

Article

Not peer-reviewed version

Joint Analysis of QTL Data Provided Insights into the Connection of Transcriptome and Proteome and the Impact of Omics Platforms

[Majid Nikpay](#)*

Posted Date: 3 April 2026

doi: 10.20944/preprints202604.0198.v1

Keywords: transcriptome; proteome; omics platform; mendelian randomization; genetic correlation; colocalization



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Joint Analysis of QTL Data Provided Insights into the Connection of Transcriptome and Proteome and the Impact of Omics Platforms

Running Title: Studying Transcriptome-Proteome Connection Using QTL Data

Majid Nikpay

Omics and Biomedical Analysis Core Facility, University of Ottawa Heart Institute, Ottawa, K1Y 4W7, Canada; mnikpay@ottawaheart.ca

Abstract

By integrating high throughput eQTL and pQTL data generated using different platforms, in this study, the relationship between transcriptome and proteome, as well as, the efficacy of platforms in measuring transcript and protein levels in blood were investigated. eQTL data were obtained from the eQTLGen study that used Microarray and INTERVAL study that relied on RNASeq platform to measure transcripts. pQTL data were obtained from UK Biobank study that used Olink and deCODE study that used SomaScan platform to measure proteins. A total of 1,162 genes that were shared between the four platforms were selected and investigated. The outcome of Mendelian randomization analysis identified 211 genes that their transcript levels significantly ($P < 5 \times 10^{-8}$) predicted their protein levels across the panels. Similarly, genetic correlation analysis identified 67 genes that share significant correlation. 12% ($N = 25$) of genes identified through Mendelian randomization and 7% of those identified through genetic correlation showed negative associations. Cross-platform analysis revealed in INTERVAL-UKBB panel the effect size of SNPs on eQTLs and pQTLs show the highest correlation; while in eQTLGen-deCODE panel this value was the lowest. Co-localization analysis further confirmed these findings and indicated genes with strong evidence of colocalization in their eQTLs and pQTLs encode intracellular proteins while those with trivial evidence of colocalization encode secretory proteins that undergo glycosylation. Integrating both transcriptome and proteome for biomarker discovery and locus annotation is important, as overall genetics of transcriptome and proteome are not the same. RNASeq and Olink platforms provide more accurate measures of RNA and protein levels.

Keywords: transcriptome; proteome; omics platform; mendelian randomization; genetic correlation; colocalization

Introduction

In recent decades, the availability of data that describe the nature of connection between SNPs and various biological features has supported research in several study fields. They enabled to investigate the mechanism whereby a genomic loci increases the risk of a disease and identify potential biomarkers through multiomics efforts that combine GWAS data with different types of QTLs ^{1,2}. They supported basic researches that seek to understand the properties of biological systems ^{3,4}.

According to the central dogma of molecular biology, the higher level of a transcript predicts higher level of its translated protein. A number of high throughput studies have systematically investigated this concept by calculating the degree of correlation between transcripts and their protein levels. As reviewed before ^{5,6} findings from these studies indicate moderate to weak correlations between the two measures. Several hypotheses have been put forward to explain this discrepancy such as the influence of post-transcriptional mechanisms (such as translation efficiency,

miRNA repression, RNA-binding proteins), protein degradation kinetics (ubiquitin–proteasome, lysosome), subcellular localization and post-translational modifications^{5–7}.

Nonetheless, gaps still exist in our knowledge including the influence of genetic factors and technological differences; furthermore, conducting such studies are important to detail the nature of relation between transcriptome and proteome and identify genes that act differently. In current study, QTLs from different sources were integrated to investigate the nature of connection between transcriptome and proteome by genetic means and study the influence of high throughput platforms in measuring RNA and protein levels.

