in baboons. To clarify this point and to address the Reviewer's comment, we have: a) provided more detail on the modelling approach applied (Methods), and b) conducted genotype x time point (i.e. diet) interaction testing for differential cis-pQTLs as a sensitivity analysis.

a) The updated text in the Methods reads as follows: (lines 505-510) "Briefly, mashr uses multivariate mixture models to learn patterns of specificity and sharing of SNP effects across multiple conditions or contexts, while accounting for data-driven correlation between effect estimates, e.g. arising from sample overlap between contexts. Following that, mashr uses adaptive shrinkage improving estimates of effects and significance enabling direct comparison of SNP effects across conditions to define patterns of specificity and sharing, without the need of additional testing."

b) For the nine differential cis-pQTLs identified through mashr, we performed genotype x time point (i.e. diet) interaction testing as a sensitivity analysis. We found that all nine showed a significant interaction effect between time points with concordant direction of effect with mashr results, suggesting agreement between analytical methods (see forest plot below, now included as Supplementary Figure 3).



Supplementary Figure 3. Sensitivity analysis of differential cis-pQTLs. Forest plot showing effect estimates of SNP x time point interaction terms for each mashr-detected differential cis-pQTL, and for each dietary group. Points represent interaction effect estimates obtained from generalized linear models, and horizontal lines indicate the corresponding 95% confidence intervals. PR, periodically restricted group NR, non-restricted group.

We note that in case of LBR, the sensitivity analysis also yielded a significant interaction in the control group, but with opposite direction of effect compared to PR group. As discussed in the submitted manuscript, LBR levels in the control group were downregulated at T2. We suggested that this downregulation is not detected in the PR group because it is likely masked by carriers of the effect allele that is associated with increased levels of LBR at T2 (outlined in lines 196-203).

We have now incorporated the sensitivity analysis in the manuscript as follows:

Methods (lines 534-541): “As a sensitivity analysis, differential cis-pQTLs identified through mashr were further tested for interaction effects between SNP genotype and time point (dietary state), for each dietary group. We fitted generalized linear models using the `gls()` function in R while, allowing for intraindividual correlation of observations using the option `corSymm(form = ~1 | sample)` (Supplementary Fig. 3)...”

Results (lines 143-145, changes in existing text highlighted in bold): “Although close to 99% of cis-pQTLs were shared between all contexts, we detected nine variants displaying strong differential effects. **These differential cis-pQTLs were also**

identified in a sensitivity analysis using SNP x diet (time point) interaction testing (Supplementary Fig. 3).”

3. Since this is a longitudinal study, the analysis should prioritize within-individual changes across time (T1 vs. T2) rather than group-based comparisons. Also, the baseline heterogeneity should be considered as confounding factors when perform this kinds of analyses.

Indeed, our study focuses on within-individual changes, occurring from T1 to T2, rather than on group-based comparisons. As described in the manuscript (Results, lines 141-142) differential cis-pQTLs reported derive from testing whether the effect of the same genetic variant changes from T1 to T2 within each dietary group. Thus, reported effects reflect within-individual variation in protein abundance over time.

We agree with the Reviewer though that this may not have been sufficiently clear in our manuscript and we have rephrased the text (highlighted in bold) to emphasize that comparisons reflect within-individual changes (lines 141-142): “Next, we **focused on within-individual comparisons** to identify cis-pQTLs exhibiting differential effects between time points for both dietary groups.”

With respect to baseline heterogeneity, while no significant differences between dietary groups exist in protein abundance at T1 (please see response to comment 1), we agree with the Reviewer that smaller trends could have some impact on our results. However, our differential cis-eQTL mapping relies on within-individual changes from T1 to T2, as explained above, and any remaining baseline differences are inherently controlled for by the study’s design.

4. The comparisons of BMI and other clinical markers (e.g., GGT, eosinophil counts, cholesterol) between T1 and T2 should be presented for both PR and TR groups. Many plasma protein changes are known to be associated with BMI and other physiological traits. If BMI also increased in the PR group at T2, this could cause the changes of plasma proteins that unrelated to diet. At present, it is unclear which confounders were included in the statistical models.

We thank the Reviewer for raising this important point and we agree that BMI is a known confounder that could affect plasma protein levels. The covariates used in our statistical models were included in the Methods (lines 475-481), but we understand that this needs to be more visible and we have now included a phrase in the Results.

Existing text from Methods (lines 475-481): “We used TensorQTL⁶⁰ to associate genotypes from 4,349,373 cis SNPs (within a 2 Mb window centred on the TSS of each protein-coding gene), with levels of 1,181 proteins (excluding proteins on chromosome X) for each context, adjusting for: (a) age, sex, BMI and medication use, (b) context-defined protein principal components (PCs) from protein values, defined by the elbow method (9, 11, 10, 11 PCs for PR at T1, PR at T2, NR at T1, NR at T2 respectively), and (c) the first three genotype PCs based on an LD-independent subset of SNPs using plink v2⁶¹ to correct for population structure.”

New text in Results (lines 113-115): “We performed cis-pQTL mapping in each dietary group and time point combination separately, using TensorQTL and adjusted for both

fixed covariates including age, sex, BMI and medication use, as well as for latent factors from both genotype and protein abundance data (Fig.1).”

Regarding the impact of animal product restriction on BMI and on clinical chemistry markers (e.g. GGT, eosinophil counts, cholesterol), we agree with the Reviewer that it is important to study these effects. Indeed, this has been covered in a previous study by our group (Loizidou et al¹⁰, Tables 1-3). We have now included a paragraph in the Results to address this comment (lines 98-102): “Previous work on this cohort has shown that short-term restriction of animal products led to changes in clinical biomarkers largely in a direction associated with positive effects on health, but with no impact on BMI. These included substantial reductions in levels of total and non-HDL cholesterol, improvements in biomarkers of liver and renal function, decreased markers of inflammation, and enhanced anti-inflammatory cytokine responses.¹⁰”

Finally, we would like to clarify that in the present manuscript, the data for clinical markers such as GGT and eosinophil counts were used as input for colocalization with differential cis-pQTL, and were retrieved from publicly available GWAS, i.e. from independent studies to the present one, which is standard practice in molecular QTL studies (please see Methods lines 558-566). We have now made this clearer in the manuscript, please refer to response to comment 5b below.

5.(a) Significant seasonal variation was observed in the results. Season should therefore be explicitly modeled as a covariate to avoid confounding. Without this adjustment, it is not possible to separate diet-related effects from seasonal fluctuations in plasma protein levels.

