Review history

First round of review

Reviewer 1

Nagura et al. present a comprehensive study on the effects of genome variants on isoform expression. They utilize Nanopore transcriptome sequencing to identify isoforms and evaluate various genomic and functional features enriched in isoform expression QTLs (ieQTLs). Additionally, they perform experimental validation of some of their findings.

Overall, this is an interesting study that highlights several differences in the genomic features distinctly associated with ieQTLs versus eQTLs, underscoring the impact of chromatin on isoform processing. This can motivate future studies to explore the relationship between chromatin elements and splicing.

While I find this study interesting, I have several concerns that should be addressed to demonstrate the robustness of the results and improve the quality and clarity of the manuscript.

Methods

An interesting aspect of the study is the utilization of Nanopore for isoform sequencing; however, very few details are provided about these data. Information on read quality (% identity to the genome) and isoform expression level distribution should be included to reveal the quantification power of the technology on the evaluated isoforms. Additionally, the authors should explain how Nanopore errors might affect the identification of SNPs and variants in their dataset.

Line 177: The analysis of ieQTL with differential effects is confusing. The expression values of different isoforms within a gene are not independent, as the upregulation of one isoform may imply the downregulation of another for genes with constant total expression. This lack of independence violates the assumption of a t-test, so I am not convinced this analysis is correct. Furthermore, the potential association of ieQTLs with differential splicing should be viewed in the context of global splicing differences. How many genes had differentially expressed isoforms, and how many of them underwent alternative splicing? The association observed with the ieQTL might simply reflect the levels of genes showing alternative splicing within the population.

Results

Line 269: Similar to my previous comment, errors in Nanopore sequencing may lead to the identification of false novel isoforms. The authors report 6,832 novel isoforms. A quality assessment of these novel sequences should be provided to assess their reliability. Additionally, the authors should indicate how many of these novel isoforms were affected by ieQTLs.

Line 291: It is not clear how the multiple testing correction was performed. Was this a general correction considering all the performed tests, or a gene-wise correction considering the

number of tests for each gene? The correction should take into account the total number of tests in the ieQTL analysis and the total number of tests in the eQTL analysis.

Most of the ieQTLs seem to affect minor isoforms. It should be reported what the expression levels of these minor isoforms are and the detected effect sizes in these cases.

Minor comments

- Moderate the abstract sentence: "The identification of eQTLs provides a profound understanding of the mechanisms governing gene expression." Replace profound by help to understand

- Figures should be substantially improved. They do not have the quality or clarity for the Genome Biology journal

- Figure 1: Panel (B) should be better explained in the caption. The yellow color is missing in Panel (B). Panel (E) is not necessary. Panel (F) is unclear; what does each dot represent? The x-axis indicates gene p-value and the y-axis indicates isoform p-value. What does each point signify? Is it indicating a gene? How are the different isoform p-values considered? Possibly, a boxplot representation of these data would be more suitable.

- Figure 2: This figure is confusing. Are only significant functional features displayed? What is the p-value of the enrichment? What do the numbers on the x-axis of the bar plots represent? These long labels are difficult to read. Use a different visual display. I assume only the indicated p-values are differential between eQTL groups? This figure should be substantially improved for clarity.

- Figure 3B: This is difficult to interpret. Do the sizes of the concentric rings have any meaning? A bar plot might provide a clearer representation.

- Figure 4A and F and Figures 5A and B are too small. Please make the isoform representations clearer.

Reviewer 2

In this manuscript, Nagura et al present the collection and analysis of long-read RNA sequencing dataset from 67 lymphoblastoid cell lines (LCLs) of Japanese origin from the 1000 Genomes Project. The primary value of the manuscript is the long-read RNA-seq dataset which I assume will be shared completely openly upon publication (as 1000 Genomes participants have already consented for completely open sharing of their genetic data). I expect this to become an important reference dataset for developing and benchmarking of novel methods for molecular QTL mapping in long-read RNA-seq datasets and I thus expect it to become a highly cited paper. However, I have several concerns about the choice of the analysis methods and specific claims made in the paper.

1. The number of isoform expression QTLs (ieQTLs) and gene expression QTLs (eQTLs) reported in this paper are much larger than would be expected for a molecular QTL study of of this sample size. I believe that this is primarily due to the unusual strategy chosen for multiple testing correction. The standard practice in the field is to select a single lead variant per

molecular trait (e.g. gene or isoform), correct these first for the number of tested variants using either permutations (see FastQTL; https://doi.org/10.1093/bioinformatics/btv722), standard Bonferroni correction, or using Bonferroni correction with the estimated number of independent tests performed (e.g. eigenMT; https://pubmed.ncbi.nlm.nih.gov/26749306/). These trait-level corrected p-values are then further corrected for the number of molecular traits (isoforms or genes) tested using either Benjamini-Hochberg FDR or the qvalue method. While both permutations-based approach and eigenMT have been demonstrated to produce similar corrected p-values given sufficiently large LD panel for eigenMT, standard Bonferroni correction is slightly over-conservative. In this paper, the the authors use LD pruning with plink to define the number if independent tests performed per gene, which I suspect is underestimating the number of independent tests. With eigenMT correction, the number of independent tests is typically ~50% less than the number of variants in the cis region. How many independent tests per gene do you estimate with your unvalidated plink-based approach? Secondly, it seems that the authors do not correct at all for the number of isoforms or genes tested, further significantly inflating the number of discoveries. Thus, to be able to faithfully compare the number of QTLs detected in this manuscript with other similar studies, the authors should report the number of significant gene-level and isoform-level lead variants while properly accounting for both LD around each gene (using either permutations, eigenMT or Bonferroni correction with the actual number of tested SNPs) AND the number molecular traits tested (using e.g. Benjamini-Hochberg FDR). I would expect these numbers to be significantly smaller compared to the numbers currently reported in the paper. I would then like to see the all of the enrichment and replication analysis reported in this paper to be repeated using the appropriately corrected lead QTL variants.