Methods

To investigate the relation between transcriptome and proteome data, three genetic approaches were used. Mendelian randomization, genetic correlation and colocalization. Mendelian randomization is a reductionist approach; it uses a specific set of SNPs (known to influence transcript level) to find out if change in the level of a transcript impacts its protein level. Genetic correlation is a holistic approach, it uses information from all SNPs to calculate the degree of genetic similarity between a transcript and its protein. Colocalization investigates if similar cis-regulatory elements govern transcription and translation of a gene.

In the following sections, nature of data obtained for the purpose of this study and the details of analyses are provided.

eQTL Data

Previously eQTLGen consortium investigated the genetic architecture of blood transcriptome. This consortium integrated eQTLs for 19,960 genes from 37 studies, encompassing over 31,000 blood and peripheral blood mononuclear cell (PBMC) samples. The majority of these samples were of European origin. Gene expression profiling was conducted mainly using Microarray (which accounted for 79.7% of the samples)⁸. The consortium provided public access to their eQTL summary association statistics, after processing these data, eQTLs ($P < 5e-8$) for 16,732 genes were obtained for this study and used as the input data for the discovery step.

A second set of eQTLs was obtained from the INTERVAL study, in which the authors recently conducted eQTL mapping in blood samples from 4,732 participants of European origin⁹. The INTERVAL study is a randomized trial of healthy blood donors who were recruited at England's National Health Service Blood and Transplant center. Gene expression profiling of the samples was achieved by whole-blood Illumina RNA sequencing and the authors provided eQTL summary association statistics for 17,362 genes. From this database, eQTLs for 15,298 genes which also reported in the eQTLGen study were obtained for the validation purpose.

pQTL Data

Summary association statistics for SNPs regulating proteins were obtained from the UK Biobank Pharma Proteomics Project¹⁰ which is a collaboration between UK Biobank and several biopharmaceutical companies. The project aims to characterize the plasma proteomic profiles of UK Biobank participants. So far, the project has provided pQTLs for 2,923 proteins in 34,557 subjects of European descent. The authors generated the results by measuring proteins using the Olink proteomics assay and linking the data to the genotypes of the participants at UK Biobank.

pQTL data from Ferkingstad et al. study¹¹ was also obtained in which the authors used the SomaScan assay to quantify protein levels in plasma samples taken from 35,559 Icelanders of European descent that participated in the deCODE health study. The authors provided pQTLs for 4,719 proteins.

Mendelian Randomization

To investigate if change in the level of a transcript influences its protein level, Mendelian randomization was used. The method requires a set of independent SNPs known to influence the level of the transcript in order to test if change in the level of the transcript influence its protein level. To conduct the test, GSMR algorithm (version 1.1.1) was used, as compared to other methods, GSMR provides higher statistical power because it takes linkage disequilibrium information and uncertainties around SNP effect sizes into account ¹. To select a set of independent SNPs for MR analysis, the following criteria were considered:

- SNPs must be significantly associated with the transcript level ($P < 5 \times 10^{-8}$)
 - SNPs must be in linkage disequilibrium ($r^2 < 0.05$)
 - SNPs must have pQTL summary association statistics
- Genotype data from the 1000 Genomes Phase 3 ($n = 503$ of European ancestry) were obtained and used to calculate linkage disequilibrium between SNPs.

Genetic Correlation

To conduct genetic correlation analysis, SNPs that had both available eQTL and pQTL data were selected. Next, a pruning step was applied to identify SNPs that are in linkage equilibrium ($r^2 < 0.05$). Furthermore, QC analysis was performed to assure summary association statistics for each SNP is aligned to the same allele in both eQTL and pQTL dataset. Finally, Pearson correlation test was performed to test if the effect sizes of SNPs (Z-scores) on transcriptome correlate with their effect sizes on proteome. This method was selected because traditional methods such as LD Score Regression require access to complete genome-wide data of a trait to estimate the parameter of heritability; however, the eQTL data used in this study were available for SNPs proximal to the gene and a limited number of distal SNPs