We appreciate this comment by the Reviewer and agree that seasonality can confound results. Indeed, our work has uncovered seasonal effects, an aspect of human biology that remains understudied, although studies demonstrating such effects in humans are emerging^{11,12}. To avoid confounding effects of seasonality in our study, both dietary groups were profiled during the exact same windows (time points). Season is therefore represented by sampling time points: T1 in autumn, when both groups were on an omnivorous diet; T2 in spring, when the PR group had practiced three-to-four weeks of animal product restriction, and the NR group was on an omnivorous diet. For each time point, sample collection was completed within a two-week period to minimize effects of seasonality, and all appointments were scheduled between 7:30-9:30 am to minimize circadian effects. Therefore, we expect seasonal effects on proteins to be homogenous between dietary groups and we do not expect that our results are confounded by seasonality. Rather, we suggest that direct effects of seasonality are captured to the best of the study’s ability, by the control (NR) group.

To clarify this key point in the text we have added the following phrase in the Results (lines 96-97): “Sample collection at each time point was completed within a two-week period to minimize seasonal effects within the sampling window.”

5.(b) The health outcome measures are not clearly defined. The manuscript often refers to potential links with obesity, liver function, or immune-related diseases, but no direct clinical endpoints were reported in this cohort. The authors should clearly state the observational variables in this study

We thank the Reviewer for this comment which helps make our manuscript clearer. The health outcomes mentioned in our study (Results lines 160-169) are from publicly available GWAS data and were used in the statistical colocalization analysis, as is standard practice in colocalization analysis between QTLs and higher level phenotypes. In this analysis, we assessed whether differential cis-pQTLs identified in our study colocalize with SNPs that have been associated with health outcomes and clinical traits from published GWAS (SNP-phenotype associations retrieved through the OpenGWAS catalog¹³ <https://opengwas.io/>). To address the Reviewer's comment and to make this clearer in our manuscript, we have now provided more information on the standard colocalization approach applied in the Methods, (lines 558-566): "Briefly, each differential cis-pQTL along with its proxies ($r^2 > 0.8$) was queried against OpenGWAS, using the in-built "phewas" function of the "ieugwasr" R package, to identify known associations between the differential cis-pQTLs and clinical outcomes with $p\text{-value} < 1e-5$, defining loci to be taken forward to colocalization analysis. For each such locus, we retrieved publicly available summary statistics for SNPs mapping within a 2-Mb window centred on the transcription start site of the pGene (i.e. the gene that corresponds to the differential cis-pQTL). We then performed single trait colocalization in each context separately, using the Coloc R package¹⁴ with default parameters."

6. It is unclear whether dietary intake was monitored during the study period. Were dietary records, recalls, or surveys collected to confirm adherence to the fasting practice? A more detailed description of the dietary protocol is needed for interpreting the observed changes.

We thank the Reviewer for this important comment. Although we did not monitor dietary intake, the dietary pattern adhered to by the study's participants is a very well-established pattern of periodic restriction of animal products that is practiced for religious reasons^{15,16}. This pattern has also been described extensively in previous work from our group, including Rouskas et al¹ (ref 9 in main text) and Loizidou et al¹⁰ (Supplementary Note 2 from ref 10 in main text). We agree however, that it is necessary to include more information in the present manuscript to help with the interpretation of observed changes. To address the Reviewer's comment, we have now added detailed information regarding the dietary pattern as Supplementary Notes 1 and 2.

Supplementary Note 1.: "FastBio study inclusion and exclusion criteria (from Rouskas et al¹)

Inclusion criteria:

- Healthy female and male subjects 18 – 75 years of age at the time of enrolment. One participant turned 76 between time of enrolment and first sampling timepoint (T1).
- Able to provide signed and dated informed consent. Willing to provide blood samples.
- Individuals who had practiced periodic animal product restriction for at least ten years (for the PR group).

- Individuals who had not practiced any kind of specific diet including veganism, vegetarianism, caloric restriction, intermittent fasting (for the NR group).

Exclusion criteria:

- Use of antibiotic, antifungal, antiviral or antiparasitic drugs six months prior to T1.
- Acute disease, defined as the presence of a moderate or severe illness with or without fever, at sampling timepoints.
- Alcohol or drug abuse two years prior to T1.
- Participants who are normally periodically abstaining from animal products, but had to alter their diet for a specific reason (e.g. pregnancy)."

And *Supplementary Note 2*: "Dietary pattern of periodic animal product restriction specified by the Greek Orthodox Church (from Rouskas et al¹)

Individuals following the dietary regimen of the Greek Orthodox Church practice restriction of animal products through abstinence from meat, fish, dairy products and eggs for 180-200 days annually. Consumption of shellfish and molluscs is permitted on all days of restriction. Restriction is practiced on Wednesdays and Fridays throughout the year (excluding the week immediately after Christmas, Easter and the Pentecost) and over four extended periods annually as follows:

- 40 days before Christmas (abstinence from meat, dairy products and eggs, consumption of fish is permitted except on Wednesdays and Fridays).
- 48 days before Easter (Lent) (abstinence from meat, fish, dairy products and eggs, consumption of fish is permitted only on March 25th and on Palm Sunday).
- 0-30 days in June (abstinence from meat, dairy products and eggs, consumption of fish is permitted).
- 15 days before August 15th (Assumption) (abstinence from meat, fish, dairy products and eggs, consumption of fish permitted only on August 6th). ..."

We have also added the following text to the Methods:

(lines 412-415): "Briefly, we interviewed over 1,000 individuals from the greater region of Thessaloniki, Greece, conducting two rounds of interviews to identify participants fulfilling the study's stringent criteria (Supplementary Notes 1-2). Following two interviews, 411 apparently healthy individuals were included in the study."

