Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Myocardial splice associations after aortic valve replacement in a secondary RNA sequencing analysis

Donglin Xu
Qilu Hospital of Shandong University
Zhen Yu
Qilu Hospital of Shandong University
Ang Li
Qilu Hospital of Shandong University
Zengkun Liu
Qilu Hospital of Shandong University
Qiduo Chen
Qilu Hospital of Shandong University
Jing Liu
Qilu Hospital of Shandong University
Hourong Sun
sunhourong2016@126.com
Qilu Hospital of Shandong University

Research Article

Keywords: aortic stenosis, alternative splicing, myocardium, aortic valve replacement, reverse remodelling, RNA sequencing

Posted Date: November 17th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11547048/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Myocardial splice associations after aortic valve replacement in a secondary RNA sequencing analysis

Donglin Xu, Zhen Yu, Ang Li, Zengkun Liu, Qiduo Chen, Jing Liu and Hourong Sun

Department of Cardiovascular Surgery, Qilu Hospital of Shandong University, 107 Wenhua Xi Road, Jinan, Shandong 250012, China.

Correspondence: Hourong Sun. Email: sunhourong2016@126.com.

Abstract

Background

Left ventricular hypertrophy does not regress equally in all patients after aortic valve replacement. We examined whether myocardial splice differences associated with remodelling labels persisted after individual samples were excluded, and what these differences could change in reference transcripts and proteins.

Results

We reanalysed 41 public myocardial RNA-sequencing samples labelled as persistent hypertrophy (HYP, n = 21) or reverse remodelling (REV, n = 20). Of 156,111 candidate splice events, 70 met the full-cohort selection criteria. Twenty-two events in 19 genes were selected with unchanged direction in all 41 leave-one-out analyses. Both local splice alternatives matched reference transcripts for 16 events. A score derived separately within each training set could be calculated for every excluded sample. Its area under the receiver operating characteristic curve was 0.648, compared with 0.998 when selection and assessment used the full cohort. *DGLUCY* long-site inclusion averaged 0.766 in HYP and 0.605 in REV; the difference remained positive after each sample was removed. Complete reference sequence comparisons linked the 15-base splice-site difference to the five-amino-acid segment AGAYK. This sequence alternative is already recorded in UniProt.

Conclusions

The analysis identified internally stable myocardial splice associations and mapped selected events to reference sequences. The modest leave-one-out discrimination and unresolved correspondence between the 41 public samples and the original 40-person clinical cohort preclude a validated prognostic interpretation. Independent outcome-linked data are needed to test these associations.

Keywords: aortic stenosis; alternative splicing; myocardium; aortic valve replacement; reverse remodelling; RNA sequencing

Background

Aortic valve replacement (AVR) relieves valvular obstruction, but left ventricular hypertrophy and fibrosis can persist after surgery [1,2]. Myocardial tissue collected during AVR provides an opportunity to examine molecular differences between patients with different subsequent remodelling patterns. Whether such differences can help assess an individual patient is a separate question.

Alternative splicing changes which exons or splice sites are used to assemble an RNA transcript. It can therefore alter a transcript without requiring a change in total gene expression. Splicing changes have been reported in human heart failure [3]. Short-read RNA sequencing measures local splice usage, while reference transcripts and proteins help identify the sequence affected by each event.

We reanalysed public myocardial RNA-sequencing data from an AVR study in three steps. We first repeated event selection after excluding each sample to identify associations that did not depend on any one sample. We then assessed a score in the excluded samples to determine whether stable event selection translated into discrimination between groups. Finally, we matched the stable events to reference transcripts and examined the complete coding and protein sequences for a *DGLUCY* splice-site event.

Methods

Public samples and remodelling labels

We used public, deidentified left ventricular myocardial RNA-sequencing data from BioProject PRJNA1098010 [1,4]. The source study collected tissue during surgical AVR for severe aortic stenosis and assessed remodelling approximately one year later. Its primary endpoint was return of body-surface-area-indexed left ventricular mass to the normal sex-specific range. The article reported 40 transcriptomic participants: 21 with persistent hypertrophy and 19 with reverse remodelling. The public paired-end mRNA archive contained 41 distinct BioSamples, labelled HYP (21) or REV (20).

The public records did not identify which REV-labelled sample lay outside the reported 40-person cohort. We therefore included all 41 archived samples, using one mRNA library per BioSample and the deposited labels for grouping. We checked run accessions, BioSample accessions and sample names against the archive. Individual baseline and follow-up ventricular mass and injury-marker measurements were unavailable, so we could not verify each clinical endpoint or adjust for these measurements.