Colocalization

The aim of the test is to find out if the top causal cis-SNP that influences the expression of a gene is also the top causal SNP influences its protein level. A significant effect indicates transcript and the protein level of a gene are governed by similar cis-regulatory SNPs. To conduct the colocalization test, the SMR program (version 1.3.1) ¹² was used that in addition to colocalization can differentiate between a pleiotropy and a linkage effect using a test known as HEIDI. In this context a $P_{HEIDI} > 0.01$ indicates pleiotropy or the situation that similar SNPs regulate the transcript and protein level and a $P_{HEIDI} \leq 0.01$ indicates linkage or the situation that different SNPs (that happen to be linkage) regulates the transcript and protein level of a gene. Following colocalization analysis, genes with $P_{SMR} < 5 \times 10^{-8}$ and $P_{HEIDI} > 0.01$ were considered as significant or genes whose their eQTLs and pQTLs colocalize.

Results

Table 1 provides an overview of QTL data investigated in this study. Transcriptome data in eQTLGen study were generated using Microarray, whereas in INTERVAL study, RNASeq was used to measure transcript levels. Olink and SomaScan platforms were used to measure protein levels in UK Biobank (UKBB) and deCODE study, respectively. All studies were conducted in participants of European descent. The sample sizes of UKBB and deCODE studies were similar (~35K); however, the sample size of INTERVAL study ($N=4,732$) was considerably smaller than eQTLGen study ($N=31,684$).

Table 1. Overview of the QTL data analyzed in this study.

Study Name	Source (PMID)	Data type	Platform	Sample Size	Number of Biomarkers
eQTLGen	34475573	eQTL	Microarray	31,684	19,960

INTERVAL	36991119	eQTL	RNAseq	4,732	15,298
UKBB	37794186	pQTL	Olink	34,557	2,923
deCODE	34857953	pQTL	SomaScan	35,559	4,719

To conduct the analyses and compare the findings, four panels were generated. These include eQTLGen-UKBB, eQTLGen-deCODE, INTERVAL-UKBB and INTERVAL-deCODE panel. 1,162 genes that were shared among the four datasets were identified and investigated for the purpose of this study.

Mendelian Randomization

To identify if change in the level of a transcript influences its protein level Mendelian randomization was performed. The analysis was applied to all four panels and the findings were compared. Transcript and protein of 436 genes showed concordant direction of association across four panels (**S1 Table**). Overall, higher levels of transcripts as determined by SNPs were associated with higher protein levels. The mean of regression coefficients across the panels was 0.28 [95% CI: 0.25=0.31]. The observed associations reached genome-wide significance level ($P<5e-8$) for 186 genes (**S1 Table**). Notably among them was *MDGA1* genes that its transcript level was strongly associated with its protein level (**S1 Figure**). For all these genes, higher levels of transcripts (**S1 Table**) were associated with higher levels of proteins. The outcome of pathway analysis based on DAVID functional annotation tool¹³ indicated these genes are involved in immune processes (**Table 2**). A set of 25 genes were also identified that showed significant negative association ($P<5e-8$) with their transcripts across the 4 panels. The outcome of functional analysis indicated that the identified genes are typically secreted extracellular proteins that undergo glycosylation (**S2 Table**). Notable among these genes were the enzymes *ABO* and *PAM*. *ABO* encodes a glycosyltransferase that determines the ABO blood group by catalyzing the terminal sugar modification of glycoconjugates on cell surfaces and in secreted proteins; whereas, *PAM* is an enzyme involved in processing several hormones into their active forms.

Table 2. Biological processes that are overrepresented among genes that the associations between their transcripts and proteins were positively significant ($P<5e-8$) following Mendelian randomization.

Term	Count in the network	Fold Enrichment	P-value	P-value*
Cell adhesion	25 in 677	4.07	1.1e-8	1.6e-5
Immune system process	28 in 953	3.24	1.2e-7	1.8e-4
Innate immune response	22 in 621	3.91	2.1e-7	3.1e-4
Negative regulation of activated T cell proliferation	6 in 17	38.91	3.2e-7	4.7e-4
Immune response	20 in 572	3.85	1.1e-6	1.6e-3

* Bonferroni corrected P-value.