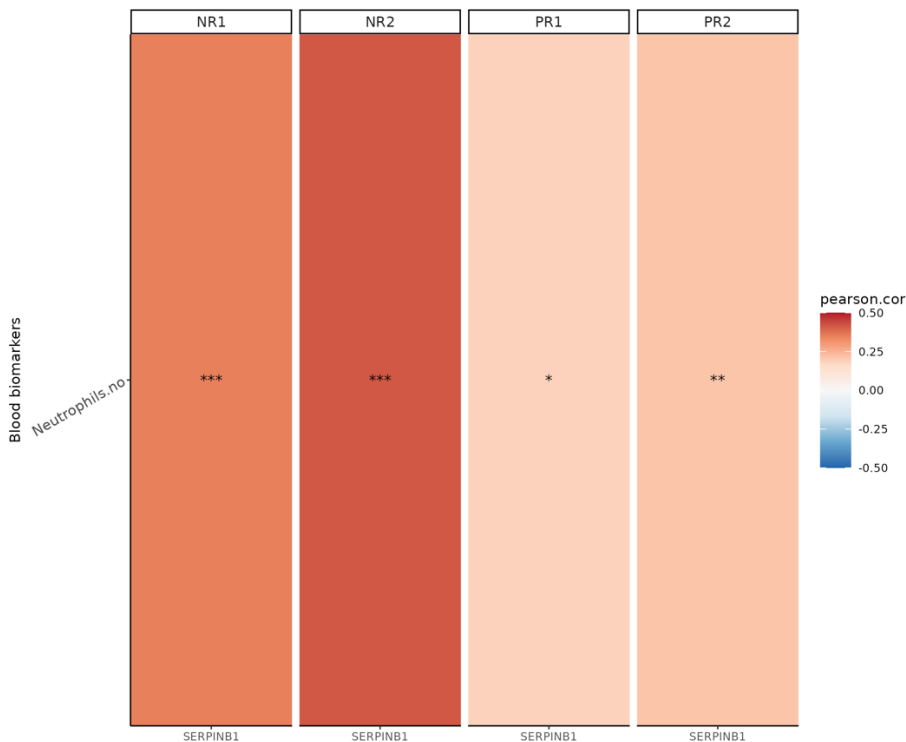
(lines 418-424): "This dietary regime is a real-world dietary pattern that is strongly motivated by religious belief, and is typically part of the culture of the family practicing individuals are born into. Adherence usually starts during childhood and involves predictable and highly consistent patterns of abstinence from meat, fish, dairy products and eggs for 180-200 days annually in a highly structured pattern involving four extended periods of abstinence during the year and abstinence on Wednesdays and Fridays (e.g. on these days practicing individuals typically consume legumes)."

7. The discussion around SERPINB1 and neutrophil counts is overstated. While SERPINB1 is known to be involved in the regulation of neutrophil survival, the manuscript does not demonstrate a direct association between SERPINB1 abundance and neutrophil counts in this study as well as the influence on the immune system.

We thank the Reviewer for raising this point and we agree that we have not demonstrated a direct association between SERPINB1 and neutrophil counts or on an influence on the immune system. Rather, we provide evidence of a genetic association between rs2293772 and plasma SERPINB1 abundance. We then show that this genetic association also extends to neutrophil counts. We agree with the Reviewer however that the manuscript would benefit from further clarification and from toning down this finding.

To support our rationale for testing whether the genetic association of rs2293772 with SERPINB1 levels also extends to neutrophil counts, we have now highlighted more that SERPINB1 is expressed in high levels in neutrophils¹⁷ (ref 26 in main text), and have also shown the existence of a positive correlation between SERPINB1 levels and neutrophil counts measured in FastBio participants (as described in Loizidou et al¹⁰).

We have now included this finding as Supplementary Fig. 6 (shown below):



Supplementary Figure 6. Correlation of SERPINB1 levels with neutrophil counts. Pearson's linear correlation between inverse rank normal transformed SERPINB1 protein levels and Neutrophil counts in each context. indicate significance as *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$. The colour gradient reflects the magnitude and direction of the correlation, with red indicating positive correlation and blue negative correlation; PR: periodically restricted group; NR: non-restricted group; T1: time point 1; T2: time point 2; pearsons.cor = Pearson's correlation.

Although this finding is only a correlation, combined with experimental evidence from the literature that shows that SERPINB1 is highly expressed in neutrophils, it supports our rationale for testing whether the genetic association detected for SERPINB1 also extends to neutrophil counts. To address this point, we have updated the Results as follows:

Included new text to improve clarity (lines 249-251): “In line with the role of SERPINB1 in neutrophils, plasma SERPINB1 levels showed a significant positive correlation with neutrophil counts across all contexts (Supplementary Fig. 6).”

The section’s subtitle has been revised from (lines 242-244): “Seasonally regulated genetic effects on SERPINB1 associated with regulation of neutrophils”) to “Shared genetic effects on key neutrophil regulator SERPINB1 show seasonal specificity” to provide a more direct and conservative description.

The finding has been toned down as (lines 256-259): “Our findings suggest that genetic regulatory effects may be modulated by seasonal factors, such as changes in pathogen exposure. Longitudinal studies spanning the whole year will be required to better characterize season-dependent genetic influences on protein levels.”

8. Given the small sample size of this study, any power analysis has been done before performing the MR and colocalization analyses? For example, I notice that only a few cis-pQTLs can explain substantial variance, most are weak instruments.

We thank the Reviewer for this comment on statistical power and acknowledge the importance of power considerations. Indeed, we did not perform MR analyses because of limited power due to the small number of strong IVs (F-statistic > 10). We have now edited the phrasing in the Results to make this point clearer (lines 169-173): “Given that **only a single strong** instrumental variable (IV) **was available for each differential pGene, our statistical power** to perform two-sample Mendelian randomization (MR) **was limited and we did not proceed with these analyses.**”

Regarding genetic colocalization, to the best of our knowledge, these analyses currently lack standardized frameworks for formal power calculation, and are typically evaluated based on effect sizes, signal strength, and posterior probabilities. In our study, we applied standard methodology to perform statistical colocalization analysis, similarly to Koprulu et al¹⁸. To that end we used a stringent process (please refer to our response to comment 5.(b)), for selecting and prioritizing loci as candidates for colocalization.

9. In general, there are a lot of speculative statements in the results section. For example, “genotype-dependent seasonal variation may contribute to individual differences in seasonal susceptibility to these diseases.”, “these signalling cascades may aid the identification of downstream effectors of FGF21 that could constitute potential therapeutics.” They should be moved to discussion. Also, the authors should be more cautious about the clinical interpretation and overstatements.

We thank the Reviewer for this suggestion, we now have made changes throughout the manuscript. Specifically, we have now edited the subtitle in line 243 (previous subtitle: “Seasonally regulated genetic effects on SERPINB1 associated with regulation of

neutrophils”) to “Shared genetic effects on key neutrophil regulator SERPINB1 show seasonal specificity” to provide a more direct and conservative description.” In addition, we have toned down the SERPINB1 finding (lines 256-259): “Our findings suggest that genetic regulatory effects may be modulated by seasonal factors, such as changes in pathogen exposure. Longitudinal studies spanning the whole year will be required to better characterize season-dependent genetic influences on protein levels.”

