

# A novel splicing graph allows a direct comparison between exon-based and splice junction-based approaches to alternative splicing detection

Jelard Aquino<sup>1</sup>, Daniel Witoslawski<sup>1</sup>, Steve Park<sup>2</sup>, Jessica Holder<sup>1</sup>, Amei Amei<sup>3</sup>, Mira V. Han<sup>1,\*</sup>

<sup>1</sup>School of Life Sciences, University of Nevada, 4505 S Maryland Pkwy, Las Vegas, NV 89154, USA

<sup>2</sup>New York Medical College, 40 Sunshine Cottage Road, Valhalla, NY 10595, USA

<sup>3</sup>Department of Mathematical Sciences, University of Nevada, 4505 S Maryland Pkwy, Las Vegas, NV 89154, USA

\*Corresponding author. School of Life Sciences, University of Nevada, Las Vegas, NV 89154, USA. E-mail: [mira.han@unlv.edu](mailto:mira.han@unlv.edu)

## Abstract

There are primarily two computational approaches to alternative splicing (AS) detection using short reads: splice junction-based and exon-based approaches. Despite their shared goal of addressing the same biological problem, these approaches have not been reconciled before. We devised a novel graph structure and algorithm aimed at mapping between the exonic parts and splicing events detected by the two different methods. Through simulations, we demonstrated disparities in sensitivity and specificity between splice junction-based and exon-based methods. When applied to empirical data, there were large discrepancies in the results, suggesting that the methods are complementary. With the discrepancies localized to individual events and exonic parts, we were able to gain insights into the strengths and weaknesses inherent in each approach. Finally, we integrated the results to generate a comprehensive list of both common and unique AS events detected by both methodologies.

**Keywords:** alternative splicing; differential exon usage; splicing graph

## Introduction

Alternative splicing (AS) enhances the transcriptomic and proteomic diversity in eukaryotic cells by generating different isoforms through the inclusion or exclusion of specific exons. The AS process is involved in essential biological processes, such as development [1], differentiation [2], and disease [3–5]. Computational tools analyzing AS in RNA-seq data using short reads broadly fall into two categories: splice junction-based and exon-based approaches. The former identifies AS events by analyzing junction reads spanning splice sites, while the latter analyzes reads mapping to exon fragments. Popular tools like rMATS [6], SUPPA2 [7], and MAJIQ [8] use the splice junction-based approach, while others like DEXSeq [9] and JunctionSeq [10] use the exon fragment-based approach. While the splice junction-based approach has a more direct correspondence to specific AS events, it relies on fewer reads that map to fewer bases at the junctions, potentially leading to lower power in case of low coverage or low expression events. The exon-based approach considers all reads overlapping the exon of interest, including those not spanning a splice junction, so it can potentially lead to higher power relying on more coverage, but because it exhaustively tests every exonic part in the gene, unaware of the splicing structure, it leads to larger numbers of tests and suffers more from the multiple testing problem [11]. Also, the exonic part that is the unit of the exon-based approach can sometimes be involved in multiple AS events that are nested

among each other, or even alternative transcription start (ATS) or alternative transcription termination (ATT) events, so there may not be a direct one-to-one relationship between the splicing event versus the exonic part usage.

Despite these complementary strengths and weaknesses, no studies have directly compared these approaches at the event and exon levels. There have been several studies that systematically and comprehensively compared multiple AS software. However, because it is not straightforward to map the “events” of event-based tools to the “exonic parts” of exon-based tools, usually, the problem is circumvented by avoiding the direct comparison at the event/exon level. The studies either ignore exon-based methods and only compare the event-based methods [12–14] or they aggregate the results to the gene level. The aggregation from event/exon-level results to the gene-level results is done by either limiting the simulation or empirical data to genes that have exactly two transcripts and only a single splicing event [15], or by summarizing the P-values of all the events identified for each gene into a single q-value per gene [11, 16].

To bridge this gap between the splice junction-based and exon fragment-based approach, we present a novel splicing graph, GrASE (Graph of Alternative Splice junctions and Exonic parts), that produces a mapping between exonic parts and the corresponding splicing events. The capability to map between the two approaches allows two important analyses that were not

Received: May 30, 2024. Revised: November 24, 2024. Accepted: April 7, 2025

© The Author(s) 2025. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

available before. First, for the simple AS events where there is direct mapping between the two, we can compare and evaluate the performance of the two approaches at the event level. Second, we can compare the extent of consistency and complementarity of the two approaches against empirical data. Using comprehensive simulated data, for the first time, we evaluated the performance of two representative software for each approach, rMATS and DEXSeq, for their specificity and sensitivity at the splicing event/exon level. Also, using cell-line and blood samples, we show how we can integrate the results from the two software and find the events that are common and unique.

## Methods

### Overview of GrASE

We introduce GrASE, a novel splicing graph. GrASE, a directed acyclic graph, includes splice junctions as nodes and three different types of edges: exonic parts, exons, and introns. Exonic parts represent the exon counting bins that are defined by flattening the exon boundaries derived from gene annotations using `dexseq_prepare_annotation.py` of DEXSeq [9]. The main difference between traditional splice graphs and GrASE are these exonic parts edges ( $E_{ex\_part}$ ) that are arranged sequentially along genomic positions without overlap. GrASE resembles the contiguous splice graph introduced in Whippet [17], but GrASE represents junctions as nodes and exonic parts as edges, integrating the exonic parts edges with edges for regular exons and introns. Below, we describe the structure of GrASE with an example graph G. The graph G comprises five nodes representing the five splice junctions, plus artificial start (R) and end (L) nodes. The edges consist of three disjoint set of edges, the intron edges ( $E_{intron}$ ), the exon edges ( $E_{exon}$ ), and the exonic part edges ( $E_{ex\_part}$ ) that connect the junctions based on the gene annotation.

$$G = (V, E)$$

$$V = \{R, 1, 2, \dots, 5, L\} \quad E = E_{intron} \cup E_{exon} \cup E_{ex\_part}$$

$$E_{intron} = \{\{R, 1\}, \{2, 3\}, \{2, 4\}, \{5, L\}\}$$

$$E_{exon} = \{\{1, 2\}, \{3, 5\}, \{4, 5\}\}$$

$$E_{ex\_part} = \{\{1, 2\}, \{3, 4\}, \{4, 5\}\}$$

We also use the notation  $intron_{R,1}$  instead of  $\{R, 1\} \in E_{intron}$  for simplicity.

On this splicing graph, one can represent a transcript as (i) a path along the exon and intron edges similar to traditional splicing graphs or (ii) a path along the exonic parts and intron edges. For example, in Fig. 1, the same transcript ENST00000446393 can be represented as both

- i)  $SENST00000446393 = \{intron_{R,1}, exon_{1,2}, intron_{2,3}, exon_{3,5}, intron_{5,L}\}$  along the intron/exon edges and
- ii)  $S'_{ENST00000446393} = \{intron_{R,1}, ex\_part_{1,2}, intron_{2,3}, ex\_part_{3,4}, ex\_part_{4,5}, intron_{5,L}\}$  along the intron/exonic part edges.

Likewise for the alternative transcript ENST00000641506, we have

- i)  $SENST00000641506 = \{intron_{R,1}, exon_{1,2}, intron_{2,4}, exon_{4,5}, intron_{5,L}\}$  along the intron/exon edges and
- ii)  $S'_{ENST00000641506} = \{intron_{R,1}, ex\_part_{1,2}, intron_{2,4}, ex\_part_{4,5}, intron_{5,L}\}$  along the intron/exonic part edges.

The unique property of exonic part edges lies in their 1D genomic arrangement, defined by flattened exon boundaries,

ensuring nonoverlapping intervals. Consequently, for any exon defined by start and end nodes, there exists a single unique path in set  $E_{ex\_part}$ . This guarantees a one-to-one mapping between exon edges ( $exon_{i,j}$ ) and corresponding paths through exonic part edges. Examples in Fig. 1 are the mappings from the exon edges ( $exon_{3,5}$ ) to exonic part edges ( $ex\_part_{3,4}$ ,  $ex\_part_{4,5}$ ) and the mapping from the exon edges ( $exon_{4,5}$ ) to exonic part edges ( $ex\_part_{4,5}$ ).

A splicing event manifests as a bubble in the splicing graph, formed by splits where a node defining an exonic part edge has multiple outgoing/incoming degrees of exon or intron edges [18]. To detect AS, we focus on nodes connected with multiple edges, as these represent potential junctions for AS. While the concept of AS as bubbles in the graph seems straightforward in theory, detecting and classifying them into different AS types in practice can be challenging, particularly due to nested bubbles. For a long time, many software tools have only addressed simple cases and ignored complex events, although recently, progress has been made [17]. An alternative approach is to disregard the whole paths defining the bubbles and focus solely on the source and sink nodes separately. Then, you can view AS as two separate events: one defined by alternative outgoing edges for the source node and the other defined by alternative incoming edges for the sink node. This approach was adopted by Vaquero-Garcia et al. [8] with their local splicing variations.