Given that the examined QTLs were generated using different platforms, next the efficacy of platforms were investigated by calculating the degree of correlation between them based on shared SNPs. For this purpose, the list of SNPs used for MR analysis were obtained and the degree of correlation between their effects on transcripts and proteins were calculated using Pearson correlation method. The outcome of analysis is displayed in **Figure 1**. The highest degree of correlation between eQTLs and pQTLs was observed in the INTERVAL-UKBB panel ($r=0.42$ [95% CI: 0.40, 0.45]). Followed by the result in the eQTLGEN-UKBB ($r=0.38$ [95% CI: 0.36, 0.40]), INTERVAL-deCODE ($r=0.36$ [95% CI: 0.33, 0.38]) and the eQTLGEN-deCODE panel ($r=0.32$ [95% CI: 0.30, 0.33]). Altogether these results indicate, RNASeq and Olink platforms provide more accurate estimates of RNA and protein levels than Microarray and SomaScan platforms.

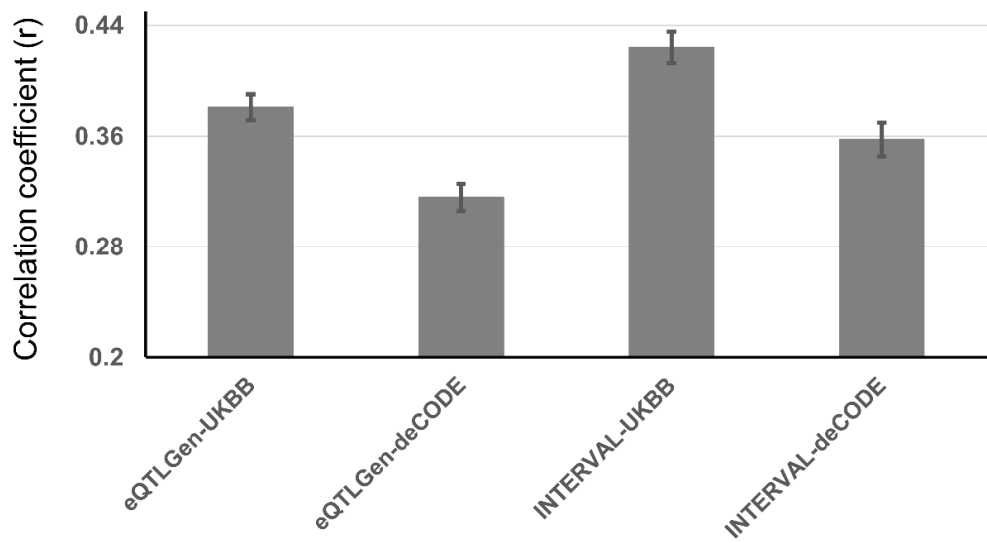


Figure 1. The degree of correlation between eQTL and pQTL effect sizes of SNPs used in Mendelian randomization.

SNPs contributing to Mendelian randomization were selected and the degree of correlation between their effect sizes (β regression coefficients) on transcript and protein level were calculated using Pearson correlation method. The highest degree of correlation was observed in the INTERVAL-UKBB panel, whereas in eQTLGen-deCODE panel, this value was the lowest. The heights of columns indicate the correlation coefficients and the error bars indicate standard errors.

Genetic Correlation

Next, genetic correlation was conducted to evaluate the degree of similarity in effect sizes of SNPs that associate with transcript and protein level of a gene. Unlike Mendelian randomization that requires specific set of SNPs. Genetic correlation takes into account the information from all available SNPs. To conduct the test, SNPs that have both eQTL and pQTL summary statistics for a gene were selected and after removing those that are in linkage disequilibrium ($r^2>0.05$). The degree of correlation between their effects (Z-scores) on transcripts and protein levels were measured using Pearson correlation method. The outcome of analysis identified 271 genes that showed concordant direction of correlation across 4 panels (**S3 Table**). Overall transcripts and proteins of genes were correlated positively. The mean of genetic correlation coefficients (r) across the panels was 0.32 [95% CI: 0.31, 0.35].