Following the Reviewer’s suggestion, we have also revised the text in lines 317-319 from: “We highlight the likely metabolic importance of eosinophils in humans and suggest that elucidating these signalling cascades may aid the identification of downstream effectors of FGF21 that could constitute potential therapeutics” to the following, which has been moved to the Discussion section (lines 356-358): “Further functional studies will be required to clarify these pathways and to assess their relevance for translational strategies targeting metabolic health.”

10. The authors should also explicitly state whether genetic ancestry PCs were included in all models to avoid population stratification bias

We agree with the Reviewer that this is an important point and needs to be more obvious in the manuscript. Indeed, the first three principal components derived from genotype data were included in all genetic association analyses, to control for potential population stratification. This was stated in the Methods (lines 475-481), but we have now included a phrase in the Results to highlight this key point (please also refer to our response to comment 4): “We performed cis-pQTL mapping in each dietary group and time point combination separately, using TensorQTL and adjusted for both fixed covariates including age, sex, BMI and medication use and for latent factors from both genotype and protein abundance data (Fig.1).” Additionally, we have now moved the PCA figure to the main text as Figure 2.

Minor comments:

1. More details in missing data handling (expression below LOD) in proteome profiling should be provided

We have now rephrased this for clarity in the Methods (lines 447-448): “Protein assays with NPX values below limit of detection (LOD) in more than 40% of the samples were excluded from downstream analysis.”.

2. Any multiple testing correction methods have been included in the genetic analysis to control the false discovery?

We agree that this is of key importance and indeed all our genetic association analyses have been corrected for multiple testing. We have now included the following phrase in the Methods to state this more clearly (lines 481-484): “All cis-pQTL analyses using TensorQTL were adjusted using (a) SNP-level correction using 1,000 permutations to account for multiple cis-SNPs tested for each protein coding gene, followed by (b) gene-level correction using q-values, controlling for FDR at 5%, correcting for multiple proteins being tested.”

For differential cis-pQTL mapping using mashr, we controlled for multiple testing by setting the local false sign rate (lfsr) to 5%. Lfsr is a measure of significance that requires true discoveries to be not only nonzero, but also correctly signed¹⁹. We have now included a more detailed section on how correction for multiple testing was performed in the Methods (lines 519-521): “Lfsr provides a more stringent measure of significance of the estimate compared to FDR as it requires true discoveries to be not only non-zero, but also correctly signed⁶⁷.”

3. The report of sample sizes after QC across text, tables, and figures should be consistent

We thank the Reviewer for pointing this out. Differences in sample sizes appeared in the Main text between lines 433-434, and line 451. This is a result of data-specific QC pipelines being applied to genotype vs. proteomic data, resulting in small differences in sample size between datasets. We have now addressed this point in the Methods by explicitly describing how these numbers arise.

Methods (lines 455-456): “Three first degree relatives were omitted from all analyses and a single individual was excluded due to low call rate and excess heterozygosity.” and (lines 468-470): “Overall, following all QC steps, 386 individuals (191 PR and 195 NR) who had measurements for both time points were considered for downstream proteogenomic analyses.”

Additionally, for sample sizes in Fig. 4, differences between panels **a** and **b** arise due to metabolite QC as described in Rouskas et al, where 14 participants were removed due to outlying observations, resulting in a total of 370 individuals with measurements at both time points.

Reviewer #2 (Remarks to the Author):

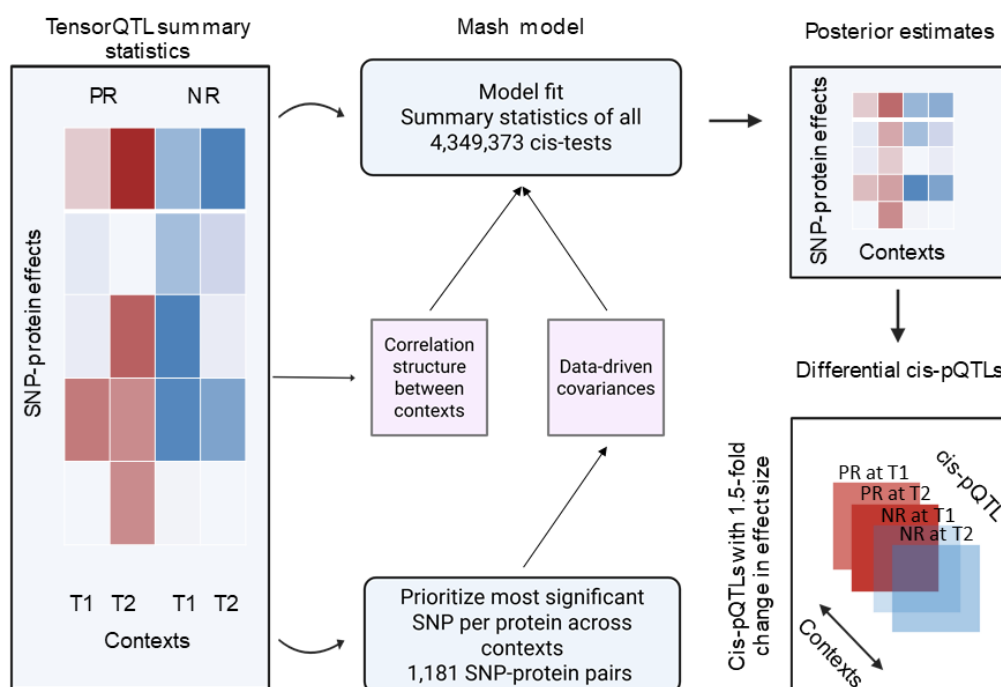
Simistiras et al. evaluate the effects of diet on genetic regulation of proteins using proteomic and genetic data on two samples from the FastBio study: individuals who alternate between omnivore and vegetarian diets, and individuals who are consistently omnivore. The study design is clever with samples from two time points for each sample to evaluate both differences in animal products in diet as well as seasonal differences. Overall, I think this is a fascinating and largely new area of study with a lot of translational potential given that diet is modifiable while the genome is not. My main concerns with the manuscript are a lack of clarity in how differential mapping was assessed as I outline further below.

Major Comments

- Please provide more information about differential cis-pQTL mapping using mashr and the other tests performed in the methods section. Even though the workflow is cited, since this is a or indeed, the, central analysis in the paper, I think all necessary information to understand the analysis should be included.**

We thank the Reviewer for this comment and agree that a better workflow description will make our manuscript easier to follow. We have now addressed this point by including a workflow figure outlining our study’s analytical strategy (Supplementary Fig. 1, shown below). In addition, we have edited the Methods and Results sections to

provide a clearer explanation of how analyses were performed, addressing the specific points raised by the Reviewer.