Splice-event selection

The input consisted of existing rMATS event-count tables from a GRCh38 workflow recorded as using GENCODE release 44. We did not realign the FASTQ files. The tables covered skipped exons (SE), mutually exclusive exons (MXE), alternative 5-prime splice sites (A5SS), alternative 3-prime splice sites (A3SS) and retained introns (RI). We defined each event by its class, chromosome, strand, gene identifier and genomic coordinates. Percent spliced in (PSI) ranged from 0 to 1. All group differences were calculated as HYP minus REV.

The input contained 310,255 event-by-count-representation records corresponding to 156,111 distinct events. Each analysis applied rMATS-turbo 4.3.0 to all eligible input records [5]. We adjusted P values by the Benjamini-Hochberg method separately for each event class and each count representation: junction counts (JC) and junction plus exon counts (JCEC) [6]. Additional file 1 specifies model eligibility, zero-count handling and input provenance.

A result met the strict statistical criterion when $q < 0.05$ and the absolute training-group PSI difference was at least 0.10. Tier S events met this criterion in both JC and JCEC, with the same direction and adequate count support in both. Tier A events met it in one representation and had $P < 0.05$ or $q < 0.10$ in the other, with the same direction and adequate support in the strict representation. Within each group, adequate support required valid PSI for at least 80% of samples, a median combined inclusion-plus-skipping count of at least 10, a combined count of at least 5 in at least half the samples, and no individual sample contributing more than 35% of group counts. Count-support filters were applied after multiple-testing correction.

Leave-one-out selection and scoring

We performed 41 leave-one-out analyses. In each analysis, one sample was excluded and all selection steps were repeated using the other 40 samples. Event selection, effect direction and score scaling used these training samples only. We counted how often each event met Tier S or Tier A criteria. Events selected in all 41 analyses were retained for annotation and the descriptive PSI heatmap. This final 22-event set was not substituted into the leave-one-out scoring procedure.

For each training set, the score used JC PSI for Tier S events and the strict representation for Tier A events. Each event PSI was centred on its training mean, divided by its training sample standard deviation and multiplied by the sign of its training-group difference. Zero-variance events were excluded. We averaged the available event values for each sample, requiring at least 80% coverage of scoreable events. We then centred and scaled this average using the mean and sample standard deviation of the scoreable training-sample averages. No missing values were imputed.

We evaluated the resulting excluded-sample scores using the empirical area under the receiver operating characteristic curve (AUC). HYP was the positive group, with larger scores indicating higher HYP-oriented splice usage. We also calculated the apparent AUC after selection and scoring in all 41 samples. We did not reverse scores, remove samples or tune thresholds to improve the leave-one-out AUC. Because the training sets overlap, we did not apply confidence intervals or tests that treat the cross-validation estimates as independent. Sample identifiers, event results and scores are supplied in Additional file 2.

Mapping events to reference sequences

We matched the 22 consistently selected events to Ensembl release 116 transcripts on GRCh38. After converting zero-based rMATS starts to one-based closed coordinates, we required matching chromosome, strand, target-exon boundaries and the adjacent splice boundaries that define the event. We tested both local splice alternatives and retained all matching transcript models. When a representative inclusion transcript was needed, we prioritised canonical status, protein-coding annotation and then transcript identifier. A local structural match was treated as annotation, without assuming that the corresponding full-length transcript was expressed in a study sample.

We assigned coding segments to amino-acid positions only when the complete coding length agreed with the annotated protein length, allowing for a terminal stop codon. We retained mismatched cases as flagged records rather than inferring a protein consequence. Ensembl supplied protein-feature annotations. For *DGLUCY*, we compared the complete long reference coding sequence and protein with three short reference transcripts having the closest coding-exon structure. We translated each complete coding sequence, checked it against the supplied protein and retrieved the 15-base variable genomic segment. We also checked UniProt Q7Z3D6 for an existing annotation of the amino-acid difference [7,8].

DGLUCY count checks and software

We extracted the *DGLUCY* JC and JCEC count rows and linked their sample positions to the fixed sample crosswalk. We reconstructed PSI from inclusion and skipping counts divided by their effective lengths. To assess dependence on individual samples, we removed each stored PSI value in turn and recalculated the HYP-minus-REV mean difference. This descriptive check did not refit rMATS or calculate a new significance test for the selected event.

Count modelling used the fixed rMATS executable; numerical processing and figures used Python and Matplotlib. Additional file 3 contains analysis launchers, frozen inputs and plotting code. Additional files 4 and 5 contain the two parts of the archived count input, which can be reassembled with the supplied utility. Generative AI assisted programming, manuscript editing, translation and interpretation checks. The authors

remain responsible for checking the data, analysis and claims. All figures were plotted from numerical data or reference coordinates.