For 62 genes, the magnitude of correlation reached GWAS significance level across 4 panels ($P<5e-8$, **S3 Table**), similar to MR analysis, these genes were overrepresented in immune processes ($P=8.3e-5$). Genes *RALB*, *EPHB4*, *RGMB*, *IL18R1*, *PAM* showed significant negative correlation across 4 panels. A common feature between these genes is that they are involved in cell-surface signaling and extracellular signal transduction.

Next, SNPs contributing to genetic correlation were pooled and the degree of correlation between their influence (Z-score) on transcript and protein levels were calculated in each panel and the results were compared (**Figure 2**). The mean of correlation coefficients was highest in the INTERVAL-UKBB panel, $r=0.44$ [95% CI: 0.43, 0.46] and lowest in the eQTLGen-deCODE panel, $r=0.049$ [95% CI: 0.048-0.05]. The degree of correlation in INTERVAL-deCODE was 0.16 [95% CI: 0.155, 0.17] and in eQTLGen-UKBB was 0.20 [95% CI: 0.2-0.21]. Altogether these findings confirm the

results from Mendelian randomization regarding the higher accuracy of RNASeq and Olink platform in measuring RNA and protein levels as compared to Microarray and SomaScan platform.

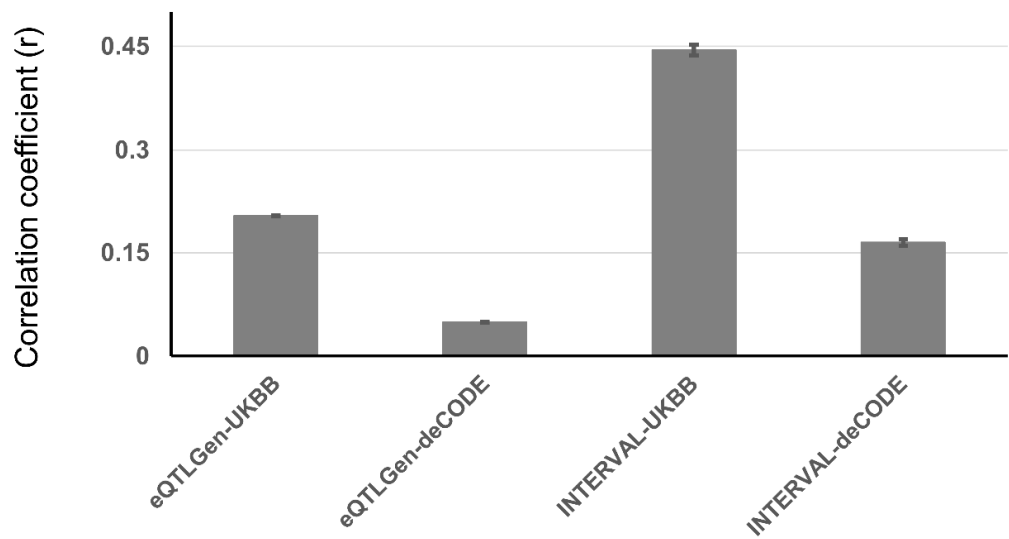


Figure 2. The degree of correlation between eQTL and pQTL effect sizes of SNPs across the platforms.

In each panel, SNPs contributing to MR analysis were selected and the degree of correlation between their influences (Z-score) on transcript and protein level were calculated. The highest degree of correlation was observed in the INTERVAL-UKBB panel, whereas the lowest correlation was observed in the eQTLGen-deCODE panel. The height of a column indicate the coefficient (r) derived from the Pearson correlation test. Error bars at the top of each column indicate standard errors.