2. On lines 284-286 the authors report that "A comparison of our eQTL list with GTEx eQTL and sQTL studies using lymphocyte samples demonstrated that 86.7% and 88.2% of the identified eQTLs were not detected, respectively (Additional file 1: Figure S2)."
These number are unplausible, given the sample size of the current study and likely reflect the issues with multiple testing I highlighted in point 1. I would like the GTEx analysis to be repeated using the significant lead molQTL variants identified after appropriate multiple testing correction (see point 1). For example, the authors could take their appropriately corrected lead variants from point 1 and ask how many of them are in high LD ($r^2 > 0.8$) with lead variants from the GTEx. Furthermore, GTEx has one of the smallest sample sizes for LCLs and multiple eQTL and sQTLs studies have been performed in LCLs such as the GEUVADIS study (Lappalainen et al 2013; https://www.nature.com/articles/nature12531) or the MAGE study (Taylor et al, 2024; https://www.nature.com/articles/s41586-024-07708-2). The authors could also take re-analysed eQTL and sQTL results for the GEUVADIS dataset from eQTL Catalogue (Kerimov et al, 2023) and perform similar replication.

3. In the discussion (lines 484-486), the authors claim that "Overall, we believe that the functional impact of variants in splice sites cannot be accurately predicted solely by software and short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing."). While this might be true, this claim is not supported by current analysis and results presented in the manuscript. In particular, the two example associations presented in the results (rs1333973 with IFI44L and rs11026723 with GAS2) were both detected by the Leafcutter method in the eQTL Catalogue re-analysis of GEUVADIS short-read RNA-seq data (Kerimov et al, 2023) and also statistically fine mapped to a single causal variant. See RNA-seq read coverage plots here:

IFI44L:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_1%3A78620571%3A78627906%3Aclu_31999_%2B_L1
GAS2:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_11%3A22668335%3A22674850%3Aclu_40188_%2B_L1

Furthermore, on lines 482-484, the author make this claim about the GAS2 sQTL: "Because this variant affected the expression level of only one isoform (Figure 4H), this pattern should be difficult to detect using short-read sequencing." I'm not sure what the authors mean by this, as the GAS2 sQTL can clearly be detected with the Leafcutter method in the GEUVADIS short-read RNA-seq data as I have highlighted above.
As a result, the strong claims about the discovery supposedly novel associations should either be significantly better substantiated or removed from the manuscript.

4. On lines 473-474, the authors state that "Furthermore, SpliceAI software predicted a low functional impact for the SNP(Δscores = 0.34) (Additional file 2: Table S18)".
While it is true that 0.34 is below the 0.5 threshold recommended in the SpliceAI paper, it is well above the "high recall" threshold of 0.2 and much higher than the background distribution of SpliceAI scores for random variants. In our experience, we often see that confidently finemapped sQTL causal variants have SpliceAI scores below 0.5, likely reflecting the fact that these variant also have relatively small effect sizes. Thus I believe that stating that SpliceAI score of 0.34 as "low" is at best misleading.

5. I would strongly encourage the authors to publicly release their quantified gene expression and isoform expression matrices together with the corresponding 1000 Genomes sample identifiers. If these had been public at the first review stage then we could have re-analysed their data using the eQTL Catalogue's standard molecular QTL analysis workflow (https://github.com/eQTL-Catalogue/qtlmap) to independently evaluate the concordance between molecular QTLs detected using long-read sequencing in this study and all of the available short-read eQTLs and sQTLs from the eQTL Catalogue.

Authors' response to reviewers

Reviewer #1: Nagura et al. present a comprehensive study on the effects of genome variants on isoform expression. They utilize Nanopore transcriptome sequencing to identify isoforms and evaluate various genomic and functional features enriched in isoform expression QTLs (ieQTLs). Additionally, they perform experimental validation of some of their findings.
Overall, this is an interesting study that highlights several differences in the genomic features distinctly associated with ieQTLs versus eQTLs, underscoring the impact of chromatin on isoform processing. This can motivate future studies to explore the relationship between chromatin elements and splicing.

While I find this study interesting, I have several concerns that should be addressed to demonstrate the robustness of the results and improve the quality and clarity of the manuscript.

Our reply: Thank you for this encouraging comment. We appreciate your insightful review.

Methods

An interesting aspect of the study is the utilization of Nanopore for isoform sequencing; however, very few details are provided about these data. Information on read quality (% identity to the genome) and isoform expression level distribution should be included to reveal the quantification power of the technology on the evaluated isoforms. Additionally, the authors should explain how Nanopore errors might affect the identification of SNPs and variants in their dataset.

Our reply: Thank you for this comment.
We added the mapping rate, error rate, and number of coding and non-coding isoforms to the Results section, as follows: "On average, 99.9% of reads were mapped, and the average mismatch rate was 7.1% (Additional file 2: Table S1)." (line 103). We also added the isoform expression level distribution to Figure S1.
In our study, we did not detect genetic variations in the nanopore sequencing data. We used genetic variation data released by the 1000 Genomes Project for the analysis. Accordingly, we changed the sentence "Based on the expression levels of each isoform and genotype data, we detected ieQTLs." to "Based on the expression levels of each isoform and genotype data obtained from the 1000 Genomes database, we detected ieQTLs" in the revised manuscript (line 111).