Results

Twenty-two events were selected after every single-sample exclusion

Full-cohort analysis selected 70 events: 39 Tier S and 31 Tier A. Across the 41 leave-one-out analyses, 240 events were selected at least once. Twenty-two events in 19 genes were selected in every analysis with unchanged direction; 46 were selected in at least 37 analyses (Fig. 1a,b). Each training set yielded 56 to 97 selected events, with a median of 73 (Fig. 1c). The median pairwise Jaccard similarity between selected sets was 0.606, with a range of 0.378 to 0.804.

The 22 consistently selected events showed considerable variation within both groups (Fig. 2). For several events, individual samples had PSI values near 0 or 1, and the groups had overlapping distributions. The heatmap retains the archive order within each group. The adjacent group means summarise the same JC PSI values; neither display constitutes an independent classification assessment.

Leave-one-out scores overlapped between HYP and REV

All 41 excluded samples met the scoring and coverage requirements. Their mean scores were 0.091 in HYP and -0.074 in REV, a difference of 0.164 (Fig. 3a). The leave-one-out AUC was 0.648, equivalent to 272 concordant pair-equivalents among 420 HYP-REV pairs (Fig. 3b). The full-cohort apparent AUC was 0.998 (Fig. 3c). Thus, near-complete separation after full-cohort event selection did not persist when each sample was excluded from model construction.

Both local splice alternatives matched reference models for sixteen events

Eighteen of the 22 events matched at least one inclusion-compatible transcript, and 16 matched models for both local alternatives. *TTN* and *MTUS1* illustrate the limits of the mapping: the target exon was present in current reference models, but no transcript met all required local splice boundaries. We left these events

without a full-length isoform assignment. Additional file 2 lists all transcript matches and coding-length checks.

DGLUCY splice-site usage mapped to a known five-amino-acid difference

The *DGLUCY* event uses two donor sites 15 bases apart on the positive strand of chromosome 14 (Fig. 4a). Mean long-site PSI was 0.766 in HYP and 0.605 in REV, giving a difference of 0.161 (Fig. 4b). Every sample had at least 42 combined inclusion and skipping counts, and 26 samples had positive counts for both branches. Reconstructed PSI differed from the stored three-decimal values by less than 0.0005. The JC and JCEC rows were identical for this event, so they represent the same count evidence.

The long reference protein contained 621 amino acids; each of the three matched short transcripts encoded a 616-amino-acid protein. Removing bases 442-456 from the long coding sequence reproduced each complete short coding sequence. Removing residues 148-152, AGAYK, reproduced each complete short protein (Fig. 4c). UniProt already records this alternative as VSP_010506. The segment overlaps a D-glutamate cyclase domain annotation, but we did not measure protein abundance or enzyme activity.

The mean PSI difference remained positive after removal of each sample, ranging from 0.129 to 0.199 (Fig. 4d). After rounding, these differences matched the effects recorded for the event in the corresponding training analyses. The observed direction therefore did not depend on a single stored sample PSI value.

Discussion

This analysis found 22 splice events that remained selected after every single-sample exclusion, yet the leave-one-out score had an AUC of only 0.648. These results answer different questions. Repeated selection indicates that the event-level associations were not lost when any one sample was removed. The overlapping excluded-sample scores show that this stability did not yield strong individual-level discrimination.

The reference sequence analysis gives one association a precise molecular description. For *DGLUCY*, the two donor sites correspond to reference proteins differing by AGAYK. Earlier work established *DGLUCY* involvement in mitochondrial D-glutamate metabolism, including in the heart [9]. Our data do not establish

whether AGAYK changes enzyme activity or ventricular recovery. Because the sequence alternative is already annotated, the finding is an association with known splice-site usage rather than a newly discovered protein isoform.

The link to clinical remodelling also requires caution. The public archive contains one more REV-labelled sample than the cohort reported in the source article. Without the clinical inclusion key, we cannot identify the additional sample or verify its endpoint. We retained all public samples to avoid choosing exclusions based on their effect on the results. The findings therefore apply to the deposited HYP and REV labels rather than a verified reconstruction of the original 40-person endpoint cohort.

The study has several other limitations. All analyses used one small cohort, and the leave-one-out training sets shared almost all samples; their agreement is not independent replication. Whole-tissue PSI may reflect differences in cell composition as well as splice regulation. Exported counts lack mapping quality, read-start and unique-alignment information, so they cannot rule out local alignment artefacts. We did not recreate the original alignment workflow. Differences between GENCODE release 44 and the current Ensembl annotation may also affect transcript matching, and a compatible local path does not identify the full-length transcripts expressed in a patient.