Colocalization

To find out if the top causal variant that influences the expression of a gene is also the top causal variant influences its protein level, colocalization analysis was performed using the SMR program. The outcome of analysis revealed limited significant results (**S4 Table**). Altogether after examining the 1,162 genes shared between the QTL datasets, 90 genes were identified that showed significant evidence of colocalization ($P_{SMR} < 5e-8$ and $P_{HEIDI} > 0.01$) in at least one panel. This indicates to large extent cis-elements that regulate transcript and protein level of a gene are different. **Figure 3** provides a visual depiction of colocalization result for TBCB gene. The eQTL and pQTL regional association plots show similar distribution pattern and in each plot, SNP rs2231569 is the top associated cis-QTL. Overall, the highest number of significant signals (N=56) was observed in INTERVAL-UKBB panel (N=56), and the lowest number was observed in eQTLGen-deCODE panel (N=26). The number of significant findings in INTERVAL-deCODE panel and eQTLGen-UKBB panel were 32 and 31 respectively. Genes *TBCB*, *TCL1A*, *S100A12*, *CNP*, *STX8* showed significant evidence of colocalization in all 4 panels. A common feature among all these genes is that they are intracellular, non-secreted proteins. In contrast, genes *LRPAP1*, *IGF1R*, and *CA6* that did not show evidence of colocalization in all panels ($P_{SMR} > 0.01$, $P_{HEIDI} < 5e-8$) were all secretory/extracellular proteins that undergo glycosylation. This indicates a contributing factor to weakening the correlation between the transcript level and its protein level is the protein destination.

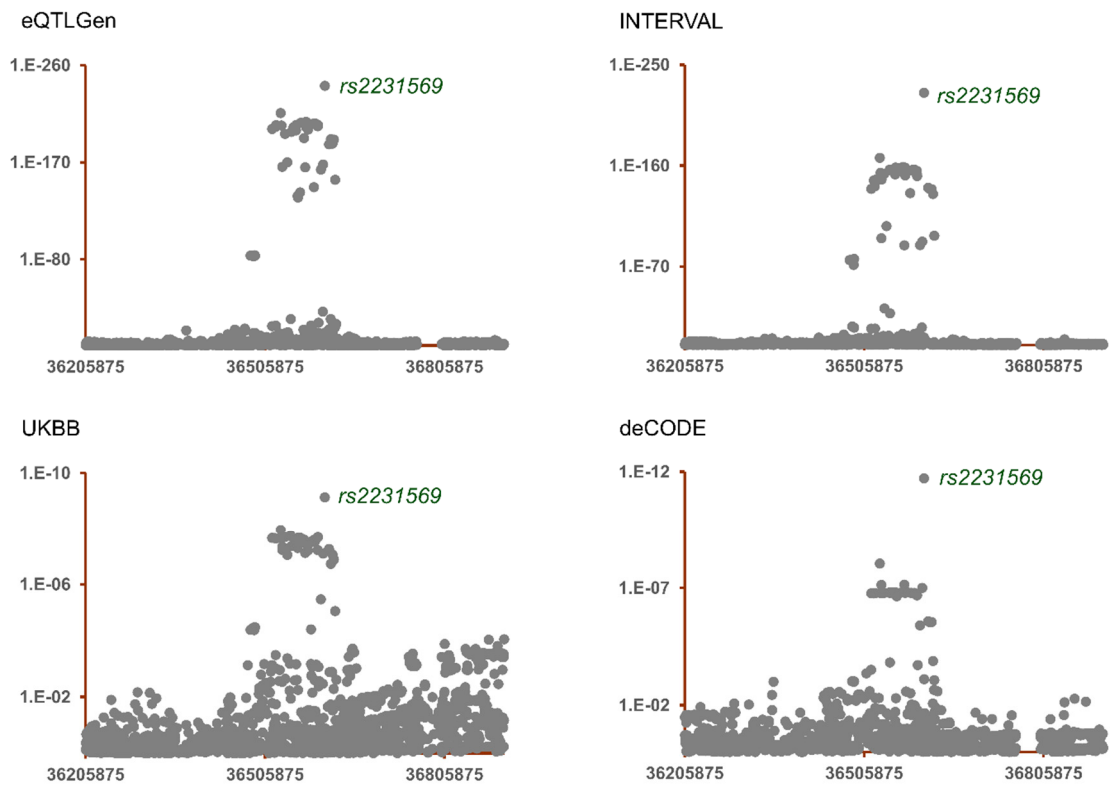


Figure 3. Regional association plots indicate colocalization of eQTLs and pQTLs.