Line 177: The analysis of ieQTL with differential effects is confusing. The expression values of different isoforms within a gene are not independent, as the upregulation of one isoform may imply the downregulation of another for genes with constant total expression. This lack of independence violates the assumption of a t-test, so I am not convinced this analysis is correct. Furthermore, the potential association of ieQTLs with differential splicing should be viewed in the context of global splicing differences. How many genes had differentially expressed isoforms, and how many of them underwent alternative splicing? The association observed with the ieQTL might simply reflect the levels of genes showing alternative splicing within the population.

Our reply: Thank you for this important comment.
As the reviewer mentions, interpreting ieQTLs with different beta values is challenging. We tested the differences in beta values using the t-test. Since the t-test assumes the independence of two beta values, the interpretation of the results requires caution; significant results may arise either from differences in beta values or from the dependency between the two beta values. Because the majority of ieQTLs with differential effects were not located in splice sites (Table S6), we concluded that most ieQTLs with differential effects are unlikely to be caused by alternative splicing. Although the molecular mechanism of ieQTLs with differential effects is currently unclear, ieQTLs are located in intergenic regions, and some may have an isoform-specific enhancer effect. We validated one candidate (Figure 6D), finding indications of an isoform-specific enhancer, but further studies are needed.
Among the ieQTLs, 1,775 were identified as differential ieQTLs, and 221 exhibited opposite effects. Differential ieQTLs affected 1,300 isoforms (492 genes), and ieQTLs with opposite effects affected 84 isoforms (25 genes). Accordingly, we added the following sentence to the Results section: "Differential ieQTLs and ieQTLs with opposite effects affected 1,300 isoforms (492 genes) and 84 isoforms (25 genes), respectively" (line130).

Results

Line 269: Similar to my previous comment, errors in Nanopore sequencing may lead to the identification of false novel isoforms. The authors report 6,832 novel isoforms. A quality assessment of these novel sequences should be provided to assess their reliability. Additionally, the authors should indicate how many of these novel isoforms were affected by ieQTLs.

Our reply: Thank you for this important comment.
Isoforms with very low expression levels may not be reliable; therefore, we selected isoforms that had ≥ 0.5 reads per million (RPM) in ≥ 5% of the samples. Additionally, we evaluated the sensitivity of the novel isoform detection method described in our previous study (Kiyose et al., PLoS Genetics, 2022), which demonstrated the second-highest sensitivity reported for novel isoforms.
We added the following sentence to the Methods section: “SPLICE software classified mapped reads to each isoform, considering sequencing errors. In our previous comparison, SPLICE showed high sensitivity for detecting novel isoforms [14]” (line 407). Because we did not describe the criteria for the minimum expression level in our previous submission, we have added "We selected isoforms that had ≥ 0.5 reads per million (RPM) in ≥ 5% of the samples" to the revised Methods section (line 415). We apologize for this mistake. In our study, 12,202 isoforms were affected by ieQTLs, of which 2620 were novel isoforms. Accordingly, we added the following sentence to the Results section: “Among the isoforms affected by ieQTLs, 21.5% (2,620/12,202) were novel isoforms.” (line 122)

Line 291: It is not clear how the multiple testing correction was performed. Was this a general correction considering all the performed tests, or a gene-wise correction considering the number of tests for each gene? The correction should take into account the total number of tests in the ieQTL analysis and the total number of tests in the eQTL analysis.

Our reply: Thank you for this important comment.
In the previous manuscript, we performed gene-wise correction. We analyzed variants with a minor allele frequency (MAF) ≥ 0.03 within a cis-window defined as 1 Mb from the transcription start site (TSS) and transcription end site (TES) of each gene. We applied linkage disequilibrium (LD) pruning for variants within the cis-window while considering variants with $r^2 > 0.5$ as linked. We counted the number of unlinked variants within the cis-window of each gene, which was used as the number of tests and applied Bonferroni correction.
As both reviewers raised concerns about the correction for multiple testing, we compared four methods of correction:
First method: Bonferroni correction using the number of variants determined by LD pruning. The results from this method were reported in the previous version of the manuscript.
Second method: Bonferroni correction using the number of variants determined by LD pruning, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes.
Third method: Bonferroni correction using the number of variants, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. This method is the most conservative.
Fourth method: Correction using eigenMT software, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes.

Comparison of the Number of Identified Isoforms

We compared the number of identified isoforms and genes among the four methods. The number of isoforms affected by ieQTLs was 12,202 for the 1st and 2nd methods (LD pruning-based methods), 5,502 for the 3rd method (Bonferroni correction with all tested variants), and 43,607 for the 4th method (eigenMT). In conclusion, eigenMT-based method identified the largest number of isoforms, and all isoforms identified by the first three methods were included in the results of eigenMT.

Comparison of the Number of Identified ieQTLs

Since eigenMT software reports a single variant for each isoform, we searched for variants linked to the variant detected by eigenMT ($r^2 > 0.8$) and considered them as ieQTLs. We compared the number of ieQTLs and gene eQTLs. The first method identified 33,928 ieQTLs and 12,668 gene eQTLs. The second method also identified 33,928 ieQTLs and 12,668 gene eQTLs, yielding the same results as the first method. The third method identified 17,119 ieQTLs and 6,350 gene eQTLs. The fourth method identified 298,175 ieQTLs.