Individual clinical covariates and follow-up ventricular mass were unavailable, preventing assessment of whether the splice measurements add information to clinical predictors. We also lacked an independent AVR cohort with linked outcomes and performed no functional experiments. The next test should use the fixed event coordinates in a cohort with verified sample-to-outcome correspondence. Direct measurements of the specific transcript or protein alternatives would then be needed to test their biological consequences.

Conclusions

Twenty-two myocardial splice events remained associated with deposited remodelling labels after every single-sample exclusion. Reference mapping identified both local alternatives for 16 events and linked *DGLUCY* splice-site usage to a known AGAYK sequence difference. These findings provide candidates for replication; the current data do not establish a prognostic test or a causal mechanism.

List of abbreviations

A3SS, alternative 3-prime splice site; A5SS, alternative 5-prime splice site; AUC, area under the receiver operating characteristic curve; AVR, aortic valve replacement; HYP, persistent-hypertrophy label; JC, junction count; JCEC, junction plus exon count; MXE, mutually exclusive exons; PSI, percent spliced in; REV, reverse-remodelling label; RI, retained intron; SE, skipped exon.

Declarations

Ethics approval and consent to participate

This secondary analysis was conducted in accordance with the principles of the Declaration of Helsinki. We analysed publicly available, deidentified sequencing data without recruiting participants or collecting tissue. The source study reported approval by the Ethics Committee of the University Hospital, written informed consent from all participants, and adherence to the Declaration of Helsinki [1].

Consent for publication

Not applicable. The manuscript contains no identifiable participant material.

Availability of data and materials

The sequencing dataset analysed during this study is available in the NCBI Sequence Read Archive under BioProject PRJNA1098010, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1098010 [4]. The derived datasets supporting the conclusions of this article are included within the article and its additional files. Additional file 2 contains sample identifiers, figure source data, event results, transcript matches and sequence comparisons. Additional file 3 contains code, frozen inputs, reference identifiers and retrieval records. Additional files 4 and 5 contain two byte-exact parts of the archived event-count input; the reassembly script and SHA256 check are in Additional file 3. Individual ventricular-mass measurements and the inclusion key for the original 40-person cohort were not publicly available.

Competing interests

The authors declare that they have no competing interests.

Funding

The Shandong Provincial Natural Science Foundation supported this work through grant ZR2020MH029, awarded to Hourong Sun. The funder had no role in study design, data collection, analysis, interpretation, manuscript preparation or the decision to submit.

Authors' contributions

DX contributed to conceptualisation, data curation, formal analysis, methodology, software, validation, visualisation, and drafting and revising the manuscript. ZY, AL, ZL and QC contributed to data curation, validation and revision. JL contributed to validation and revision. HS contributed to conceptualisation, funding acquisition, project administration, supervision and revision.

Acknowledgements

We thank the source-study investigators and participants, and the repositories that made the data available.

Additional files

Additional file 1. File name: Additional_file_1.docx. Format: .docx. Title: Detailed analysis methods. Description: Count modelling, event selection, leave-one-out scoring, transcript matching and sequence checks.

Additional file 2. File name: Additional_file_2.xlsx. Format: .xlsx. Title: Figure source data and extended results. Description: Sample crosswalk, selected events, training analyses, held out scores, transcript matches and *DGLUCY* counts and sequences.

Additional file 3. File name: Additional_file_3.zip. Format: .zip. Title: Analysis code and frozen inputs. Description: Stage 14-17 launchers, plotting code, frozen inputs, reassembly utility and dependencies. Standard ZIP extraction is required.

Additional file 4. File name: Additional_file_4.zip. Format: .zip. Title: Archived count input part one. Description: First byte part of the compressed event-count input. Extract together with Additional file 5 and run reassemble_counts.py from Additional file 3.

Additional file 5. File name: Additional_file_5.zip. Format: .zip. Title: Archived count input part two. Description: Second byte part of the compressed event-count input. Together with Additional file 4 it restores the original file; the utility checks SHA256.

References

1. Gabriel AF, Costa MC, Caldeira D, Plácido R, Rigueira J, Carrilho-Ferreira P, Gonçalves S, Ferreira R, Nobre Â, Pinto FJ, Enguita FJ, Almeida AG. Role of myocardial microRNAs in the long-term ventricular remodelling of patients with aortic stenosis. Eur Heart J Open. 2024;4:oeae060. doi:10.1093/ehjopen/oeae060.

2. Weidemann F, Herrmann S, Störk S, Niemann M, Frantz S, Lange V, Beer M, Gattenlöhner S, Voelker W, Ertl G, Strotmann JM. Impact of myocardial fibrosis in patients with symptomatic severe aortic stenosis. Circulation. 2009;120:577-584. doi:10.1161/CIRCULATIONAHA.108.847772.