Supplementary Figure 2. Workflow of mash model fit adapted from Sarah M. Urbut et al (2019). Using the mashr package we jointly analysed the summary statistics from all the cis-tests performed using TensorQTL (4,349,373 cis-tests resulting from the association of ~5.1 million SNPs with 1,181 plasma proteins). Effect size estimates and standard errors from all 4,349,373 cis-tests across all four contexts (PR at T1, PR at T2, NR at T1, NR at T2) were used to learn the correlation structure of SNP-protein effects across contexts. A set of strong tests, defined as the most significant SNP per protein across contexts (1,181 SNP-protein pairs), was then used to estimate data-driven covariances. The mash model was subsequently fitted using the set of total summary statistics, and posterior effect estimates were obtained for the strong set. Lastly, we compared posterior estimates of the strong set to identify differential cis-pQTLs. Figure created using BioRender.

In addition to more detail in general, some specific questions include

- **What does a training set mean in this context? In general, I think of a training dataset as a subset of individuals, but it appears that this training set is the total number of summary statistics. For the training set, what information is provided? Summary statistics? If so, which summary statistics? More clarity here will benefit the reader.**

We agree with the Reviewer that “training set” is a misnomer since we did indeed use the total number of cis-tests performed (i.e total SNP-protein pairs 4,349,373 tests in

each context) to fit the mash model. To address this point, we have replaced the term “training set” with “**set of total summary statistics**” throughout the text.

We have now updated the Methods as follows (changes to the original text highlighted in bold, lines 510-516): “Using the default workflow (https://stephenslab.github.io/mashr/articles/eQTL_outline.html) we distinguished SNP-protein summary statistics **(effect sizes and standard errors) from all the cis-tests performed into two sets: (a) those corresponding to the total number of tests (4,349,373) for each context, and (b) those corresponding to a strong set of tests defined as the most significant (smallest p-value) SNP-protein association across the four contexts, for each of the 1,181 proteins studied.**”

- **What groups are the mixture weights for? The four groups (PR1, PR2, NR1, NR2)? Are the mixture weights estimated on average across all tests and then cis-pQTLs that differ from those mixture are identified? Something else? I don't think a lot of additional language is needed here, just half a sentence to clarify.**

We thank the Reviewer for this comment which helps improve the clarity of our analysis. The mixture weights have been estimated by maximum likelihood across all covariance combinations, using the set of total summary statistics across all contexts (PR at T1, PR at T2, NR at T1, NR at T2). Specifically, for the model fit we used both canonical and data driven covariances representing different patterns of specificity and sharing of SNP effects across contexts, while accounting for correlation of SNP estimates between contexts since we had paired observations, i.e. sample overlap between time points. After fitting the model, we extracted posterior estimates for the strong set and defined cis-pQTLs as SNPs with a $lfsr < 5\%$ in at least one context.

To make this clear in our manuscript we have added the following phrase (changes to the original text highlighted in bold) in the Methods (lines 516-519): “We fitted the mash model **to the set of total summary statistics (a) as specified above**, to learn the mixture weights **across all contexts**, and scale the coefficients, and subsequently extracted the posterior estimates for the **strong set (b)**. Significant cis-pQTLs were defined **as SNPs that had a $(lfsr) < 0.05$ in at least one context.**”

- **How is local false sign rate determined? What does it mean? Again, a one sentence explainer should suffice with the citation.**

We agree that some more information on $lfsr$ would be helpful. Local false sign rate ($lfsr$)¹⁹ is estimated for each SNP and reflects the posterior probability that the effect has the wrong sign, based on the posterior distribution obtained from the fitted empirical Bayes model across all contexts. This point has been also addressed in our response to Reviewer 1 in comment 2, Methods (lines 519-521) “ $Lfsr$ provides a more stringent measure of significance of the estimate compared to FDR as it requires true discoveries to be not only non-zero, but also correctly signed.⁶⁷”

- **What does “manually tested” mean?**

We agree with the Reviewer that this phrase, used to describe how we defined differential cis-pQTLs, is misleading as we did not perform any additional statistical

test. We have therefore replaced “manually tested” with “defined” to reflect how the differential cis-pQTLs were discovered. We have revised the Methods section accordingly (lines 521-524): “We defined differential cis-pQTLs, as cis-pQTLs that had a significant effect ($lfsr < 0.05$) in at least one time point (T1, T2) and the change in their posterior effect estimate between time points was greater than a factor of 1.5 within each dietary group, similar to Lin et al⁸. We further grouped differential cis-pQTLs as PR-specific if the differential cis-pQTL was only evident in the PR group, NR-specific if the differential cis-pQTL was only evident in the control group (NR) and shared if the differential cis-pQTL was evident in both dietary groups. Only differential cis-pQTLs that were also supported by TensorQTL were retained for downstream analyses.”

- How is the statistical test performed to assess the differential mapping?

Following mashr analysis, no further statistical testing was necessary to assess differential mapping, since the posterior estimates of effect sizes and significance ($lfsr$) for each SNP are directly comparable between contexts. To highlight this in the manuscript we have added the following phrase in Methods (lines 508-510): “Following that, mashr uses adaptive shrinkage improving estimates of effects and significance enabling direct comparison of SNP effects across conditions to define patterns of specificity and sharing, without the need of additional testing.”

- Were the comparisons adjusted for multiple testing?

Given that no further statistical testing was performed there was no need to adjust for multiple testing. Importantly, the $lfsr$ measure of significance from the mashr joint analysis (described in our response to Reviewer 1 minor comment 2) controls for the comparison of effects of multiple SNPs across contexts, in a similar way to FDR.

- I really like the pirate plots! However, I think displaying the plots for small groups is misleading (e.g., ≤ 6) – perhaps just, or in addition, plot the individual points as well?