SNPs surrounding the *TBCB* gene were selected and plotted to visualize their distributions. X-axis indicates genomic position in base pair (hg19 sequence assembly). Y-axis indicates the magnitude of association $[-\log_{10}(\text{P-value})]$. Each plot indicates the distribution of QTLs from a platform. Distribution of eQTLs and pQTLs from each platform show similar pattern with rs2231569 (upstream of *TBCB*) being the top cis-eQTL.

Discussion

Studying the relation between transcriptome and proteome is important from the basic science point of view and from the clinical perspective to identify biomarkers. Previous studies relied on direct measures of transcript/protein data to investigate the nature of connection between transcriptome and proteome data. In this study, QTL data and genetic methods were used to investigate the connections between them. This approach shields against the impact of environmental factors because distribution of genotypes of SNPs (that influence transcript or protein level of a gene) is a random process (due to the nature of meiosis and homologous recombination). The findings based on Mendelian randomization and genetic correlation analysis indicated genetic factors that influence transcript levels also influence their protein levels; however, the correlation is moderate to weak. This is in agreement with the previous estimates based on direct measures of protein and mRNA levels in the same samples ^{5,6}.

Previously, Wang et al. ¹⁴ investigated colocalization of eQTL and pQTL signals in human blood samples from 1,405 human subjects. The authors reported poor colocalization of eQTLs and pQTLs and concluded TWAS and PWAS have distinct genetic architecture. Findings from this study support their conclusion. From, the total of 1,162 gene shared between the QTL datasets, only 8% (N=90) showed significant evidence of colocalization in at least one panel. Furthermore, genes that showed strong evidence of colocalization were intracellular, non-secreted proteins. In contrast, genes that did not show evidence of colocalization were secretory/extracellular proteins that undergo glycosylation. I investigated this further by pooling the data from Mendelian randomization and genetic correlation

analysis and investigating the function of genes whose their transcript and protein levels negatively correlate (N=27) versus those that showed positive correlation (N=200). The findings indicate slight enrichment of genes (**S5 Table**) with negative correlation in candidate processes such as cell signaling (P=0.04) and glycosylation (P=0.07). Another factor that could contribute to decoupling of transcript and protein level in current study is the nature of blood QTL data. Blood transcriptome are typically generated using the immune cells, while, proteins are measured in plasma. This could create a decoupling state for genes that their plasma proteins originate from non-immune cells. Therefore, it is important to investigate findings from this study in other tissues, in order to better elucidate the mechanisms contributing to decoupling of transcription and translation.

Various molecular assays are being used to measure RNA and protein levels. At high throughput scale Microarray and RNAseq are the two primary platforms to measure RNA levels, while Olink and SomaScan are two main platforms to quantify protein levels. Previous studies attributed part of the discrepancy in correlation between transcriptome and proteome to differences in platforms that are used to measure RNA and protein levels. Findings from this study confirm this assumption. Mendelian randomization, genetic correlation and colocalization results concordantly indicated RNAseq and Olink platforms provide more accurate measure of RNA and protein levels than Microarray and SomaScan platforms. Although the sample size of INTERVAL study was seven times smaller than eQTLGen study, the eQTL data in this study showed better correlation with pQTLs in both UKBB and deCODE this further supports the consensus that overall RNAseq provides better quantification of transcripts than Microarray. Discussion on the efficacy of Olink and SomaScan platforms in measuring protein levels is an ongoing research. SomaScan provides a higher coverage of plasma proteome; furthermore, it has been reported that following repeated quantifications, SomaScan shows the highest precision (i.e., lower coefficients of variation) among common protein platforms¹⁵. Olink is reported to provide higher specificity¹⁵. Eldjarn et al.¹⁶ reported a larger proportion of proteins in Olink assay have cis-pQTLs (72% vs ~43%) than proteins on SomaScan assay; moreover, it has been reported¹⁷ samples measured using Olink show stronger intraclass correlation coefficients compared to samples processed with SomaScan platform.