3. Kong SW, Hu YW, Ho JW, Ikeda S, Polster S, John R, Hall JL, Bisping E, Pieske B, dos Remedios CG, Pu WT. Heart failure-associated changes in RNA splicing of sarcomere genes. Circ Cardiovasc Genet. 2010;3:138-146. doi:10.1161/CIRCGENETICS.109.904698.

4. NCBI Sequence Read Archive. BioProject PRJNA1098010. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1098010. Accessed 26 October 2026.

5. Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A. 2014;111:E5593-601. doi:10.1073/pnas.1419161111.

6. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B. 1995;57:289-300. doi:10.1111/j.2517-6161.1995.tb02031.x.

7. Ensembl. REST API and human GRCh38 reference annotation, release 116. https://rest.ensembl.org/. Accessed 26 October 2026.

8. UniProt Consortium. UniProtKB Q7Z3D6, D-glutamate cyclase, mitochondrial. Alternative sequence VSP_010506. https://www.uniprot.org/uniprotkb/Q7Z3D6/entry. Accessed 26 October 2026.

9. Ariyoshi M, Katane M, Hamase K, Miyoshi Y, Nakane M, Hoshino A, Okawa Y, Mita Y, Kaimoto S, Uchihashi M, Fukai K, Ono K, Tateishi S, Hato D, Yamanaka R, Honda S, Fushimura Y, Iwai-Kanai E, Ishihara N, Mita M, Homma H, Matoba S. D-Glutamate is metabolized in the heart mitochondria. Sci Rep. 2017;7:43911. doi:10.1038/srep43911.

Figure legends

Figure 1. Splice-event selection after single-sample exclusion

(a) Event counts by splice class for the full cohort and the 22 events selected in all 41 leave-one-out training analyses. Numbers beside points give event counts. (b) Selection frequencies for the 240 events selected at least once. (c) Event counts in each training analysis; the dashed line marks the 70 full-cohort events. Each analysis repeated selection from the full candidate universe using its 40 training samples. Training sets overlap and do not represent independent cohorts. SE, skipped exon; MXE, mutually exclusive exons; A5SS and A3SS, alternative 5-prime and 3-prime splice sites; RI, retained intron.

Figure 2. Sample-level splice usage and group means for twenty-two stable events

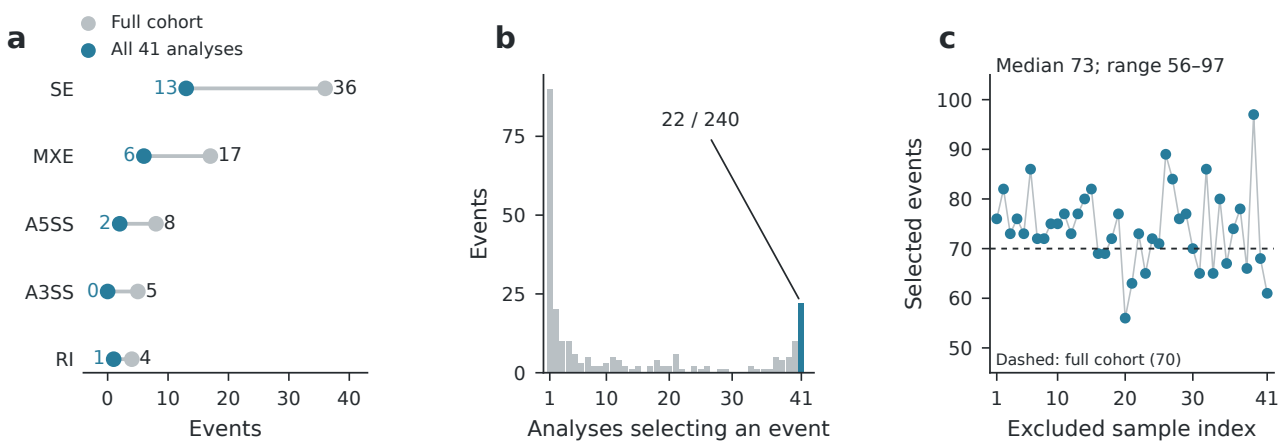
The heatmap shows JC PSI for the 22 consistently selected events across all 41 samples. Rows are ordered by gene and canonical event identifier; E01-E22 correspond to the PSI_22x41 sheet in Additional file 2. Samples retain the fixed archive order within HYP and REV. Values use the original 0-1 scale, without row standardisation. The right panel shows group means calculated from the displayed nonmissing values, with a line joining the two means for each event. These are descriptive summaries of events selected in the same cohort, not independent validation.