We are glad that the Reviewer likes that pirate/dosage plots and appreciate this suggestion. We have now included the individual points in all pirate plots (Figures 4-7 and Supplementary Figures 5,7 and 9-10) and have replaced asterisks with exact p-values in scientific format with two digits. Please see updated example below for Figure 4:

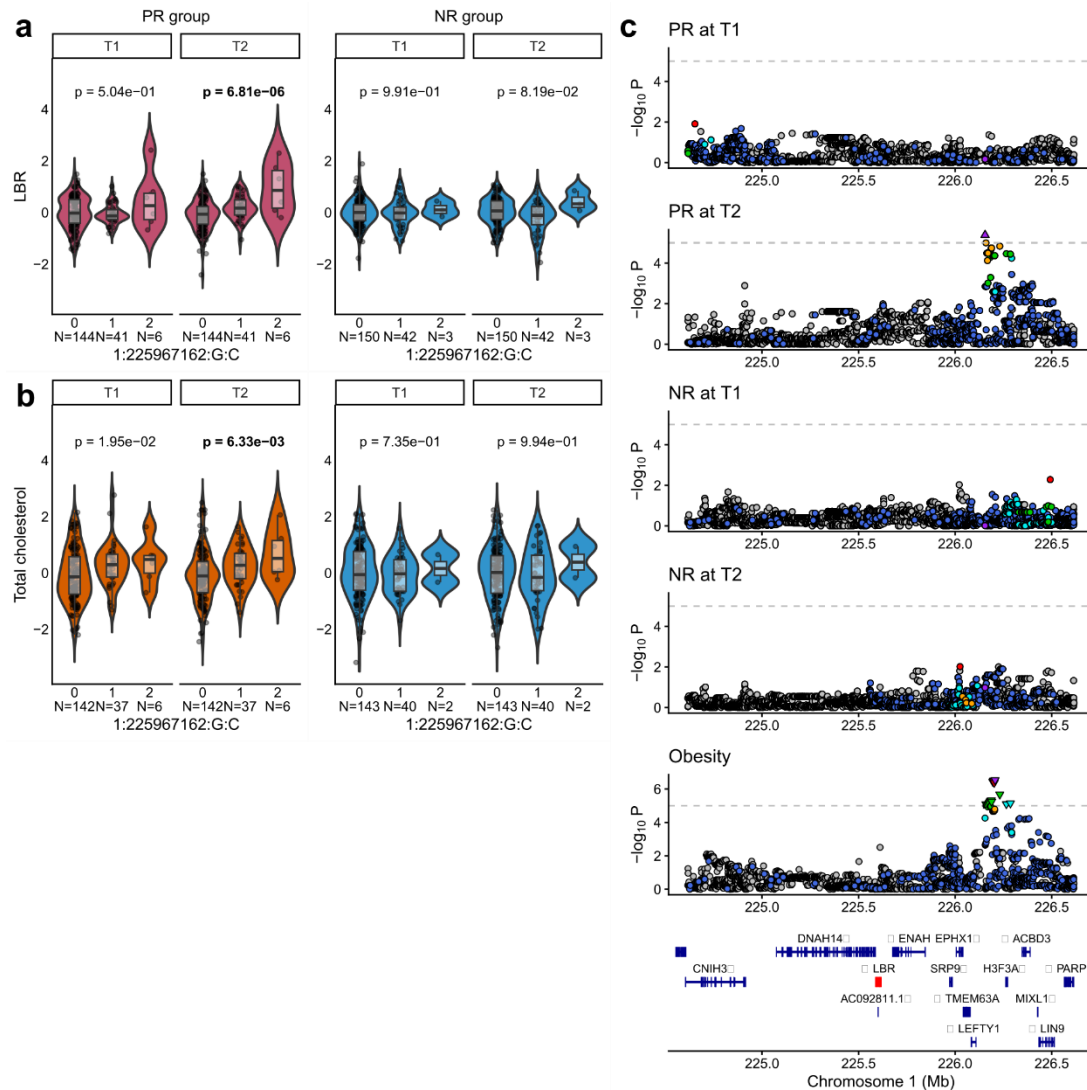


Figure 4. Diet-responsive cis-pQTL rs74148404 is associated with LBR levels and colocalizes with obesity only during animal product restriction. **a**, Dosage effect of differential cis-pQTL 1:225967162:G:C (rs74148404) on LBR levels in PR and NR groups. **b**, Validation of rs74148404 genetic effects on total cholesterol levels in PR individuals during animal product restriction (T2), no effect was detected in the NR group. **c**, Stacked regional plots for colocalization results of rs74148404 and obesity in each context; rs74148404 has suggestive evidence of colocalization with obesity only during dietary restriction (PR at T2, PPH4 = 0.55) and in no other context (PPH4 range: 0.01 to 0.06 in other contexts). “p” in panel **a** refers to the nominal p-value of the genetic association, while bold indicates significance after 1,000 permutations and controlling for 5% FDR using TensorQTL. “p” in panel **b** refers to the nominal p-value of the genetic association, while bold indicates significance after controlling for 5% FDR. PR: periodically restricted group; NR: non-restricted group; T1: time point 1; T2: time point 2.

- Please always report p-values in text and figures even if not significant (e.g., text for rs74148404) and all figures. This helps provide an idea of the strength, or lack thereof, of the association

We thank the Reviewer for pointing this out, we have now replaced all asterisks with p-values in scientific format with two digits throughout the text.

- How were adjusted p-values calculated for figures?

We thank the Reviewer for the opportunity to clarify this point. The p-values shown in the pirate plots were labelled as “adjusted” by mistake. Specifically, these correspond to nominal p-values and are provided for visual reference. TensorQTL reports permutation-based adjusted p-values only for the top (most significant) variant per protein. Since the top cis-SNP can differ across contexts, adjusted p-values are not available for the same variant across conditions. We have revised the descriptions in the legends of main Figures 4-7 and Supplementary Figures 5,7 and 9-10 to clarify the interpretation of the reported p-values. The figure legends have been updated to explicitly state that “ ‘p’ refers to the nominal p-value of the genetic association, while ‘bold’ indicates significance after respective multiple testing adjustment.”

- Line 357-358 – more information would be helpful about the imputation. What does “dietary group” refer to here? PR and NR only? It is not quite clear to me why we couldn’t impute everything together and why that might result in bias. Alternatively, if imputing separately is important why not impute separately for the different seasons too?

We thank the Reviewer for giving us the opportunity to clarify this. Indeed, we imputed NPX values that had a QC warning for each context separately, i.e. PR at T1, PR at T2, NR at T1 and NR at T2. Briefly, Olink assigns a QC warning when the value (i.e. relative protein abundance) deviates more than 0.3 units from the median value of all samples on the plate. We adopted the context specific imputation approach since we planned to perform a separate cis-pQTL mapping in each context and therefore aimed to preserve any context specific variability on protein abundance. However, we acknowledge that other strategies could be also appropriate including using the whole dataset for imputation depending on the downstream analytical strategy.

To clarify this, we have now included the following phrase in the Methods (lines 440-444): “QC was performed at both sample and protein analyte level. Samples with QC warnings were imputed using the mean NPX value of the corresponding protein across individuals. Given that our first goal was to perform cis-pQTL mapping for each context, imputation was performed separately within each context (PR at T1, PR at T2, NR at T1 and NR at T2) to preserve context- specific variation in protein levels.”