Since the identification and classification of events are beyond the scope of our paper, we relied on existing approaches to identify and classify events. First, we utilized the splicing events that are empirically identified in a given RNAseq dataset by rMATS for downstream analyses. The events identified by rMATS encompass bubbles that have only two alternative paths between the source and sink, which we call binary bubbles. Second, we utilized the SplicingGraphs of Bioconductor [19] to identify complex bubbles that have more than two alternative paths between the source and the sink and applied a heuristic to explore a subset of the potential disjoint sets of alternative paths.

### Mapping between splicing events and exonic parts for binary bubbles

For the binary bubbles, the initial mapping was done starting from the four types of events (alternative 3' splice site [A3SS], alternative 5' splice site [A5SS], retained intron [RI], and skipped exon [SE]) found by rMATS in their `fromGTF[AS].txt` files, which were mapped to the exonic fragments defined by DEXSeq using the algorithm described below. Subsequently, we merged the mapping retrieved for all the rMATS events with the DEXSeq results.

In Fig. 1B, Node 2 splits into two outgoing intron edges and Node 5 splits into two incoming exon edges. Those alternative paths create bubbles in the splicing graph and define an AS event as a set of subpaths  $\{s_1, s_2\}$  that start at the same node (Node 2) and end at the same node (Node 5).

$$s_1 = \{intron_{2,3}, exon_{3,5}\} \text{ vs. } s_2 = \{intron_{2,4}, exon_{4,5}\}.$$

To find the mapping between the exonic part and the splicing event, we leverage the previously described mapping between the exon paths and exonic part paths. For a binary splicing event that is defined as two alternative paths along the intron/exon edges, we identify the corresponding paths along the intron/exonic part edges. In Fig. 1B, the corresponding alternative subpaths would be  $\{s'_1, s'_2\}$ ,

$$s'_1 = \{intron_{2,3}, ex\_part_{3,4}, ex\_part_{4,5}\} \text{ vs. } s'_2 = \{intron_{2,4}, ex\_part_{4,5}\}.$$

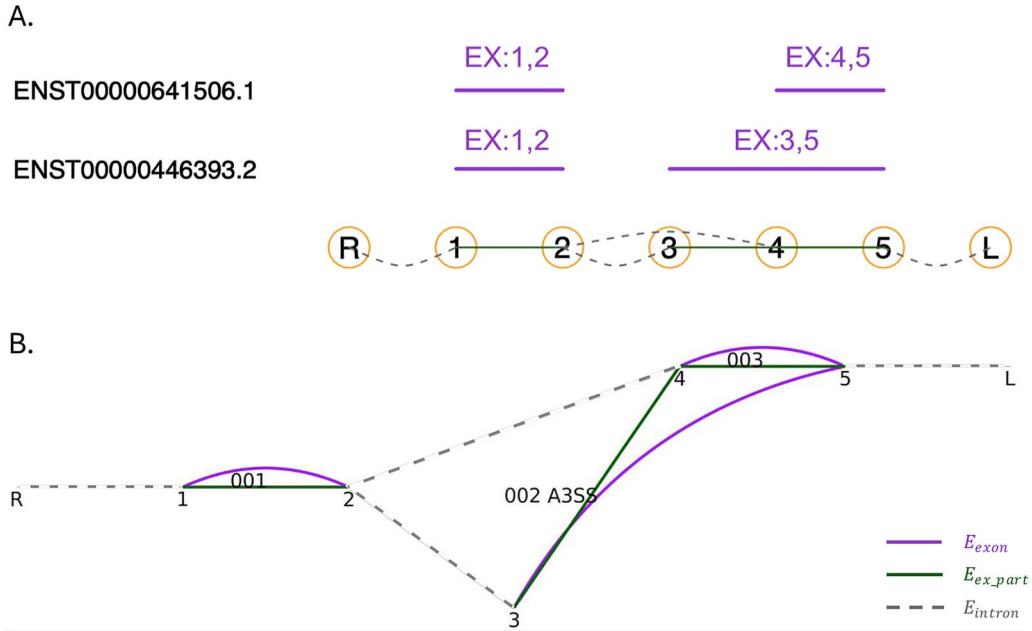


Figure 1. (A) The transcript structure of the two transcript variants of gene OR9H1P (ENSG00000228336). (B) GrASE graph representation of the gene. The nodes are the splice junctions, the intron edges ( $E_{intron}$ ) are represented by the dotted lines, the exon edges ( $E_{exon}$ ) and the exonic part edges ( $E_{ex\_part}$ ) are represented by the solid lines. The exonic part edges are also labeled by the DEXSeq exonic part ID. We can see the splicing event identified as a bubble from Node 2 to Node 5, and the differential exonic part,  $ex\_part_{3,4}$  (DEXSeq ID 002) that maps to the splicing event marked with the label A3SS.

We translate these alternative paths along the intron/exonic parts  $s'_1$  and  $s'_2$  into sets. Exonic parts not common between the two paths represent the exonic parts used differently depending on the isoform usage. Thus, we find the symmetric difference between the two sets, and the exonic part edges that are remaining in the symmetric difference are the focal exonic parts that should be tested for differential exon usage for the splicing event  $(s'_1 \Delta s'_2) \cap E_{ex\_part} = \{ex\_part_{3,4}\}$ .

We have now defined a mapping between the AS event  $\{s_1, s_2\}$  and the exonic parts  $(s'_1 \Delta s'_2) \cap E_{ex\_part} = \{ex\_part_{3,4}\}$ .

### Mapping between splicing events and exonic parts for complex bubbles

In the previous section, we described the approach of representing AS as two different sets of edges, each representing the alternative paths in the graph, and using set difference to find the focal exonic part edges that are different between the two paths. In the case of binary bubbles, there is only one way to partition the two alternative paths into two sets. However, with complex bubbles that have  $n$  ( $n > 2$ ) paths, there are potentially  $2^n - 1$  ways to partition the alternative paths into two nonempty disjoint subsets. Instead of trying to map the differential exonic parts across all the various possible partitions, we decided to use a heuristic to explore a subset of those partitions. We first found the mapping for the binary bubbles; then, we collapsed the binary bubble and reduced the number of paths systematically. As bubbles are collapsed, formerly complex bubbles became binary bubbles and were mapped recursively. This way, there are some potential partitions that are not examined, but the partitioning naturally follows the nesting hierarchy of the bubbles.

Figure 2 demonstrates an example of a complex bubble found in the gene HTR7 (ENSG00000148680), where there are three unique paths from the source (Node 4) to the sink (Node 9). This bubble can be represented as a set of three paths:  $\{4, 5, 7, 8, 9\}$ ,  $\{4, 5, 9\}$ ,  $\{4, 6, 9\}$  where each path is a list of sequential nodes spanning that path, from source to sink (Fig. 2A). The paths can

also be represented with edges.

$$\{4, 5, 7, 8, 9\} = \{exon_{4,5}, intron_{5,7}, exon_{7,8}, intron_{8,9}\}$$

$$\{4, 5, 9\} = \{exon_{4,5}, intron_{5,9}\}$$

$$\{4, 6, 9\} = \{exon_{4,6}, intron_{6,9}\}$$

Within this complex bubble, there exists a nested binary bubble with two paths,  $\{5, 9\}$ ,  $\{5, 7, 8, 9\}$ . For every binary bubble that exists in the gene, we flattened it into a single path of exonic parts. To accomplish this, we deleted the edges from the GrASE graph that exist along the binary bubble paths, e.g.  $\{intron_{5,9}\}$  and  $\{intron_{5,7}, exon_{7,8}, intron_{8,9}\}$ , and replaced them with edges that represent a flattened exonic part path from source to sink  $\{intron_{5,7}, ex\_part_{7,8}, intron_{8,9}\}$ . We recorded the number of exons flattened onto each exonic part as the weight of the edge.

For every complex bubble that is affected by the flattening of an internal binary bubble, we kept track of those changes in its paths. Three paths  $\{4, 5, 7, 8, 9\}$ ,  $\{4, 5, 9\}$ ,  $\{4, 6, 9\}$  existed before (Fig. 2A), and the flattening of an internal binary bubble  $\{5, 7, 8, 9\}$ ,  $\{5, 9\} \rightarrow \{5, 7, 8, 9\}$  reduced two paths into one, leaving two total paths  $\{4, 5, 7, 8, 9\}$ ,  $\{4, 6, 9\}$  (Fig. 2C). Once the complex bubble became binary, we used the same algorithm described in the previous section to extract the focal exonic parts  $\{ex\_part_{5,6}\}$  and  $\{ex\_part_{7,8}\}$ . We repeated this process on the updated graph, recursively collapsing all the new binary bubbles that were created through the process.