In summary in this study by combining QTL data from several independent sources, the nature of relation between transcriptome and proteome was investigated. The genetic results confirm the findings from studies that by measuring the levels of transcripts and proteins reported moderate to weak correlation between transcriptome and proteome. eQTLs generated using RNAseq and pQTLs generated using Olink platform showed the strongest association, while those generated by Microarray and SomaScan platform showed the lowest association following Mendelian randomization and genetic correlation analysis. The outcome of colocalization analysis indicate cis-regulatory elements that regulate the transcript and protein level of a gene, poorly shared. Findings from this study call for integrating both transcriptome and proteome data in multomics effort that aim to identify biomarkers or investigate the function of a genomic loci.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: eQTL summary association statistics were obtained from the eQTLGen (<https://www.eqtlgen.org/phase1.html>) and INTERVAL study (<https://www.omicspred.org/downloads>). 1000 Genomes genotype data (phase 3) were obtained from https://www.cog-genomics.org/plink/2.0/resources#phase3_1kg.

Acknowledgments: This research work was enabled in part by computational resources and support provided by the Compute Ontario and the Digital Research Alliance of Canada.

Competing interests: The author declares no competing interests.

References

1. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
2. Pasaniuc, B. & Price, A. L. Dissecting the genetics of complex traits using summary association statistics. *Nat. Rev. Genet.* **18**, 117–127 (2017).
3. Boyle, E. A., Li, Y. I. & Pritchard, J. K. An expanded view of complex traits: from polygenic to omnigenic. *Cell* **169**, 1177–1186 (2017).
4. Price, A. L. et al. Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. *PLoS Genet.* **7**, e1001317 (2011).
5. Jiang, D., Cope, A. L., Zhang, J. & Pennell, M. On the Decoupling of Evolutionary Changes in mRNA and Protein Levels. *Mol. Biol. Evol.* **40**, (2023).
6. Liu, Y., Beyer, A. & Aebersold, R. On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535–550 (2016).
7. Ponomarenko, E. A. et al. Workability of mRNA Sequencing for Predicting Protein Abundance. *Genes* **14**, 2065 (2023).
8. Võsa, U. et al. Large-scale cis- and trans-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).
9. Xu, Y. et al. An atlas of genetic scores to predict multi-omic traits. *Nature* **616**, 123–131 (2023).
10. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
11. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat Genet* **53**, 1712–1721 (2021).
12. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* **48**, 481–487 (2016).
13. Sherman, B. T. et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res.* **50**, W216–W221 (2022).
14. Wang, Q. S. et al. Statistically and functionally fine-mapped blood eQTLs and pQTLs from 1,405 humans reveal distinct regulation patterns and disease relevance. *Nat. Genet.* **56**, 2054–2067 (2024).
15. Kirsher, D. Y. et al. The Current Landscape of Plasma Proteomics: Technical Advances, Biological Insights, and Biomarker Discovery. 2025.02.14.638375 Preprint at <https://doi.org/10.1101/2025.02.14.638375> (2025).
16. Eldjarn, G. H. et al. Large-scale plasma proteomics comparisons through genetics and disease associations. *Nature* **622**, 348–358 (2023).
17. Haslam, D. E. et al. Stability and reproducibility of proteomic profiles in epidemiological studies: comparing the Olink and SOMAscan platforms. *Proteomics* **22**, e2100170 (2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.