Minor Comments

How do the non-restricted and restricted PCs (genetic ancestry) compare? Can these be highlighted differently on the PC plot? Or if completed in another paper, please reference the results

We thank the Reviewer for this useful suggestion. To address this point, we have now updated the PCA plot, which we have now moved to the main text as Figure 2, to distinguish between PR and NR participants. We now show that individuals from the two dietary groups cluster together, suggesting that their genetic ancestry is very

similar. We have also added a phrase to highlight this point in the Results (lines 109-111, changes to the original text highlighted in bold): “When considering genetic ancestry, FastBio participants clustered close to individuals of Tuscan, Cypriot, Bulgarian, and Turkish descent, **with no distinct separation between PR and NR individuals** (Figure 2).”.

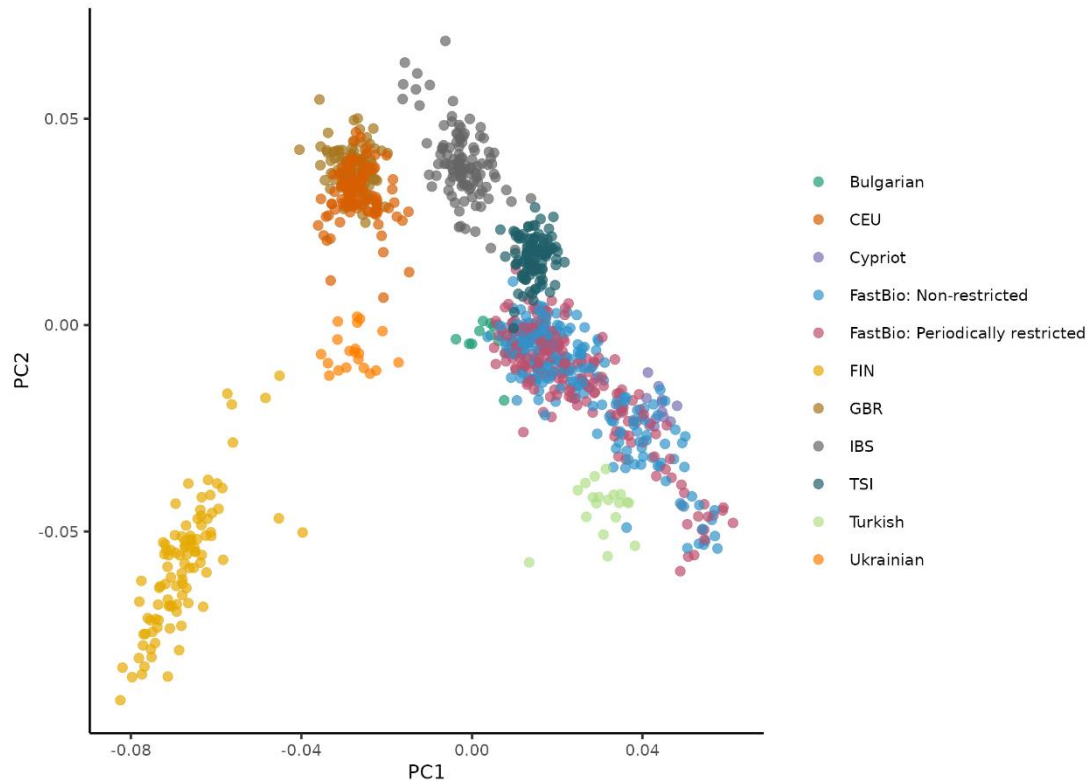


Figure 2. Principal component analysis (PCA) plot of genotypes from FastBio and reference populations. PCA plot depicting PC1, PC2 derived from genotype data from 37,922 SNPs. Colored data-points represent different populations. CEU, FIN, GBR, IBS and TSI populations are from the 1000 Genome project²⁰, while Bulgarian, Cypriot, Turkish and Ukrainian individuals are from Bánfai, Z. et al²¹. Abbreviations: CEU, Utah residents (CEPH) with Northern and Western European ancestry; FIN, Finnish in Finland; GBR, British in England and Scotland; IBS, Iberian populations in Spain; TSI, Toscani in Italia.

- (line 98) It is not clear to me what TSS is – please define. Transcription start site?

Indeed, TSS stands for “transcription start site”. We have now included this in the manuscript.

Reviewer #2 (Remarks on code availability):

The github site appears to be thorough including detailed readme file and commented code. It is easy to navigate. I did could not easily find the pirate plots, but that is a minor point. I did not attempt to run the code. While there are no explicit directions for installing and running the code, it is clear what is needed

through the code dependencies and packages including version numbers and the software needed is commonly available.

We thank the Reviewer for checking the GitHub repository. Initially, we did not include the code for the dosage plots since it had been tailored to the structure of our data and results across multiple analyses. We have now provided a dedicated script to produce these plots, under the name “dosage.effect.plot.function.R” available at (https://github.com/AlexSimis/FastBio_pqtl_project).

Reviewer #3 (Remarks to the Author):

Simistiras and et al. studies the proteogenomic effects of short-term restriction of animal products in humans. it provides some important insights on the dietary restriction in 200 healthy humans. The authors considered this as "the first study to demonstrate that short-term, real-world dietary patterns can modify genetically driven protein abundance in humans and link these diet-responsive regulatory effects to broader health-related phenotypes." The same group has previously published another work based on the same human studies in NPJ Metabolic Health and Disease (2025), explores how short-term, real-world dietary patterns—specifically, periodic restriction of animal products—can modulate metabolic pathways and influence genetically driven protein abundance. in addition, a preprint at medRxiv is also largely on the same topic. The novelty becomes questionable, and the authors should address and highlight the distinct focus of these studies if there is any and how all these data can advance the science.

We thank the Reviewer for this comment and for taking the time to go over our group's studies on animal product restriction in humans using the FastBio cohort. Indeed, our group has published an article by Rouskas et al¹. In this article, we have demonstrated that animal product restriction causes substantial shifts in the levels of metabolites and in the abundance of proteins in plasma. We revealed prominent reductions in levels of multiple lipid classes, of branched-chain amino acids and substantial shifts in eight proteins with key immunometabolic functions, including potent metabolic regulator FGF21. We would like to point out however that in this study we did not investigate genetic regulation of plasma metabolite or protein abundance. Rather, by focusing on the eight proteins affected the most by animal product restriction, we investigated potential causal effects on complex traits using Mendelian randomization (MR) analysis. To do this, we used cis-pQTLs from the UK Biobank, as instrumental variables and performed two-sample MR to assess the causal effects of these proteins on specific phenotypes. Therefore, cis-pQTLs referred to in Rouskas et al¹, are not from the FastBio cohort, but from the UK Biobank.