### Simulation of differential isoform expression

For the simulation of binary bubbles, we selected three genes, each containing at least two binary splicing events positioned distantly within the gene structure, to highlight the functionality of splice-junction/exon-level mapping. The gene MPL contained an A3SS event at exon 3 and an SE event at exon 14; The gene RBBP5

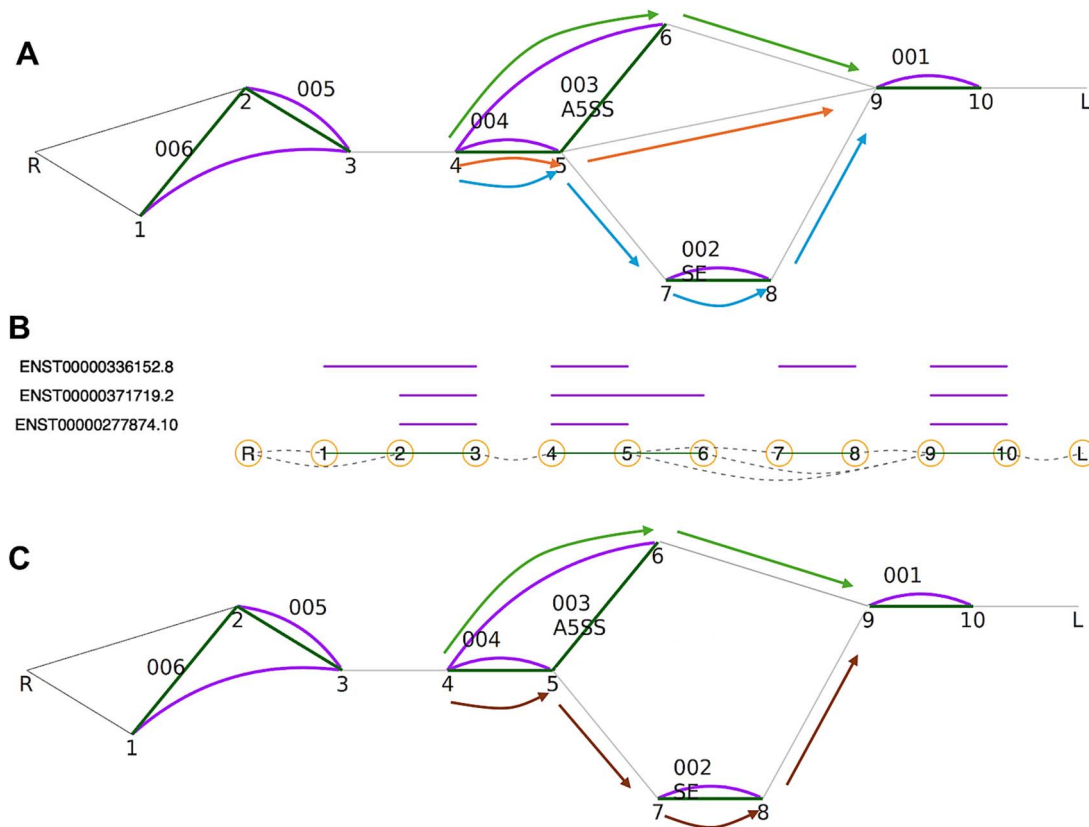


Figure 2. (A) GrASE splicing graph of the *HTR7* gene showing a complex splicing event. This graph shows the three unique paths represented with arrows that a transcript of this gene can take from Nodes 4 to 9. (B) Splicing graph of the three transcript structures of *HTR7*. The transcript ENST00000336152.8 is represented by {4,5,7,8,9}; ENST00000371719.2 is represented by {4,6,9}; and ENST00000277874.10 is represented by {4,5,9}. (C) Flattened splicing graph showing one binary bubble. The path {4,5,7,8,9} signifies the flattened path that encompasses both ENST00000336152.8 and ENST00000277874.10.

contained an A5SS event at exon 2 and an SE event at exon 15; and the gene *TRERF1* contained an A3SS event at exon 6 and two A5SS events at exon 15 and exon 23 (Supplementary Fig. S1). For *TRERF1*, we focused on the A3SS and the second A5SS event annotated in Supplementary Fig. S1. We used Polyester [20] to simulate raw RNA-seq reads for five samples each across two conditions. Polyester models isoform-level differential expression across biological replicates. For each event, we manipulated a randomly chosen single transcript encompassing the AS event to have the fold change and read depth parameters described below. In total, we generated 500 000 (2 conditions  $\times$  5 samples  $\times$  100 fold changes  $\times$  500 read depths) sets of simulated RNA-seq data for each event, covering all combinations of read depth and fold change parameters.

To determine simulation parameters, we analyzed five replicates for B-naive cells and five replicates for CD8-naive cells from Bonnal et al. (Accession ID: PRJEB5468) [21]. Based on the estimated distribution of log2 fold change between the two cell types (Supplementary Fig. S2), we designed the fold change parameters seen in Supplementary Table S1. Based on the estimated distribution of read counts mapped to transcripts for each sample (Supplementary Fig. S3), we designed a similar read depth per transcript distribution for the simulation (Supplementary Table S2). Here, the “read depth” is used as a multiplier to determine the average number of reads per transcript, using the formula  $\text{read depth} \times \text{length of transcript} / \text{length of read}$  in our simulation. For the dispersion parameter, we used two values representing large dispersion (smaller variance) and small dispersion (larger variance). For the large dispersion, we used the

default size parameter in Polyester [20] (1/3 of the transcript’s mean), and we used a value about 10-fold smaller (1/30 of the transcript’s mean) for the small dispersion.

For the simulation of complex bubbles, we selected the gene *HTR7* (ENSG00000148680.16) that contains complex bubbles. We used the same software and parameters for the simulation. To test whether we are identifying the correct exons before and after the collapsing of bubbles, we first simulated the internal binary bubble by allowing differential expression of ENST00000336152.8 versus ENST00000277874.10 and tested whether the focal exon is detected by rMATS and DEXseq. Then, we simulated the complex bubble after the collapsing of binary bubbles, where both ENST00000336152.8 and ENST00000277874.10 transcripts are partitioned and manipulated together against transcript ENST00000371719.2.

### Performance evaluation on simulated splicing events

The simulated reads were aligned using STAR against the particular gene region of the human reference genome (Gencode v34). Splice junction counts (JCs) and splice junction exon counts (JCECs) were produced using rMATS (rmats.py), while exonic part counts were produced using DEXSeq. For each event and each sample, we combined the counts from each of the 50 000 simulated runs to create a synthetic “genome” with 50 000 “genes,” where each gene corresponds to a particular simulation with a specific fold change and read depth. This aggregation resulted in a single count matrix for each event, containing 50 000 rows and 10 columns for the 10 simulated samples. We applied rMATS’s



statistical model on the JC and JCEC counts with the “—task stat” option. We ran DEXSeq on the combined exonic part counts.

For each simulated event, we utilized GrASE to identify the corresponding exonic parts mapping to the AS event. We compared the adjusted P-values provided by DEXSeq for the mapped exonic part to the False Discovery Rate (FDR) values provided by rMATS for the splicing event. The results based on the raw P-values before FDR correction are also reported for comparison. Simulations with fold change = 0 were considered real negatives, while simulations with fold change  $\neq$  0 were considered real positives for calculating sensitivity and specificity.

## Comparison of rMATS versus DEX-seq on empirical data

We constructed GrASE for all the gene models annotated in the human reference genome (Gencode v34). Using *dexseq\_prepare\_annotation.py* from DEXSeq [9], we identified all the exonic part boundaries. We used the SplicingGraphs R package [19] to construct the basic splicing graph consisting of introns and exons. We exported the basic graph as a graphml object and incorporated all the exonic part edges defined by DEXSeq using the iGraph package in python.

To compare results on empirical data, we analyzed AS between RNAseq of purified naive B cells and CD8+ naive T cells (Accession ID: PRJEB5468) [21]. The data were aligned against GRCh38 and Gencode v34 using STAR. We applied rMATS and DEXseq to quantify the counts and detect AS. Additionally, we compared RNA-seq on cell lines Raji and Jurkat for further comparison of homogenous samples.

## Results

### Direct comparison of the individual AS events detected by rMATS and DEXSeq on simulated data shows higher sensitivity for rMATS and higher specificity for DEXSeq

Using GrASE, we were able to map the events and exonic parts in our simulation and compare the detection. Here, we present the results from the small dispersion (high-variance) simulations, while results for large dispersion are available in supplementary figures (Supplementary Figs S4–S6). Sensitivity plots in Figs 3–5 illustrate a pattern where rMATS outperforms DEXSeq in detecting true positives, especially at lower read depths. The sensitivity correlates with read depth and fold change as expected. However, there is variation in the overall sensitivity across the same type of events in different genes. This variation is more pronounced for DEXseq and is not associated with the length of the exonic part as hypothesized. Instead, it follows an unpredictable pattern dependent on which transcript was chosen for over/underexpression, whether the exonic part was contained in the chosen transcript, and the length of the chosen transcript. We describe this in more detail in the section titled DEXSeq misidentifies significant exons in certain conditions. This unpredictability affects the sensitivity of DEXSeq more severely but has little effect on rMATS regardless of whether we are counting JCs or JCECs.