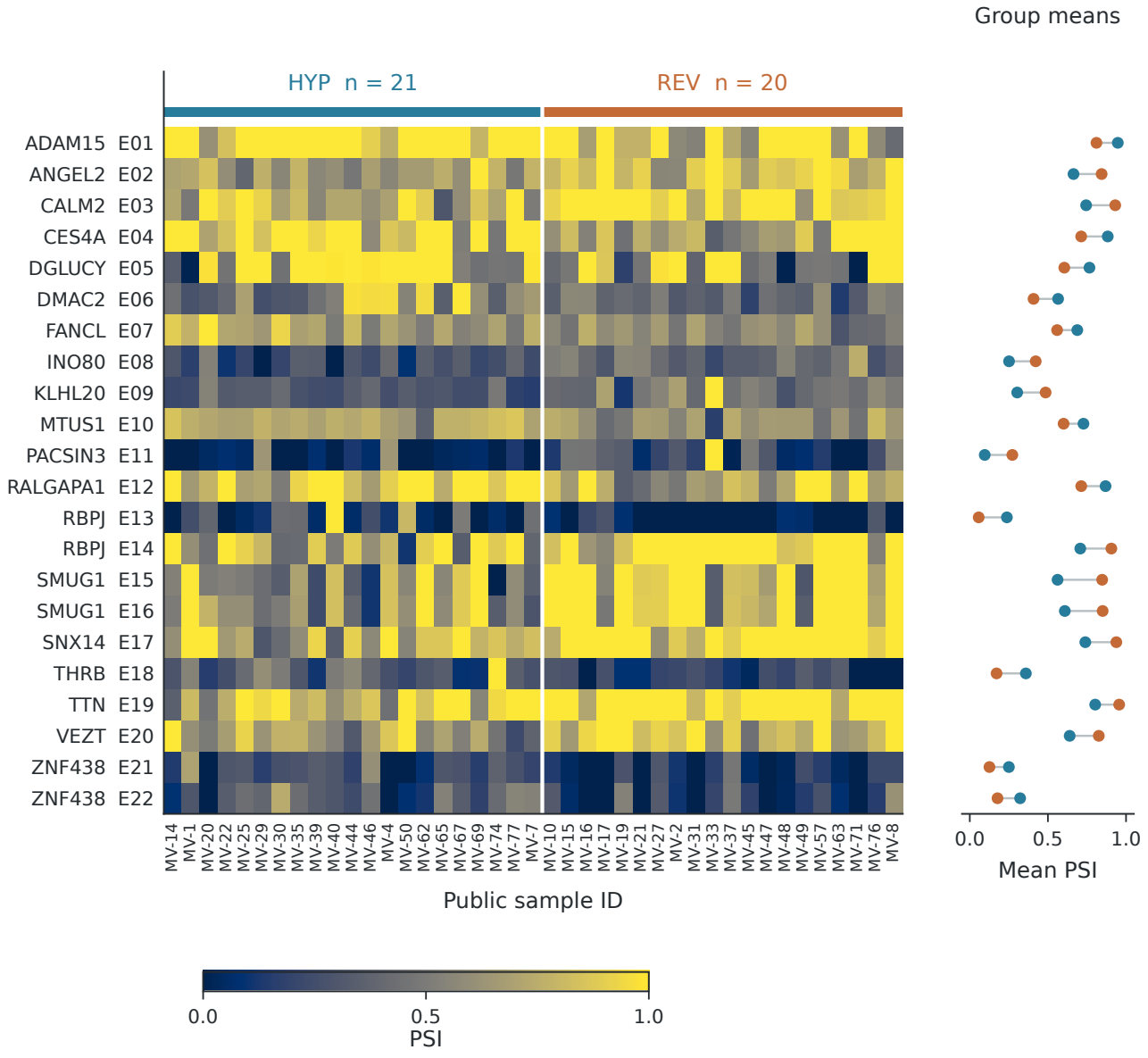
Figure 3. Leave-one-out scores provide modest group discrimination