Enrichment analysis

We performed an enrichment analysis of the variants detected by the third method, and the results were similar to that of the first method. We added the results of the enrichment analysis to Tables S6-S13.

Correction of the text

The correction for multiple testing strongly affects the results; however, the description of multiple results obtained by different correction methods in the text is confusing. Therefore, we provided lists of variants from ZENDO (https://zenodo.org/records/13856612) and described the results in the Supplementary Information. We also explained the four correction methods in Figure S2.

Accordingly, we added the following sentences to the Methods section (line 451): "We applied four methods to correct multiple tests (Additional file 1: Supplementary information, Figure S2). The first method used Bonferroni correction on the number of variants determined by LD pruning for each cis-window. The second method applied Bonferroni correction on the number of variants determined by LD pruning for each cis-window and then performed the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. The third method applied Bonferroni correction on the total number of variants for each cis-window and then performed the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. The fourth method used eigenMT software, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes [40]."

The results of the second method were the same as those of the first method. The third method is more conservative and identified a smaller number of eQTLs. Therefore, we added the following sentences to the Results section (line 113) and mainly discussed the results of the first method in the revised manuscript as follows: "In this study, we applied four methods for correcting multiple tests (Additional file 1: Supplementary information, Figure S2). Two LD pruning-based methods (correction methods 1 and 2) identified the same eQTLs. Bonferroni correction with the total number of variants (correction method 3) identified a smaller number of eQTLs. The eigenMT-based method (correction method 4) identified a larger number of isoforms. Among them, the third method is conservative and thus may miss true eQTLs, while the fourth method would be progressive. Therefore, we mainly analyzed the results of the first method in the subsequent analysis (Additional file 2: Table S3)."

Most of the ieQTLs seem to affect minor isoforms. It should be reported what the expression levels of these minor isoforms are and the detected effect sizes in these cases.

Our reply: Thank you for your comment.
As the reviewer mentioned, the percentage of minor isoforms was quite different between common and ieQTL-specific variants. Specifically, 28.3% (3,779/13,374) of common eQTLs affected minor isoforms, while 83.7% (21,074/25,173) of ieQTL-specific variants affected minor isoforms. Additionally, we found that the mean expression level of isoforms affected by ieQTL-specific variants was higher than affected by common variants (ieQTL-specific: 12 TPM; common: 24 TPM; p-value by Wilcoxon signed-rank test = 1.9×10^{-5}).
Accordingly, we added the following sentence to the Results section (line 127): "Specifically, 28.3% (3,779/13,374) of common eQTLs affected minor isoforms, while 83.7% (21,074/25,173) of ieQTL-specific variants affected minor isoforms."

Minor comments

- Moderate the abstract sentence: "The identification of eQTLs provides a profound understanding of the mechanisms governing gene expression." Replace profound by help to understand

Our reply: Thank you for this comment. We corrected "provides a profound understanding" to "help to understand" (line 32) in the revised manuscript.

- Figures should be substantially improved. They do not have the quality or clarity for the Genome Biology journal

Our reply: We apologize for the quality of the figures. In response to the reviewer's comments, we modified Figures 1, 2, and 3. We also separated Figure 4 into two figures (new Fig. 4 and 5).

- Figure 1: Panel (B) should be better explained in the caption. The yellow color is missing in Panel (B). Panel (E) is not necessary. Panel (F) is unclear; what does each dot represent? The x-axis indicates gene p-value and the y-axis indicates isoform p-value. What does each point signify? Is it indicating a gene? How are the different isoform p-values considered? Possibly, a boxplot representation of these data would be more suitable.

Our reply: We apologize for the problems with Figure 1. Below is a summary of our revisions.
• Panel (B): We added an explanation.
• The comment regarding the yellow color in Panel (B) should actually apply to Panel (D). We corrected the figure to show a Venn diagram without color (Fig. 1D in the revised manuscript).
• Panel (E): We removed Panel (E).
• Panel (F): This panel represents ieQTL p-values and gene eQTL p-values for variants detected in both analyses. Because we identified a larger number of ieQTLs than gene eQTLs, we suspected that the larger number of isoforms may cause the higher number of ieQTLs. To examine this possibility, we performed a correction for multiple tests while considering the number of isoforms in each gene. This result is described in the Results section (line 319-328). Additionally, we compared the p-values of ieQTLs with those of gene eQTLs. If the larger

number of ieQTLs is a by-product of gene eQTLs, the p-values of ieQTLs should be similar to those of gene eQTLs. Figure 1E (previously 1F) shows this result. As shown in this figure, the p-values of ieQTL-specific variants are much lower than those of gene eQTLs. This result indicates that isoform-specific eQTLs truly exist.
Because our previous description was insufficient, we added the following sentences to the Results section: "We also compared the p-values of ieQTLs and gene eQTLs (Figure 1E). The distribution of p-values revealed that numerous variants displayed lower ieQTL p-values than gene eQTL p-values, suggesting that ieQTLs are distinct and not statistical artifacts arising from gene eQTL identification" (line 144).

- Figure 2: This figure is confusing. Are only significant functional features displayed? What is the p-value of the enrichment? What do the numbers on the x-axis of the bar plots represent? These long labels are difficult to read. Use a different visual display. I assume only the indicated p-values are differential between eQTL groups? This figure should be substantially improved for clarity.