In contrast, the loss in sensitivity for DEXSeq was compensated by a higher specificity, especially in the small-dispersion/high-variance condition. We also observed that DEXSeq suffers more severely from multiple testing as the sensitivity drops drastically when the P-values are corrected. The observed pattern of performance was consistent across different isoform structures and different type of events.

This observed pattern is also reflected in accuracy metrics, with rMATS demonstrating higher accuracy at lower fold changes and DEXSeq demonstrating higher accuracy at higher fold changes. As expected, DEXSeq shows relatively higher accuracy than rMATS in the smaller dispersions (Supplementary Figs S7–S9) and relatively lower accuracy than rMATS in the larger dispersions (Supplementary Figs S10–S12). Overall accuracy levels are high because simulations of a small number of positive cases with different fold changes and read depths are combined with simulations of a large number of negative cases with zero fold change to calculate the accuracy. This leads to negative cases dominating accuracy measures, and specificity is weighted more importantly than sensitivity. However, we believe that this better reflects the imbalanced empirical positive rate and is a more realistic scenario.

We further investigated whether focusing on the exons/events detected by both rMATS and DEXSeq after deploying GrASE can reduce the false positives from rMATS and DEXSeq. Our simulation studies revealed that GrASE effectively reduces false positives, especially when analyzing splice junction counts from rMATS (Supplementary Fig. S13). Notably, DEXSeq’s conservative approach to multiple testing adjustments remained largely unaffected by this filtering, indicating its robustness against false discoveries. We did observe an improvement in false-positive rates when integrating GrASE’s splice junction-based and exon-based methods, particularly with rMATS’s splice junction counts, but this filtering incurred a cost by also increasing the false-negative rate substantially.

### Simulations of complex alternative splicing bubbles show that GrASE can map exons differentially used in complex alternative splicing events

The simulations showed that we can identify the focal exons for complex events. HTR7 has a complex AS event composed of two nested events, one A5SS and one SE (Fig. 2). We first simulated the internal binary bubble corresponding to the SE event and showed that both rMATS and DEXSeq detects the event under favorable parameter values (Fig. 6). Then, we collapsed the binary bubble and simulated the outer A5SS event by manipulating both ENST00000336152.8 and ENST00000277874.10 transcripts together against transcript ENST00000371719.2. We found that in this scenario, both DEXSeq and rMATS detect the A5SS event, but while DEXSeq detects the SE event as well, rMATS could not detect the SE event. This is expected, since our simulation for the complex event (A5SS) is designed to consider the binary event (SE) to be collapsed, and thus, both ENST00000336152.8 and ENST00000277874.10 were differentially expressed in tandem as a group against ENST00000371719.2. From the simulation of complex events, we were able to confirm the general pattern we observed earlier with the binary events, where rMATS showed higher sensitivity and DEXSeq showed higher specificity.

### DEXSeq misidentifies significant exons in certain conditions

In our simulation, we encountered a specific case in the skipped exon of RBBP5 (Fig. 3), where even at large fold changes of up to eight times and high read depths of 10 reads per base, DEXSeq failed to detect the event. Curious about this result, we investigated it further and discovered that such occurrences arise when the combination of exon lengths and transcript length in the gene structure results in a consistent ratio between the “this” and “others” components, even after significant fold changes. In the case

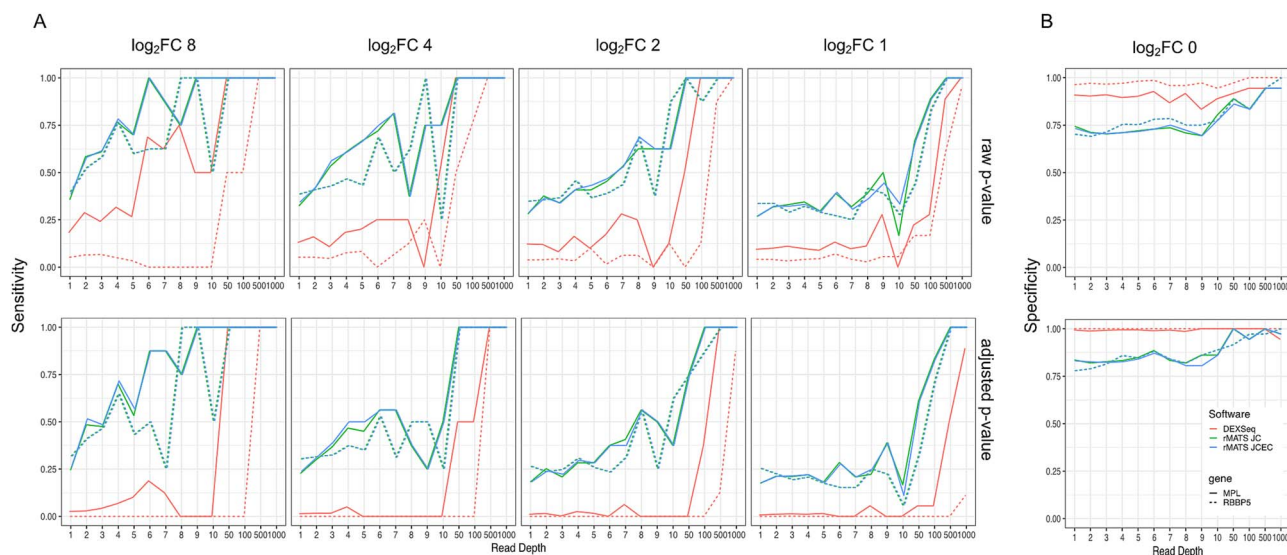


Figure 3. Sensitivity and specificity of skipped exon (SE) binary events simulated at dispersion ( $\sigma \approx 0.033$ ). (A) Sensitivity results of the SE event in genes MPL (solid line) and RBBP5 (dotted line). Each panel represents the sensitivity (y-axis) versus the read depth (x-axis) for different  $\log_2$  fold changes (8, 4, 2, 1). The lines represent the performance based on rMATS's junction counts (JCs), rMATS's junction and exon counts (JCECs), and DEXSeq. (B) Specificity results ( $\log_2$  FC 0) of the SE event in MPL and RBBP5. The top panels represent the metric based on the raw  $P$ -value of the test, while the bottom panel represents the adjusted  $P$ -value.

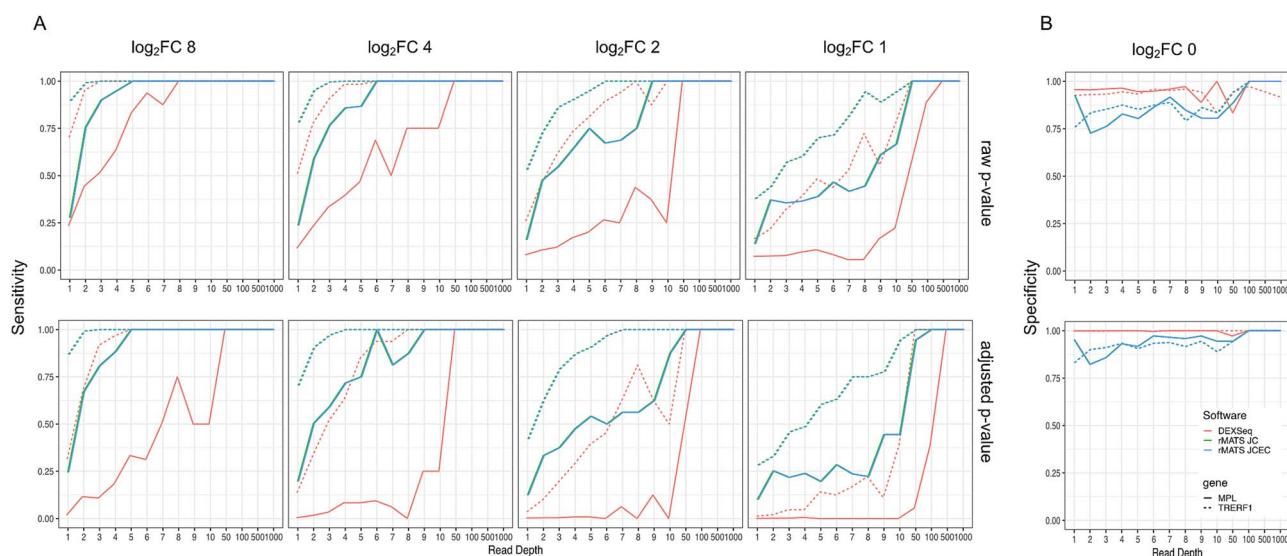


Figure 4. Sensitivity and specificity of alternative 3' splice site (A3SS) binary events simulated at dispersion ( $\sigma \approx 0.033$ ). (A) Sensitivity results of the A3SS event in genes MPL (solid line) and TRERF1 (dotted line). Each panel represents the sensitivity (y-axis) versus the read depth (x-axis) for different  $\log_2$  fold changes (8, 4, 2, 1). The lines represent the performance based on rMATS's JCs, rMATS's JCECs, and DEXSeq. (B) Specificity results ( $\log_2$  FC 0) of the A3SS event in MPL and TRERF1. The top panels represent the metric based on the raw  $P$ -value of the test, while the bottom panel represents the adjusted  $P$ -value.

of RBBP5, the manipulated transcript includes the exonic part and exhibited a large difference in overall length due to an alternative transcription termination site (TTS). Due to a particular length relationship, the ratio between “this” and “other” ended up being similar despite a large fold change.