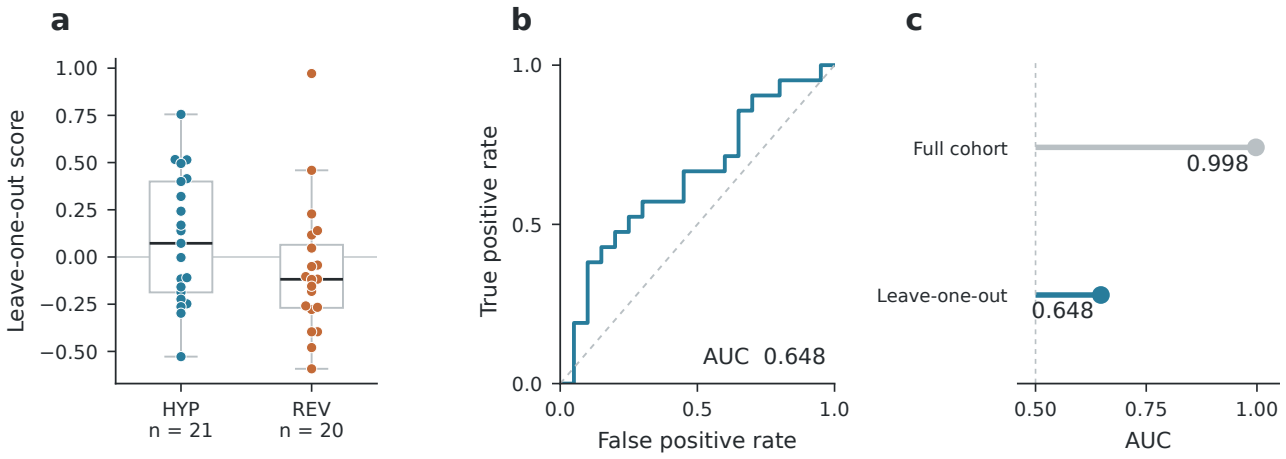
(a) Excluded-sample scores. Each point represents one sample; boxes span the interquartile range, the central line is the median and whiskers extend to observations within 1.5 interquartile ranges. All observations are plotted. (b) Empirical receiver operating characteristic curve, with HYP positive and larger scores HYP-oriented. The dashed line indicates chance discrimination. (c) Apparent full-cohort AUC and leave-one-out AUC, shown without confidence intervals. Each leave-one-out score used event selection, direction and scaling determined from the other 40 samples.

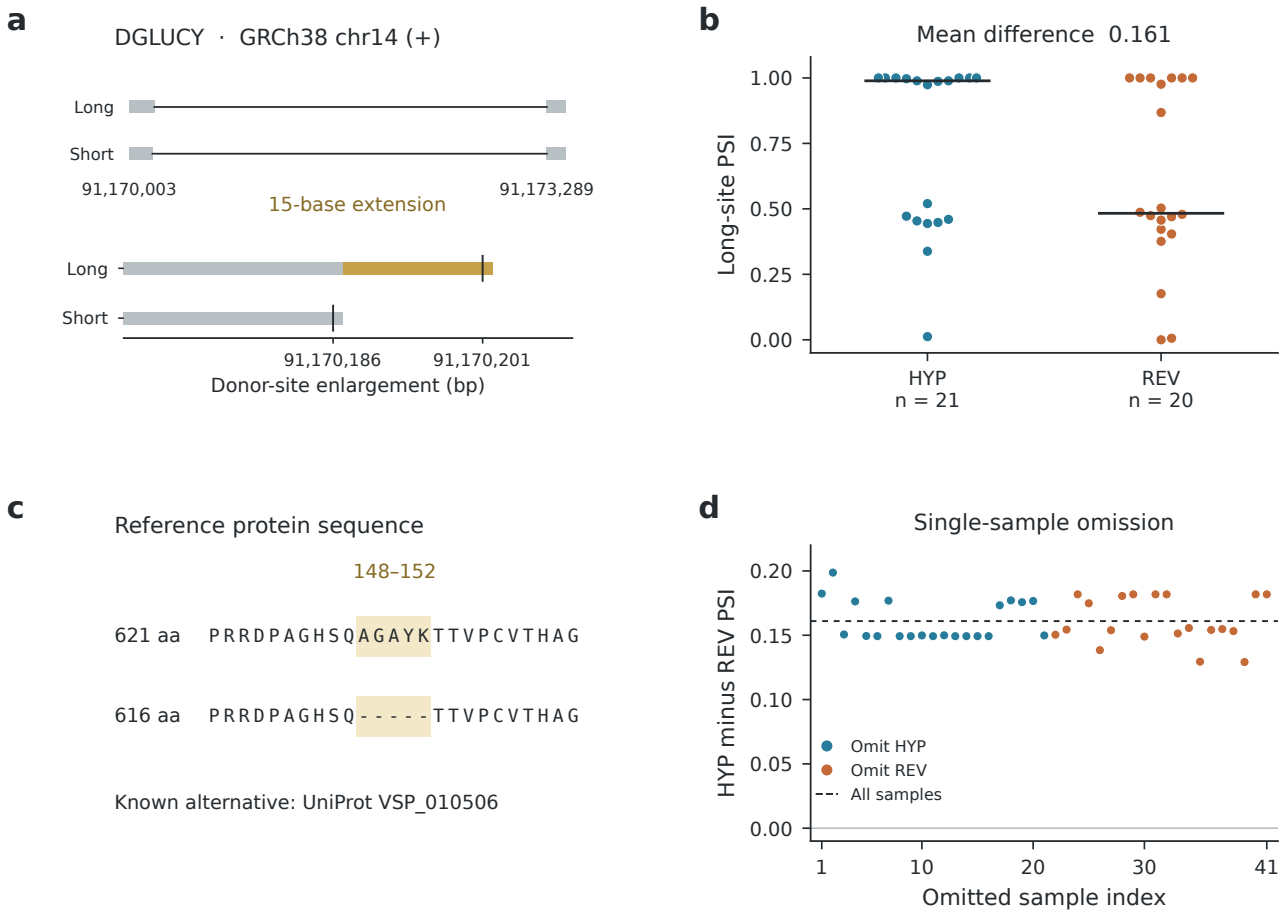
Figure 4. *DGLUCY* donor-site usage and the corresponding reference sequence difference