Our reply: We apologize for the problems with the figure.
This figure reports all categories related to the features of genomic locations (exon, promoter, splice site, etc.) and epigenetics (histone marks, chromatin state, etc.) with the actual values shown in Tables S6-S13. In the revised manuscript, we changed the format of the figure and simplified the labels. The p-values were obtained using Fisher's exact test on the number of variants. For example, 43 gene eQTL-specific variants (out of a total of 2,418 variants) were in the H3K36me3 region, as were 4,283 ieQTL-specific variants (out of a total of 23,408 variants). We generated a 2x2 table (43 and 2,418 - 43 vs. 4,283 and 23,408 - 4,283) and conducted Fisher's exact test. The p-values are presented in Tables S6-S13. However, the p-values may be confusing, and the difference in odds ratios was obvious; therefore, we removed the p-values from the figure.

- Figure 3B: This is difficult to interpret. Do the sizes of the concentric rings have any meaning? A bar plot might provide a clearer representation.

Our reply: We apologize for the confusing figure.
We changed the figure to a bar plot.

- Figure 4A and F and Figures 5A and B are too small. Please make the isoform representations clearer.

Our reply: We apologize for the problems in the Figure.
We separated Figure 4 into two figures. Figure 4F-I is Figure 5 in the revised manuscript.

Reviewer #2: In this manuscript, Nagura et al present the collection and analysis of long-read RNA sequencing dataset from 67 lymphoblastoid cell lines (LCLs) of Japanese origin from the 1000 Genomes Project. The primary value of the manuscript is the long-read RNA-seq dataset which I assume will be shared completely openly upon publication (as 1000 Genomes participants have already consented for completely open sharing of their genetic data). I expect this to become an important reference dataset for developing and benchmarking of novel methods for molecular QTL mapping in long-read RNA-seq datasets and I thus expect it to become a highly cited paper. However, I have several concerns about the choice of the analysis

methods and specific claims made in the paper.

Our reply: We appreciate the insightful review and comments.

1. The number of isoform expression QTLs (ieQTLs) and gene expression QTLs (eQTLs) reported in this paper are much larger than would be expected for a molecular QTL study of of this sample size. I believe that this is primarily due to the unusual strategy chosen for multiple testing correction. The standard practice in the field is to select a single lead variant per molecular trait (e.g. gene or isoform), correct these first for the number of tested variants using either permutations (see FastQTL; https://doi.org/10.1093/bioinformatics/btv722), standard Bonferroni correction, or using Bonferroni correction with the estimated number of independent tests performed (e.g. eigenMT; https://pubmed.ncbi.nlm.nih.gov/26749306/). These trait-level corrected p-values are then further corrected for the number of molecular traits (isoforms or genes) tested using either Benjamini-Hochberg FDR or the qvalue method. While both permutations-based approach and eigenMT have been demonstrated to produce similar corrected p-values given sufficiently large LD panel for eigenMT, standard Bonferroni correction is slightly over-conservative. In this paper, the the authors use LD pruning with plink to define the number if independent tests performed per gene, which I suspect is underestimating the number of independent tests. With eigenMT correction, the number of independent tests is typically ~50% less than the number of variants in the cis region. How many independent tests per gene do you estimate with your unvalidated plink-based approach? Secondly, it seems that the authors do not correct at all for the number of isoforms or genes tested, further significantly inflating the number of discoveries. Thus, to be able to faithfully compare the number of QTLs detected in this manuscript with other similar studies, the authors should report the number of significant gene-level and isoform-level lead variants while properly accounting for both LD around each gene (using either permutations, eigenMT or Bonferroni correction with the actual number of tested SNPs) AND the number molecular traits tested (using e.g. Benjamini-Hochberg FDR). I would expect these numbers to be significantly smaller compared to the numbers currently reported in the paper. I would then like to see the all of the enrichment and replication analysis reported in this paper to be repeated using the appropriately corrected lead QTL variants.

Our reply: Thank you for this very important comment.
In the previous manuscript, we performed gene-wise correction. We analyzed variants with a minor allele frequency (MAF) ≥ 0.03 within a cis-window of 1 Mb from the transcription start site (TSS) and transcription end site (TES) of each gene. We applied linkage disequilibrium (LD) pruning for variants within the cis-window while considering variants with $r^2 > 0.5$ as linked. We counted the number of unlinked variants within the cis-window of each gene, which was used as the number of tests and applied Bonferroni correction.
Because both reviewers raised concerns about the correction for multiple testing, we compared four methods of correction:
First method: Bonferroni correction using the number of variants determined by LD pruning. The results from this method were reported in the previous version of the manuscript.
Second method: Bonferroni correction using the number of variants determined by LD pruning, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes.
Third method: Bonferroni correction using the number of variants, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. This method is the most conservative.

Fourth method: Correction using eigenMT software, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes.

Comparison of the Number of Identified Isoforms

We compared the numbers of identified isoforms and genes among the four methods. The number of isoforms affected by ieQTLs was 12,202 for the 1st and 2nd methods (LD pruning-based methods), 5,502 for the 3rd method (Bonferroni correction with all tested variants), and 43,607 for the 4th method (eigenMT). eigenMT identified the largest number of isoforms, and all isoforms identified by the first three methods were included in the results of eigenMT.