Since this case was encountered serendipitously, we sought to identify similar instances. Utilizing a condition where the ratio between “this” and “other” remains equal after a large fold change, we developed a formula (Supplementary Fig. S14) that, when minimized, allowed us to identify potential cases where even a large fold change wouldn't lead to a significant interaction term. We selected 25 cases of gene structures with the smallest values for our condition formula based on the Gencode annotations and

simulated a favorable scenario of large fold change (8 $\times$ ) and high read depth (100) for these genes. Among the 25 cases, we found that DEXSeq could not detect the skipped exon in eight instances (Supplementary Table S3). However, in all these cases, DEXSeq did identify significant exon usage differences in the adjacent exons neighboring the skipped exon. This limitation has been discussed in the original publication, wherein they noted that the test's output with regard to which of the counting bins are affected can be unreliable if the isoform regulation affects a large fraction of the exons [22]. By directly mapping the splicing event to the relevant exon between rMATS and DEXSeq, we could detect this misidentification of the exonic part and confirm its impact even in cases with simple gene structure featuring few splicing events.

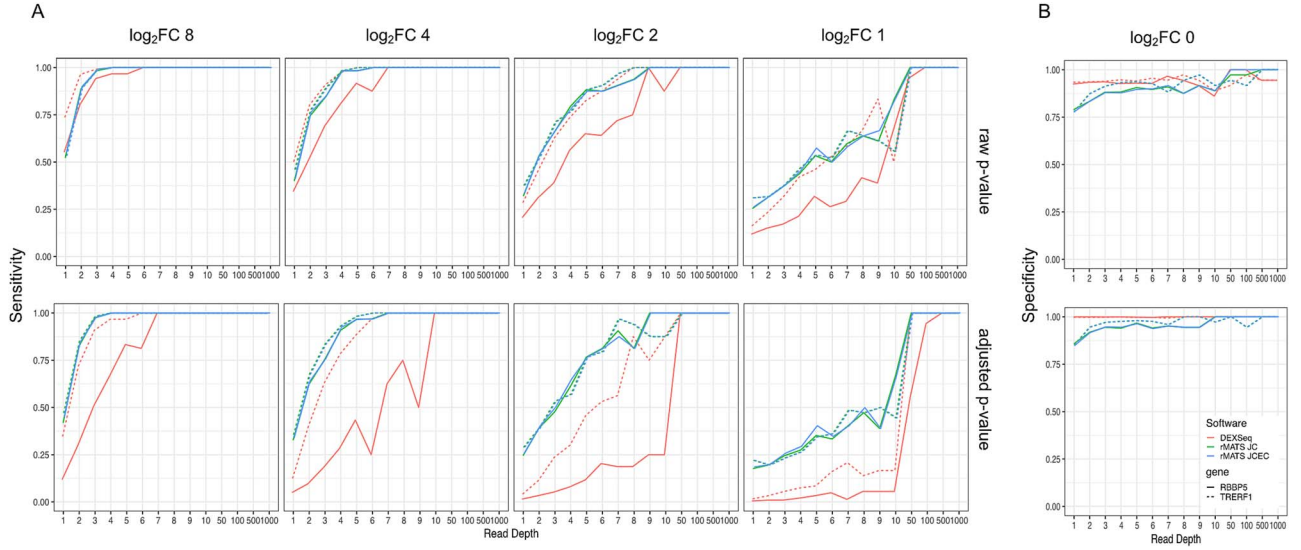


Figure 5. Sensitivity and specificity of alternative 5' splice site (A5SS) binary events simulated at dispersion ( $\sigma$ )  $\approx$  0.033. (A) Sensitivity results of the A3SS event in genes *RBBP5* (solid line) and *TRERF1* (dotted line). Each panel represents the sensitivity (y-axis) versus the read depth (x-axis) for different  $\log_2$  fold changes (8, 4, 2, 1). The lines represent the performance based on rMATS's JCs, rMATS's JCECs, and DEXSeq. (B) Specificity results ( $\log_2$  FC 0) of the A5SS event in *RBBP5* and *TRERF1*. The top panels represent the metric based on the raw P-value of the test, while the bottom panel represents the adjusted P-value.

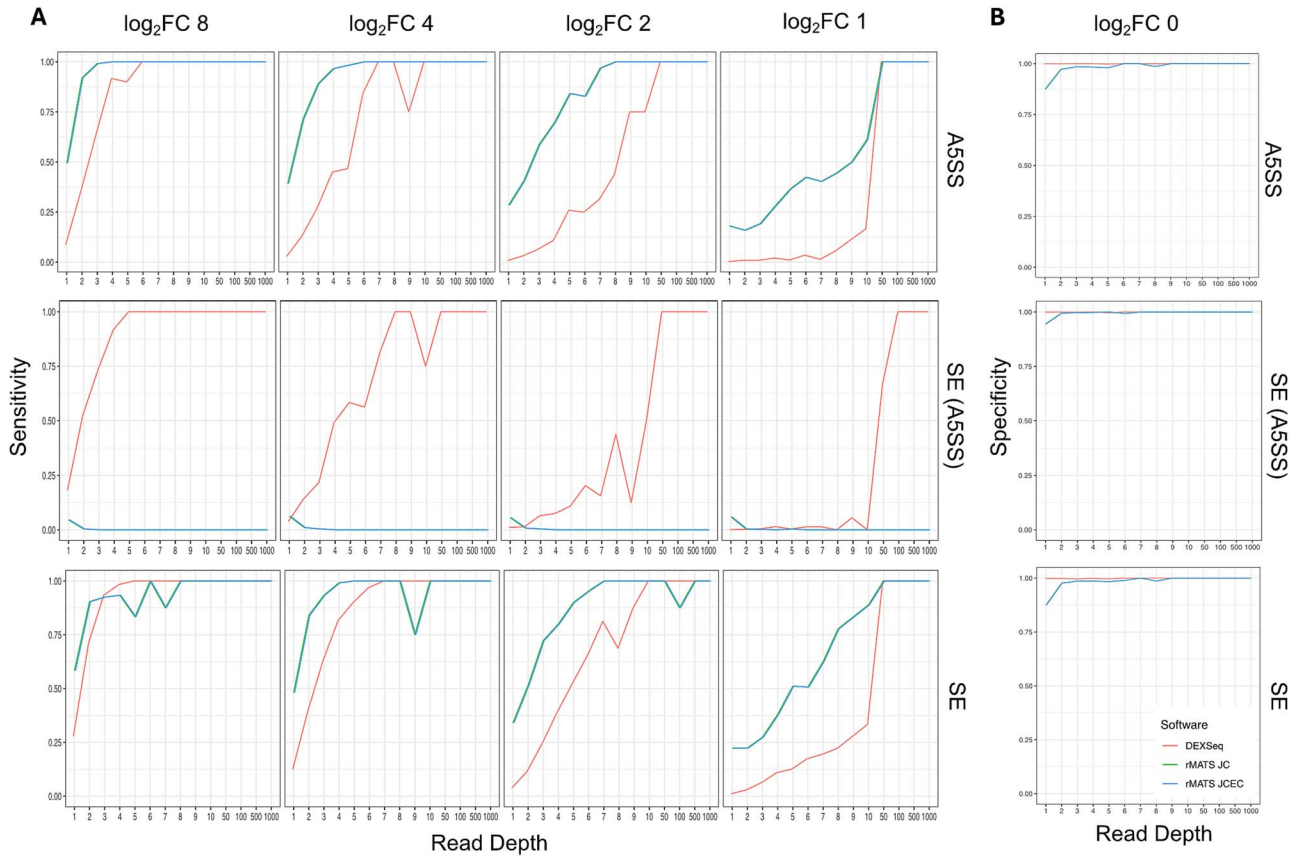


Figure 6. Sensitivity and specificity of complex events found in gene *HTR7* that include both A5SS and SE within nested bubbles. Simulations used dispersion ( $\sigma$ )  $\approx$  0.033. The top panel represents the detection of the outer A5SS event when simulation was manipulating transcripts for the A5SS event. The middle panel represents the detection of the inner SE event when simulation was manipulating transcripts for the outer A5SS event. The bottom panel represents the detection of the inner SE event when simulation was manipulating the transcripts for the SE event. (A) Each panel represents the sensitivity (y-axis) versus the read depth (x-axis) for different  $\log_2$  fold changes (8, 4, 2, 1). The lines show detection based on rMATS's JCECs and DEXSeq. (B) Specificity results ( $\log_2$  FC 0) for the events.