(a) Positive-strand GRCh38 exon coordinates. The upper track shows the local event on a common genomic scale; the lower view enlarges the two donor sites, 15 bases apart. Shared sequence is grey and the 15-base extension is gold. (b) Individual long-site PSI values, with group medians shown as black bars. (c) Reference protein alignment around AGAYK; the highlighted residues occupy positions 148-152 in the 621-residue protein. Removing these residues reproduces the complete 616-residue reference proteins. UniProt records the alternative as VSP_010506. (d) HYP-minus-REV mean PSI difference after each sample is removed. Point colour identifies the omitted sample group. The dashed line gives the full-sample difference; the solid grey line marks zero. This panel reuses stored PSI without model refitting. Reference sequence matching does not demonstrate protein expression or functional effects.


a
Full cohort
All 41 analyses
SE
13
36
MXE
6
17
A5SS
2
8
A3SS
0
5
RI
1
4
0
10
20
30
40
Events
b
22 / 240
Events
75
50
25
0
1
10
20
30
41
Analyses selecting an event
c
Median 73; range 56–97
Selected events
100
90
80
70
60
50
Dashed: full cohort (70)
1
10
20
30
41
Excluded sample index


Group means
HYP n = 21
REV n = 20
ADAM15 E01
ANGEL2 E02
CALM2 E03
CES4A E04
DGLUCY E05
DMAC2 E06
FANCL E07
INO80 E08
KLHL20 E09
MTUS1 E10
PACSIN3 E11
RALGAPA1 E12
RBPJ E13
RBPJ E14
SMUG1 E15
SMUG1 E16
SNX14 E17
THRB E18
TTN E19
VEZT E20
ZNF438 E21
ZNF438 E22
MV-14 MV-1 MV-20 MV-22 MV-25 MV-29 MV-30 MV-35 MV-39 MV-40 MV-44 MV-46 MV-4 MV-50 MV-62 MV-65 MV-67 MV-69 MV-74 MV-77 MV-7 MV-10 MV-15 MV-16 MV-17 MV-19 MV-21 MV-27 MV-2 MV-31 MV-33 MV-37 MV-45 MV-47 MV-48 MV-49 MV-57 MV-63 MV-71 MV-76 MV-8
Public sample ID
0.0 0.5 1.0
Mean PSI
0.0 0.5 1.0
PSI


a
Leave-one-out score
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
HYP
n = 21
REV
n = 20
b
True positive rate
1.0
0.5
0.0
0.0
0.5
1.0
False positive rate
AUC 0.648
c
Full cohort
0.998
Leave-one-out
0.648
0.50
0.75
1.00
AUC


a
DGLUCY · GRCh38 chr14 (+)
Long
Short
91,170,003
91,173,289
15-base extension
Long
Short
91,170,186
91,170,201
Donor-site enlargement (bp)
b
Mean difference 0.161
Long-site PSI
1.00
0.75
0.50
0.25
0.00
HYP
n = 21
REV
n = 20
c
Reference protein sequence
148–152
621 aa PRRDPAGHSQAGAYKTTVPCVTHAG
616 aa PRRDPAGHSQ-----TTVPCVTHAG
Known alternative: UniProt VSP_010506
d
Single-sample omission
HYP minus REV PSI
0.20
0.15
0.10
0.05
0.00
Omit HYP
Omit REV
All samples
1
10
20
30
41
Omitted sample index

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.xlsx
- Additionalfile3.zip
- Additionalfile5.zip
- Additionalfile4.zip