Comparison of the Number of Identified ieQTLs

Since eigenMT software reports a single variant for each isoform, we searched for variants linked to the variants detected by eigenMT ($r^2 > 0.8$) and considered them ieQTLs. We then compared the number of ieQTLs and gene eQTLs. The first method identified 33,928 ieQTLs and 12,668 gene eQTLs. The second method also identified 33,928 ieQTLs and 12,668 gene eQTLs, yielding the same results as the first method. The third method identified 17,119 ieQTLs and 6,350 gene eQTLs. The fourth method identified 298,175 ieQTLs.

Enrichment analysis

We performed an enrichment analysis of the variants detected by the third method, and the results were similar to that of the first method. We added the result of the enrichment analysis to Tables S6-S13.

Correction of the text

The correction for multiple testing strongly affects the results; however, the description of multiple results obtained by different correction methods in the text is confusing. Therefore, we provided lists of variants from ZENDO (https://zenodo.org/records/13856612) and described the results in the Supplementary Information. We also explained the four correction methods in Figure S2.

Accordingly, we added the following sentences to the Methods section (line 451): "We applied four methods to correct multiple tests (Additional file 1: Supplementary information, Figure S2). The first method used Bonferroni correction on the number of variants determined by LD pruning for each cis-window. The second method applied Bonferroni correction on the number of variants determined by LD pruning for each cis-window and then performed the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. The third method applied Bonferroni correction on the total number of variants for each cis-window and then performed the Benjamini-Hochberg FDR correction based on the number of isoforms or genes. The fourth method used eigenMT software, followed by the Benjamini-Hochberg FDR correction based on the number of isoforms or genes [40]."

The results of the second method were the same as those of the first method. The third method is more conservative and identified a smaller number of eQTLs. Therefore, we added the following sentences to the Results section (line 113) and mainly discussed the results of the first method in the revised manuscript as follows: "In this study, we applied four methods for correcting multiple tests (Additional file 1: Supplementary information, Figure S2). Two LD pruning-based methods (correction methods 1 and 2) identified the same eQTLs. Bonferroni correction with the total number of variants (correction method 3) identified a smaller number of eQTLs. The eigenMT-based method (correction method 4) identified a larger number of isoforms. Among them, the third method is conservative and thus may miss true eQTLs, while

the fourth method would be progressive. Therefore, we mainly analyzed the results of the first method in the subsequent analysis (Additional file 2: Table S3)."

2. On lines 284-286 the authors report that "A comparison of our eQTL list with GTEx eQTL and sQTL studies using lymphocyte samples demonstrated that 86.7% and 88.2% of the identified eQTLs were not detected, respectively (Additional file 1: Figure S2)."
These number are unplausible, given the sample size of the current study and likely reflect the issues with multiple testing I highlighted in point 1. I would like the GTEx analysis to be repeated using the significant lead molQTL variants identified after appropriate multiple testing correction (see point 1). For example, the authors could take their appropriately corrected lead variants from point 1 and ask how many of them are in high LD (r2 > 0.8) with lead variants from the GTEx. Furthermore, GTEx has one of the smallest sample sizes for LCLs and multiple eQTL and sQTLs studies have been performed in LCLs such as the GEUVADIS study (Lappalainen et al 2013; https://www.nature.com/articles/nature12531) or the MAGE study (Taylor et al, 2024; https://www.nature.com/articles/s41586-024-07708-2). The authors could also take re-analysed eQTL and sQTL results for the GEUVADIS dataset from eQTL Catalogue (Kerimov et al, 2023) and perform similar replication.

Our reply: Thank you for this comment.
In the previous comparison, we only compared our eQTLs with GTEx eQTLs. However, following the reviewer's comment, we performed the analysis again. Since the eQTL list from the GEUVADIS study was unavailable, we used eQTL and sQTL data from the MAGE study. We selected variants in high LD ($r^2 > 0.8$) with lead variants and compared the MAGE eQTLs and sQTLs with our gene eQTLs and ieQTLs. Of our ieQTLs, 26.8% (10,343/38,547) and 22.6% (2,976/13,158) of gene eQTLs were detected in the sQTL and eQTL data, respectively, from the MAGE study.
Accordingly, in the revised manuscript, we changed the sentence from "A comparison of our eQTL list with GTEx eQTL and sQTL studies using lymphocyte samples demonstrated that 86.7% and 88.2% of the identified eQTLs were not detected, respectively (Additional file 1: Figure S2)" to "A comparison of our eQTL list with previous eQTL and sQTL studies using lymphocyte samples demonstrated that 26.8% (10,343/38,547) of ieQTLs and 22.6% (2,976/13,158) of gene eQTLs were detected, respectively [21]" (line 131).
Finally, we compared our eQTL list with MAGE variants and variants in high LD ($r^2 > 0.8$). Since the number of linked variants was excessively large and may mislead the results of the comparison, we removed Figure S2 from the revised manuscript.