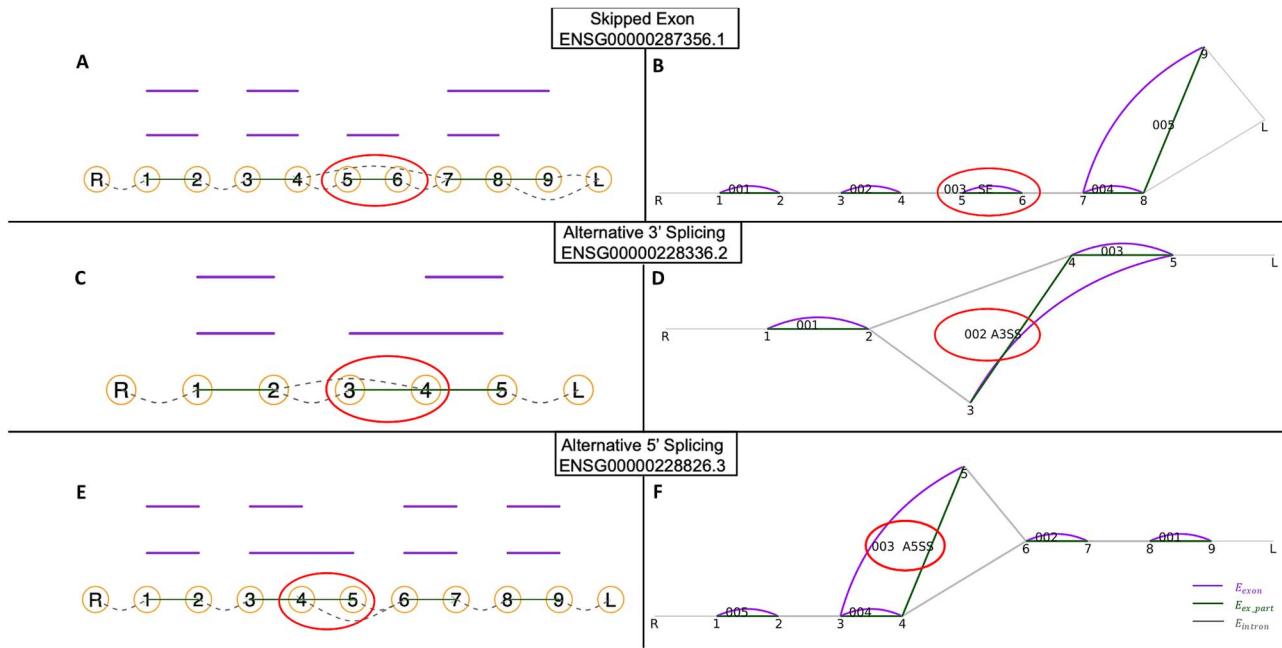


Figure 7. Mapping between the AS event and the focal exonic part. Panels on the left describe the AS event and the transcripts involved, while the right panels visualize GrASE and highlight the focal exonic part that is differentially included between the two alternative splice forms. (A, B) ENSG00000287356 with an SE event. (C, D) ENSG00000228336 (OR9H1P) with an A3SS event. (E, F) ENSG00000228826 with an A5SS event.

### GrASE allows mapping between rMATS events and DEXSeq exonic parts in empirical data

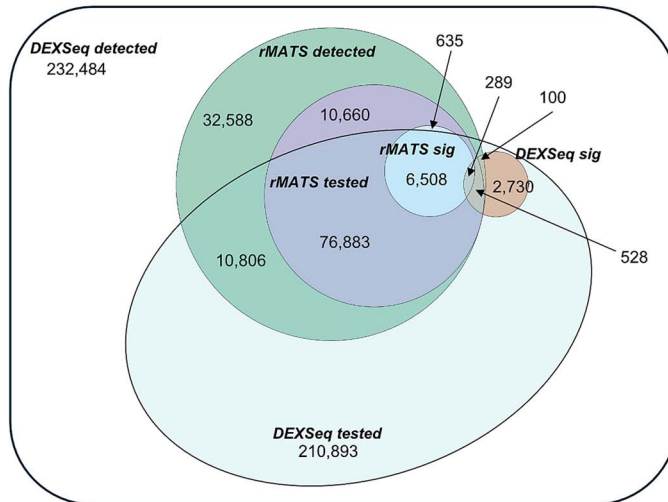
After examining simulated data, we compared the detection in empirical data. We conducted differential splicing analysis using rMATS and differential exon usage analysis using DEXSeq on RNA-seq of highly purified naive B and naive CD8<sup>+</sup> T cells from Bonnal *et al.* [21]. For all AS events detected by rMATS, we successfully identified the corresponding exonic parts that differed between the alternative splicing events. Figure 7 illustrates three genes as examples, each exhibiting different kinds of splicing events (SE, A3SS, and A5SS, respectively). Based on the Gencode v34 annotation, there were a total of 585 104 exonic parts. Comparing the naive B and naive T cells, rMATS detected 135 880 AS events, which mapped to 138 997 differential exonic parts using GrASE. Meanwhile, DEXSeq tested 308 737 exonic parts out of the 585 104 after independent filtering. Overall, there was a discrepancy in the events and exonic parts tested between the two software, with only 83 502 events (equivalent to 84 208 exonic parts) tested by both software. There was a considerable number of exonic parts that are tested by DEXSeq but were ignored by rMATS. Because DEXSeq tests all exonic parts that pass the independent filtering, the discrepancy in the total number of tests may arise not only from exonic parts that are not involved in any isoform switch but also from exonic parts that are involved in isoform switch that rMATS does not consider. The latter includes (i) alternative transcript start site (TSS) or alternative TTS [23] that rMATS disregards and (ii) complex events that are missed by rMATS [8]. The gene with the highest number of events detected was the long noncoding RNA *Linc-GALH*, which had a total of 146 different events, encompassing 178 different exonic parts in various combinations. Six genes had >100 events detected and mapped to exonic parts, including *TEX41*, *Linc-GALH*, *PVT1*, *SPG7*, *SNHG14*, and *DDX3X*, four of which are long noncoding RNAs.

### Direct comparison of the individual alternative splicing events detected by rMATS and DEXSeq on empirical data shows a large discrepancy between methods

We identified 3504 differential splicing events ( $FDR \leq 0.05$ ) through rMATS. For differential exon usage, we identified 3647 exonic parts that were differentially used (adjusted  $P$ -value  $\leq 0.05$ ) based on DEXSeq. After running each software separately and comparing the results, we mapped the splicing events from rMATS to their corresponding the exonic parts from DEXSeq using GrASE, to examine the concordance between the two software. We found that the two software are largely complementary, with only a minor proportion of significant calls from each software showing common detection (Fig. 8). Reporting on the intersection and union of events and exonic parts turned out to be slightly complicated due to the 1:n or n:1 mapping that exists between events and exonic parts, when a single event encompasses several exonic parts or when a single exonic part is involved in several different splicing events. So, here we report it in two different ways, first through the number of events, and second, through the number of exonic parts. To find the intersection, for cases of one-to-many mapping, if at least one of the many mapped events/exonic parts was called significant, we considered it to be called significant. Based on the number of events, there were a total of 135 880 events that were identified and tested by rMATS and we found corresponding exonic parts to all of those cases. Of those, 268 were called significant with both rMATS and DEXSeq, 3236 were only called with rMATS, and 1205 were only called with DEXSeq. Based on the number of exonic parts, there were a total of 308 737 exonic parts tested with DEXSeq after independent filtering, and we mapped 138 997 of them to the corresponding events detected with rMATS. Of those tested with both rMATS and DEXSeq, 289 were called significant with both rMATS and DEXSeq, 7143 were only called with rMATS, and 528



## A Number of Exonic Parts



## B Number of Splicing Events

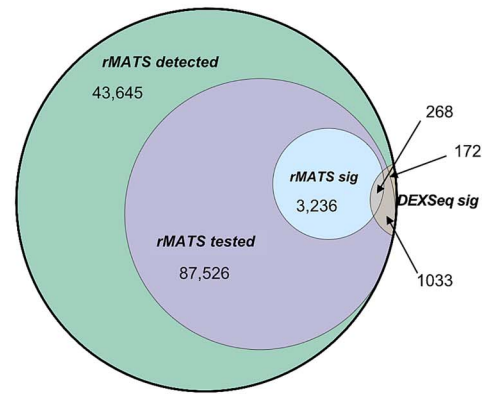


Figure 8. Venn diagrams for exonic parts and splicing events that are detected, tested, and called positive. (A) Overlapping in terms of DEXSeq exonic parts. (B) Overlapping in terms of rMATS splicing events. rMATS detected=all of the splicing events annotated by rMATS; rMATS tested=all of the splicing events rMATS statistically tested; rMATS sig=all the significant splicing events in rMATS with  $FDR \leq 0.05$ ; DEXSeq detected=all the exonic parts annotated by DEXSeq; DEXSeq tested=all exons DEXSeq statistically tested that were given an adjusted P-value; DEXSeq sig=all the significant exons DEXSeq detected with adjusted P-value  $\leq 0.05$ .

were only called with DEXSeq. Additionally, there were 100 exonic parts that were never tested by rMATS but called significant with DEXSeq.