In our study by Loizidou et al¹⁰ (previously in medRxiv, now published in *Communications Medicine*) we have explored the impact of animal product restriction on clinical chemistry measurements, demonstrating that most changes are in a direction suggesting positive effects on health. We also revealed shifts in immune cell types through flow cytometry, and differences in cytokine responses following immune challenge *in vitro*.

Although conducted in the same cohort (FastBio), we believe that the scope of the above studies is very different to that of the present study. In the present manuscript by Simistiras et al, we have investigated the genetic regulation of plasma protein

abundance in FastBio individuals, and have explored how these genetic effects are modulated by animal product restriction. As described in (lines 66-69) “In the present study, we explored the impact of diet on proteogenomic regulation by profiling a unique group of apparently healthy individuals who alternate between omnivory and dietary restriction of animal products for religious reasons, in a consistent and highly structured pattern involving restriction for 180-200 days annually.”

To make the novelty of this study explicitly clear in the manuscript we have now included the following phrases in the Results (lines 98-109) “Previous work on this cohort has shown that short-term restriction of animal products led to changes in clinical biomarkers largely in a direction associated with positive effects on health. These included substantial reductions in levels of total and non-HDL cholesterol, improvements in biomarkers of liver and renal function, decreased markers of inflammation, and enhanced anti-inflammatory cytokine responses¹⁰. At the proteomic level, no significant differences were detected between groups when both followed an omnivorous diet (T1), however pronounced differences emerged at T2, when the PR group had practiced short-term animal product restriction⁹. Dietary restriction resulted in extensive metabolic reprogramming, with mostly beneficial effects on health⁹, including reductions in plasma levels of multiple lipid classes and, of branched-chain amino acids, as well as altered abundance of key immunometabolic proteins, including potent metabolic hormone FGF21.⁹”

Since the study focused on diet restriction, short-term fasting while sleeping should be one important factor. are the samples analyzed from humans under fed or fasting states? would feed and fasting state affect the outcome of the analysis?

We thank the Reviewer for raising this important point and agree that this should be made clear in the manuscript. Although mentioned in Rouskas et al¹ and in Loizidou et al¹⁰, we had omitted to mention in the present manuscript that all samples (from both dietary groups and at both time points) were taken following overnight fasting. To address this point, we have now included the following phrase in the Results (lines 89-92) “Participants were profiled at two time points, **following overnight fasting**: T1 in autumn, covering a period of omnivory for both dietary groups, and T2 in early spring, covering a three-to-four-week period of restriction for PR individuals, during Lent”.

References

1. Rouskas, K. *et al.* Periodic dietary restriction of animal products induces metabolic reprogramming in humans with effects on cardiometabolic health. *npj Metab Health Dis* **3**, 14 (2025).
2. Han, S. S. & Chatterjee, N. Review of Statistical Methods for Gene-Environment Interaction Analysis. *Curr Epidemiol Rep* **5**, 39–45 (2018).
3. Flutre, T., Wen, X., Pritchard, J. & Stephens, M. A Statistical Framework for Joint eQTL Analysis in Multiple Tissues. *PLoS Genet* **9**, e1003486 (2013).
4. Petretto, E. *et al.* New Insights into the Genetic Control of Gene Expression using a Bayesian Multi-tissue Approach. *PLoS Comput Biol* **6**, e1000737 (2010).
5. Uebachs, S. M., Wang, G., Carbonetto, P. & Stephens, M. Flexible statistical methods for estimating and testing effects in genomic studies with multiple conditions. *Nat Genet* **51**, 187–195 (2019).
6. Yang, C. *et al.* Genomic atlas of the proteome from brain, CSF and plasma prioritizes proteins implicated in neurological disorders. *Nat Neurosci* **24**, 1302–1312 (2021).
7. Lea, A. J., Peng, J. & Ayroles, J. F. Diverse environmental perturbations reveal the evolution and context-dependency of genetic effects on gene expression levels. *Genome Res.* genome;gr.276430.121v1 (2022) doi:10.1101/gr.276430.121.
8. Lin, W. *et al.* Genetic regulatory effects in response to a high-cholesterol, high-fat diet in baboons. *Cell Genomics* **4**, 100509 (2024).
9. Panousis, N. I. *et al.* Gene expression QTL mapping in stimulated iPSC-derived macrophages provides insights into common complex diseases. Preprint at <https://doi.org/10.1038/s41467-025-61670-9> (2023).

10. Loizidou, E. M. *et al.* Short-term animal product dietary restriction alters metabolic profiles and modulates immune function. *Commun Med* **6**, 16 (2025).
11. Dopico, X. C. *et al.* Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nat Commun* **6**, 7000 (2015).
12. Wucher, V., Sodaei, R., Amador, R., Irimia, M. & Guigó, R. Day-night and seasonal variation of human gene expression across tissues. *PLoS Biol* **21**, e3001986 (2023).
13. Elsworth, B. *et al.* The MRC IEU OpenGWAS data infrastructure. Preprint at <https://doi.org/10.1101/2020.08.10.244293> (2020).
14. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genet* **10**, e1004383 (2014).
15. Sarri, K. O., Linardakis, M. K., Bervanaki, F. N., Tzanakis, N. E. & Kafatos, A. G. Greek Orthodox fasting rituals: a hidden characteristic of the Mediterranean diet of Crete. *Br J Nutr* **92**, 277–284 (2004).
16. Georgakouli, K. *et al.* The Effects of Greek Orthodox Christian Fasting during Holy Week on Body Composition and Cardiometabolic Parameters in Overweight Adults. *Diseases* **10**, 120 (2022).
17. Baumann, M., Pham, C. T. N. & Benarafa, C. SerpinB1 is critical for neutrophil survival through cell-autonomous inhibition of cathepsin G. *Blood* **121**, 3900–3907 (2013).
18. Koprulu, M. *et al.* Proteogenomic links to human metabolic diseases. *Nat Metab* **5**, 516–528 (2023).

19. Stephens, M. False discovery rates: a new deal. *Biostat* kxw041 (2016)
doi:10.1093/biostatistics/kxw041.
20. The 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
21. Bánfai, Z. *et al.* Revealing the Genetic Impact of the Ottoman Occupation on Ethnic Groups of East-Central Europe and on the Roma Population of the Area. *Front. Genet.* **10**, 558 (2019).