3. In the discussion (lines 484-486), the authors claim that "Overall, we believe that the functional impact of variants in splice sites cannot be accurately predicted solely by software and short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing."). While this might be true, this claim is not supported by current analysis and results presented in the manuscript. In particular, the two example associations presented in the results (rs1333973 with IFI44L and rs11026723 with GAS2) were both detected by the Leafcutter method in the eQTL Catalogue re-analysis of GEUVADIS short-read RNA-seq data (Kerimov et al, 2023) and also statistically fine mapped to a single causal variant. See RNA-seq read coverage plots here:
IFI44L:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_1%3A78620571%3A78627906%3Aclu_31999_%2B_L1

GAS2:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_11%3A22668335%3A22674850%3Aclu_40188_%2B_L1

Our reply: We apologize for the insufficient description and appreciate the reviewer remarking on the previous study, which we had not checked.
We believe the strongest advantage of long reads is the identification of full-length transcripts. Therefore, in the revised manuscript, we changed the sentence "Overall, we believe that the functional impact of variants in splice sites cannot be accurately predicted solely by software and short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing" to "Overall, we believe that the change in isoform caused by variants in splice sites cannot be accurately predicted solely by short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing" (line 336).
We also revised the following sentence: "Although a prior study suggested that IFI44L IVS2+3T/A induces splicing alterations [37], experimental validation has not been conducted" to "A prior study provided strong evidence that IFI44L IVS2+3T/A induces splicing alterations based on informatic analyses and qRT-PCR [27]. However, a mini-gene assay of the variant has not been conducted" (line 322). The previous study (reference number 37) performed qRT-PCR to evaluate isoform expression levels, but the experiment could not rule out the possibility that variants linked to IFI44L IVS2+3T/A rather than IFI44L IVS2+3T/A itself are the true causative variants. However, a mini-gene assay can directly examine the functional impact of a single variant, and we believe that such an assay provides stronger evidence of the functional impact of IFI44L IVS2+3T/A.

Furthermore, on lines 482-484, the author make this claim about the GAS2 sQTL: "Because this variant affected the expression level of only one isoform (Figure 4H), this pattern should be difficult to detect using short-read sequencing." I'm not sure what the authors mean by this, as the GAS2 sQTL can clearly be detected with the Leafcutter method in the GEUVADIS short-read RNA-seq data as I have highlighted above.
As a result, the strong claims about the discovery supposedly novel associations should either be significantly better substantiated or removed from the manuscript.

Our reply: We apologize for the unclear description. In the analysis of GAS2, the p-value for NM_177553.2 was very low and remained significant after correction for multiple tests, while the associations for other isoforms were not significant after correction. Accordingly, in the revised manuscript, we changed the sentence "Because this variant affected the expression level of only one isoform (Figure 4H), this pattern should be difficult to detect using short-read sequencing" to "Further, the impact of this variant was different among isoforms, and one isoform (NM_177553.2) was strongly affected (Figure 5C), a feature that would be difficult to detect using short-read sequencing." (line 334).

4. On lines 473-474, the authors state that "Furthermore, SpliceAI software predicted a low functional impact for the SNP(Δscores = 0.34) (Additional file 2: Table S18)".
While it is true that 0.34 is below the 0.5 threshold recommended in the SpliceAI paper, it is well above the "high recall" threshold of 0.2 and much higher than the background distribution of SpliceAI scores for random variants. In our experience, we often see that confidently finemapped sQTL causal variants have SpliceAI scores below 0.5, likely reflecting the fact that these variant also have relatively small effect sizes. Thus I believe that stating that SpliceAI

score of 0.34 as "low" is at best misleading.

Our reply: Thank you very much for this comment. Our previous description may have been misleading. Therefore, in the revised manuscript, we changed "Furthermore, SpliceAI software predicted a low functional impact for the SNP (Δscores = 0.34) (Additional file 2: Table S18)" to "Furthermore, SpliceAI did not assign a high prediction score to the SNP (Δscores = 0.34) (Additional file 2: Table S16)." (line 324). We also corrected the sentence in the abstract from "Moreover, an ieQTL analysis and minigene splicing assays unveiled functionally crucial variants in splicing, which software-based predictions failed to anticipate" to "Moreover, an ieQTL analysis and minigene splicing assays unveiled functionally crucial variants in splicing, which splicing prediction software did not assign a high prediction score" (line 44). We corrected “To investigate whether variants with low SpliceAI scores affect splicing, we chose two variants in the 3rd position of introns with low scores” to “To investigate whether variants with two SpliceAI scores < 0.5 affect splicing, we chose two variants in the 3rd position of introns” (line 242). Finally, we changed "Another variant (GAS2 IVS1+3G/A), which also exhibited a low SpliceAI Δscore (Δscores = 0.35), significantly influenced the splicing pattern" to "Another variant (GAS2 IVS1+3G/A), which SpliceAI did not assign a high prediction score (Δscores = 0.35), significantly influenced the splicing pattern" (line 329).

5. I would strongly encourage the authors to publicly release their quantified gene expression and isoform expression matrices together with the corresponding 1000 Genomes sample identifiers. If these had been public at the first review stage then we could have re-analysed their data using the eQTL Catalogue's standard molecular QTL analysis workflow (https://github.com/eQTL-Catalogue/qtlmap) to independently evaluate the concordance between molecular QTLs detected using long-read sequencing in this study and all of the available short-read eQTLs and sQTLs from the eQTL Catalogue.

Our reply: Thank you very much for this comment. We have released the raw read numbers for each isoform on ZENDO (https://zenodo.org/records/13856612). Additionally, we have made the FASTQ files available through the DDBJ database under the accession number: PRJDB17558.

Second round of review

Reviewer 2

I believe that the authors have not faithfully addressed my comments.

* The authors have still not demonstrated that their ad hoc multiple correction procedure (based on LD pruning) works as intended and correctly accounts for multiple testing. The authors should use either an established procedure (e.g. Bonferroni correction or permutations) for all of their primary analyses or provide robust evidence that their arbitrary procedure works as intended (e.g. by performing simulation to demonstrated that FDR is well calibrated under various scenarious). I agree that results that the authors obtained from the EigenMT method are too lenient, suggesting that there is either a technical issue with running EigenMT or the sample size is too small for EigenMT to work reliably.