### Reducing the number of tests DEXSeq runs to exons with potential alternative splicing events did not improve the performance for DEXSeq

As seen in our simulated results, DEXSeq exhibited lower sensitivity due to the large number of multiple tests it conducts for all the exons. We hypothesized that by reducing the number of tests that DEXSeq performs, we can boost the sensitivity and identify more exons involved in alternative splicing as differentially used. This idea was originally explored in Sonesson *et al.* where they filtered out lowly expressed transcript isoforms to reduce the number of tests and demonstrated improved performance [11]. However, in practice, the true expression level of an isoform is unknown. Here, we adopted a different approach where instead of filtering the isoforms that we cannot directly count, we filtered the exonic parts. Leveraging the mapping between the exonic parts and rMATS, we can identify the exonic parts involved in the candidate splicing events that are identified by rMATS and only test those relevant exonic parts. We tested this hypothesis by filtering the results of the DEXSeq tests to only retain the exonic parts involved in an AS event that is tested in rMATS. Subsequently, we applied DEXSeq's multiple testing correction function, including the independent filtering to the retained list of relevant exonic parts. Although the total number of exons to be tested was reduced to about one-fifth of the original list, it did not have the effect that we hypothesized. With the reduced list of relevant exonic parts, the independent filtering step of DEXSeq removed more exons from testing, but it also reduced the number of positive calls. In the end, the proportion of positives that DEXSeq detected among the relevant exonic parts were slightly lower than before filtering. (Fig. 9).

### Examination of the events missed by DEXSeq shows its lower sensitivity on exons that are part of multiple transcripts

We already know that the misidentification of the exons described in the DEXSeq Misidentifies Significant Exons in Certain

Conditions section and the multiple testing described in the Direct Comparison of the Individual Alternative Splicing Events Detected by rMATS and DEXSeq on Empirical Data Shows a Large Discrepancy between Methods section are contributing to the lower sensitivity observed in DEXSeq, but we wanted to examine the data further in relation to isoform complexity. DEXSeq fits a linear model that tests the effect of the interaction term condition:exon, where the exon is a factor with two levels: the number of reads mapping to the exonic part in question (level *this*) and the sum of the read counts from all other exonic parts of the same gene (level *others*). We formulated two hypotheses that could affect the sensitivity: first, the introduction of noise due to combining counts across many different isoforms into a summary count of "others," especially if there are multiple isoforms differentially used simultaneously; second, the collapsing of signal when the exonic count is part of multiple isoforms, and only one of the isoforms is differentially used. To test these hypotheses, we looked at the relationship between the proportion of significant events detected versus the number of transcripts that are involved in "this" versus "others."

Figure 10A shows the proportion of significant events based on the number of isoforms that a gene has. Regardless of the number of isoforms, rMATS detects more significant AS events compared to DEXSeq. However, there was no change in the difference in proportion between rMATS and DEXSeq in relation to isoform complexity. Figure 10B shows the proportion of significant events based on the number of transcripts this exonic part is a member of. Again, we confirm that rMATS detects more significant events than DEXSeq. However, here, the overall proportion is even higher for rMATS than DEXSeq when the exonic part is a member of a large number of transcripts. We divided the transcript membership into two groups,  $<30$  and  $>30$ . The difference in the proportion of significant events between rMATS and DEXSeq was significantly larger, when the exonic part was a member of  $>30$  transcripts ( $P = .04585$ ). Given that this experiment is based on empirical data without ground truth, it is hard to conclude whether the lower detection seen in DEXSeq with higher transcript complexity is false negatives for DEXSeq or false positives for rMATS. At most, we can conclude that the isoform

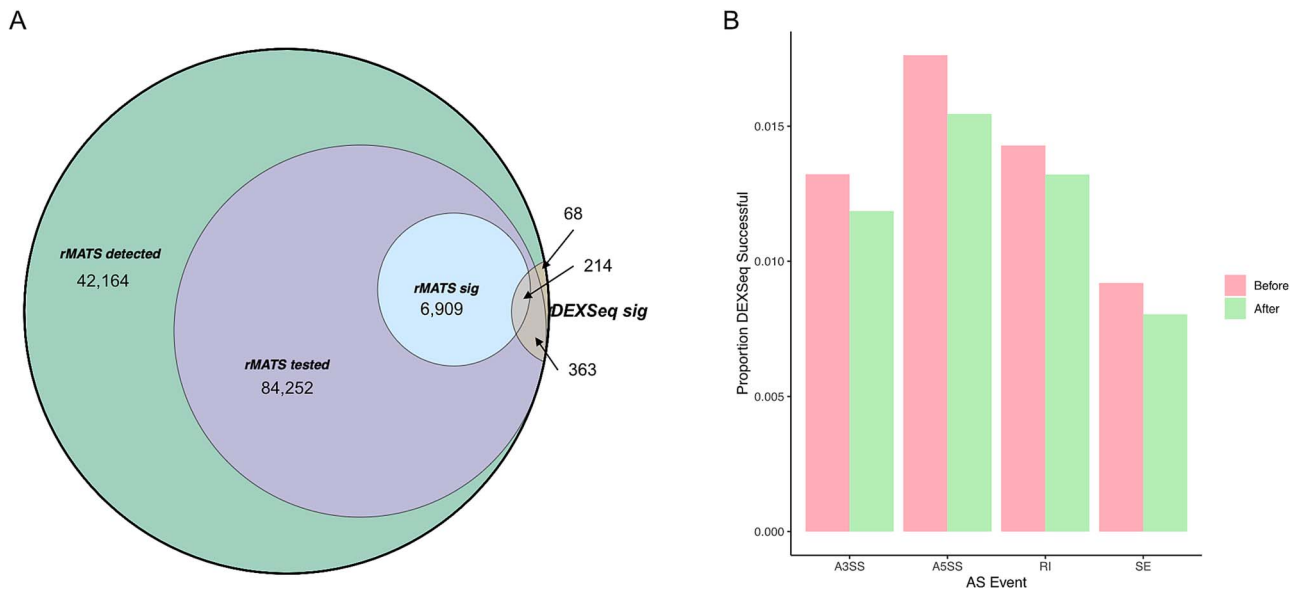


Figure 9. Direct comparison of AS events detected by DEXSeq before and after filtering for exons involved in AS. (A) Venn diagram categorizing DEXSeq exonic parts after filtering DEXSeq exons involved in AS. (B) Barplot of SEs before and after filtering for exons involved in AS. The bars represent the proportion of exons that DEXSeq detected as significant (adjusted  $P$ -value  $\leq .05$ ) over the total number of exons involved in the AS event before and after filtering the DEXSeq tested exons involved in AS.

complexity surrounding the focal exon impacts the detection differently between the two approaches.

### Examination of the events missed by rMATS shows its lower sensitivity on junctions that have low read coverage

In turn, we investigated the events that are missed by rMATS but identified by DEXSeq to understand the weakness of rMATS. One obvious difference is that exons involved in alternative TSS and TTS will be detected by DEXSeq but missed by rMATS. This is likely the largest source of discrepancy. However, if we limit ourselves just to events that are true AS events, i.e. the events that are detected and tested by rMATS, we hypothesized that the failure to detect by rMATS may be due to the sparse read coverage at splice junctions. To test this hypothesis, we categorized all the AS events tested by rMATS into three categories: events detected by both software, events detected by DEXSeq but missed by rMATS, and events detected by rMATS but missed by DEXSeq. We then examined how many events had zero read counts when summed across all five samples, which we termed zero-sum cases. Figure 10C shows the number of events with at least one zero-sum count among the four counts (B-cell inclusion, B-cell exclusion, T-cell inclusion, T-cell exclusion) for rMATS, and the number of exonic parts with at least one zero-sum count among the two counts (B cell exonic part, T cell exonic part) for DEXSeq. As expected, there were more zero-sum cases among inclusion and exclusion counts in rMATS compared to exonic part counts in DEXSeq. Nevertheless, there was an even stronger excess of zero-sum cases among events called significant by DEXSeq but missed by rMATS, with 48% of these events having at least one zero-sum count. In contrast, among events, there were called significant by rMATS but missed by DEXSeq, only 0.37% of exonic parts had at least one zero-sum count. This disparity underscores the vulnerability of rMATS to the sparsity of reads at splice junctions, a limitation not observed in DEXSeq (Fig. 10C). We also examined at the proportion of zero-sum counts among all possible counts and found a parallel trend (Fig. 10D).