In me previous review, I commented the following:

“3. In the discussion (lines 484-486), the authors claim that "Overall, we believe that the functional impact of variants in splice sites cannot be accurately predicted solely by software and short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing."). While this might be true, this claim is not supported by current analysis and results presented in the manuscript. In particular, the two example associations presented in the results (rs1333973 with IFI44L and rs11026723 with GAS2) were both detected by the Leafcutter method in the eQTL Catalogue re-analysis of GEUVADIS short-read RNA-seq data (Kerimov et al, 2023) and also statistically fine mapped to a single causal variant. See RNA-seq read coverage plots here:

IFI44L:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_1%3A78620571%3A78627906%3Aclu_31999_%2B_L1

GAS2:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_11%3A22668335%3A22674850%3Aclu_40188_%2B_L1
”
The updated wording still does not address this comment. In particular, the authors now write: “Further, the impact of this variant was different among isoforms, and one isoform (NM_177553.2) was strongly affected (Figure 5C), a feature that would be difficult to detect using short-read sequencing." (line 334).”

This sentence is still clearly false as the GAS2 sQTL is clearly also detected in the GEUVADIS short-read RNA-seq data. Please remove “a feature that would be difficult to detect using short-read sequencing" or provide another example that cannot be detected by short-read data.

* The DDJB accession PRJDB17558 for the raw data does not seem to exist. Could you provide a direct link to PRJDB17558 where data data could be downloaded?

Authors’ response to reviewers

* The authors have still not demonstrated that their ad hoc multiple correction procedure (based on LD pruning) works as intended and correctly accounts for multiple testing. The authors should use either an established procedure (e.g. Bonferroni correction or permutations) for all of their primary analyses or provide robust evidence that their arbitrary procedure works as intended (e.g. by performing simulation to demonstrated that FDR is well calibrated under various scenarious). I agree that results that the authors obtained from the EigenMT method are too lenient, suggesting that there is either a technical issue with running EigenMT or the sample size is too small for EigenMT to work reliably.

Reply: We deeply apologize for misunderstanding the reviewer's valuable comments.
To address the reviewer's concerns, we performed the Bonferroni correction based on the number of tested variants and FDR correction with number of isoforms or genes and reanalyzed all the data. Additionally, we removed all results obtained using the LD pruning method and eigenMT.
As a result, the total number of ieQTLs and gene eQTLs changed from 33,928 to 17,119 and from 12,668 to 6,350, respectively (please see Identification of ieQTL and gene eQTL in the Results). Accordingly, almost all eQTL counts and the results of the statistical analyses have

been updated (please see same section). The eQTLs of IFI44L, GAS2, and ATP5MK remained significant after applying the Bonferroni correction with the total number of tested variants; therefore, the results of the functional analysis were not affected.
We believe these revisions provide more reliable results for the readers. We sincerely appreciate the reviewer's insightful comments and once again apologize for the insufficiencies in our previous revision.

In me previous review, I commented the following:

“3. In the discussion (lines 484-486), the authors claim that "Overall, we believe that the functional impact of variants in splice sites cannot be accurately predicted solely by software and short-read analysis. This underscores the significant advantage of eQTL analysis employing long-read sequencing."). While this might be true, this claim is not supported by current analysis and results presented in the manuscript. In particular, the two example associations presented in the results (rs1333973 with IFI44L and rs11026723 with GAS2) were both detected by the Leafcutter method in the eQTL Catalogue re-analysis of GEUVADIS short-read RNA-seq data (Kerimov et al, 2023) and also statistically fine mapped to a single causal variant. See RNA-seq read coverage plots here:

IFI44L:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_1%3A78620571%3A78627906%3Aclu_31999_%2B_L1

GAS2:
https://elixir.ut.ee/eqtl/?credible_set=QTD000114_11%3A22668335%3A22674850%3Aclu_40188_%2B_L1
”
The updated wording still does not address this comment. In particular, the authors now write: “Further, the impact of this variant was different among isoforms, and one isoform (NM_177553.2) was strongly affected (Figure 5C), a feature that would be difficult to detect using short-read sequencing." (line 334).”

This sentence is still clearly false as the GAS2 sQTL is clearly also detected in the GEUVADIS short-read RNA-seq data. Please remove “a feature that would be difficult to detect using short-read sequencing" or provide another example that cannot be detected by short-read data.

Reply: We deeply apologize for the inaccuracies in our previous reply. As the reviewer correctly pointed out, our description was incorrect and overstated. Consequently, we have removed the following sentences from our manuscript: “Further, the impact of this variant was different among isoforms, and one isoform (NM_177553.2) was strongly affected (Figure 5C), a feature that would be difficult to detect using short-read sequencing. Overall, we believe that the change in isoform caused by variants in splice sites cannot be accurately predicted solely by short-read analysis. This conclusion underscores the significant advantage of eQTL analysis employing long-read sequencing.”
We sincerely apologize for the insufficient revision in the previous version of the manuscript.

* The DDJB accession PRJDB17558 for the raw data does not seem to exist. Could you provide a direct link to PRJDB17558 where data data could be downloaded?

Reply: Thank you for the comment. We have added a link to our data in the “Availability of data and materials” section.

Third round of review

Reviewer 2

I would like to thank the authors for carefully adressing all of my comments. I have now further concerns.