### Integrating the results from the two approaches produces a comprehensive list of alternative splicing detected between B cells and T cells

Based on the mapping between the events and exonic parts, we were able to integrate the results from DEXSeq and rMATS. This produced a comprehensive list of AS events and exons that are differentially used between purified B cells and T cells, including the events and exons that are detected by both software and are high-confidence events (Fig. 11). This list included some genes that were previously reported to express various isoforms in B cells or T cells, including CD37 [24], PIK3R1 [25], IRF7 [26], and several other genes that have been reported in mice (NCF1, NOSIP, TPP1, ZFAND1, etc.) [21]. We expected to identify common genes that are differentially spliced in the comparison of B cells versus T cells and the comparison of Raji versus Jurkat cell lines, given that these cell lines derive originally from B cells and T cells. However, the list was largely different. The common genes that were found between the two comparisons were ERAP2 and WARS1. There are several genes with functions in DNA repair that were uniquely differentially spliced in the Raji versus Jurkat comparison, which included ERCC1, FANCC, FANCA, and FNA1 (Supplementary Fig. S15).

We also applied GrASE to an independent RNAseq dataset of purified B cells and T cells from Monaco et al. [27] (Supplementary Fig. S16). By comparing the results from two different software across two different independent studies, we were able to examine if events/exons detected by both software are more likely to be replicated across independent studies. Indeed, we found that the proportion of events/exons that are significant by at least one software in one study was much higher if the events/exons were detected by both software in the other study (Fig. 12). We also tested the replication of results across a third study based on long-read data by Inamo et al. [28], but the replication rate was not different across categories in this comparison (Supplementary Fig. S17).

To show that GrASE can be applied to results from additional software other than rMATS, we also applied GrASE to the results

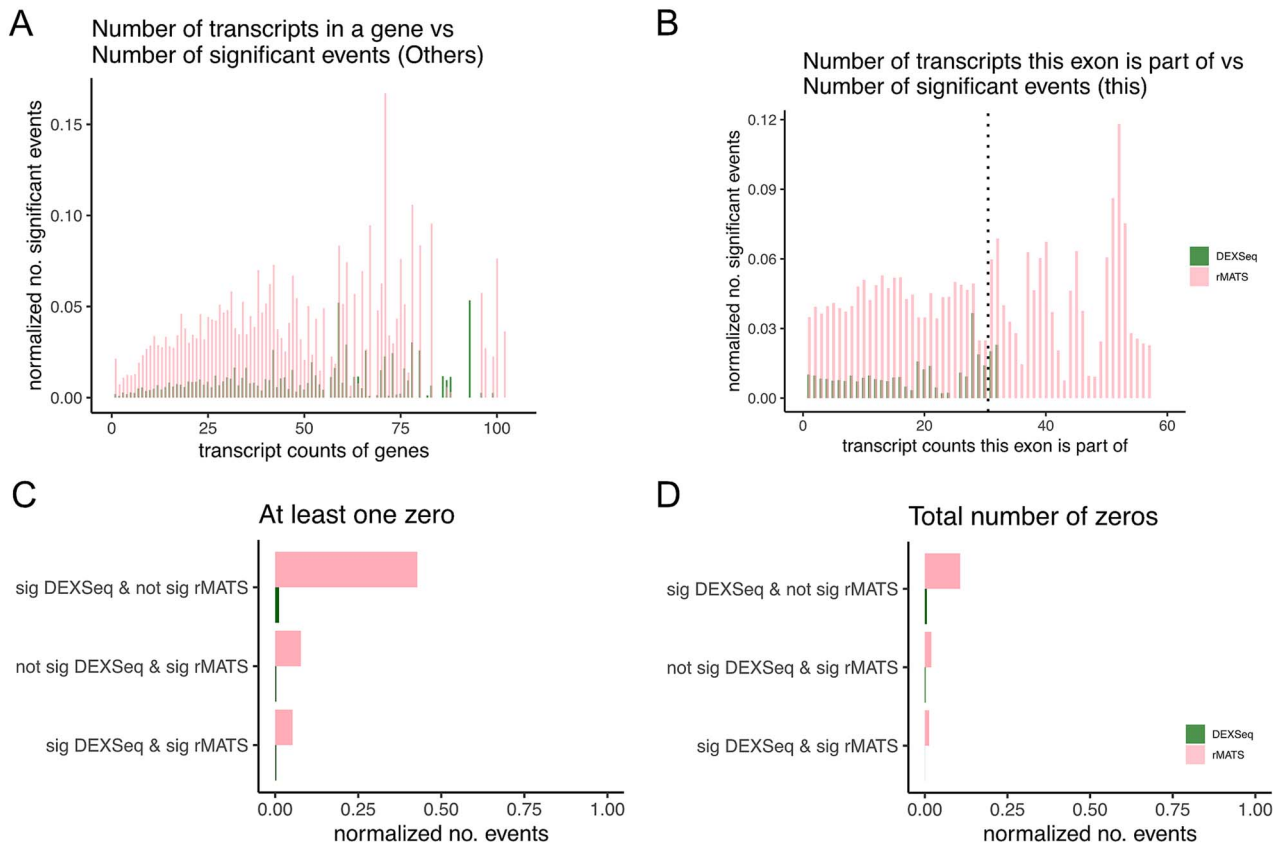


Figure 10. Detection by rMATS and DEXSeq in relation to isoform complexity and sparsity of read counts. (A) Barplot showing the number of significant AS events ( $FDR \leq 0.05$ ) detected by rMATS and DEXSeq based on the number of transcripts in a gene. (B) Barplot showing the number of significant events ( $FDR \leq 0.05$ ) based on the number of transcripts an exon is part of. The dotted line represents the division between exons part of 30 or less transcripts versus exons part of  $>30$  transcripts. (C) Barplot showing the number of events that are significant ( $FDR \leq 0.05$ ) in DEXSeq and not significant ( $FDR > 0.05$ ) in rMATS (top), not significant in DEXSeq and significant in rMATS (middle), and significant in DEXSeq and rMATS (bottom) based on the events having at least one zero sum read counts in B-naive or CD8<sup>+</sup>-naive cells. (D) Barplot showing the number of events based on the total number of zero-sum read counts in B-naive or CD8<sup>+</sup>-naive cells.

obtained by MAJIQ on the same dataset of Bonnal *et al.* Because the definitions of local splice variants used in MAJIQ are different from the events defined in rMATS, we filtered the results to only include the binary events annotated as (A5SS, A3SS, SE, and RI) by MAJIQ. This would negate the strengths of handling complex events by MAJIQ, but it would allow us to compare the results against rMATS. We used the condition that the probability of delta Percent Spliced In (PSI) is 0.20 or greater is at least 0.90 as the criterion to call the event significant for MAJIQ. Our conditions filtered out a large number of events, and we ended up with a small number of calls. The comparison showed that there was a higher proportion of significant calls in MAJIQ, if they are also called by DEXSeq or rMATS, and the number of congruent calls was larger for rMATS and MAJIQ than for DEXSeq and MAJIQ, as expected (Supplementary Fig. S18).

## Discussion

In this study, we introduced a novel splicing graph, GrASE, that offers a means to directly compare splice junction-based and exon-based methods at both the event and exon levels. Another software, Whippet [17], has proposed a similar graph that has a reverse representation of junctions as edges and exonic parts as nodes. The problem of mapping between splice junctions and exonic parts can be solved through Whippet as well, although the paper did not frame or examine the problem in such way. Our

novel approach enabled the precise localization of discrepancies between the two approaches, providing insights into their respective strengths and limitations.

DEXSeq, despite its name suggesting a focus on exons, occupies an intermediate space between exon-level and isoform-level detection. It is capable of detecting differential usage due to not only AS but also alternative TSS or alternative TTS events. Designed to test interaction between the experimental condition and *this versus others* count levels, we observed that the software will frequently detect the interaction driven by changes in “others” rather than “this,” especially when there are large changes in “others” due to the differences in overall length of the isoforms that are undergoing switches due to alternative TSS or TTS. This characteristic of the model has been clearly described in previous publications by the authors [22, 23], but this nature of the model or how much “others” can differ depending on the isoforms in empirical data may not be well understood or appreciated by the users of the software. At the gene level, DEXSeq typically performs as expected, detecting changes in the isoform usage that manifests at some of the exons. However, in some cases, the model misidentified the focal exons that were involved in the isoform switch because there is nothing in the design to prioritize change in “this” over “others.” This resulted in unpredictability in the detection of the focal exon and discrepancies between the results of event-based “versus” exon-based detection.



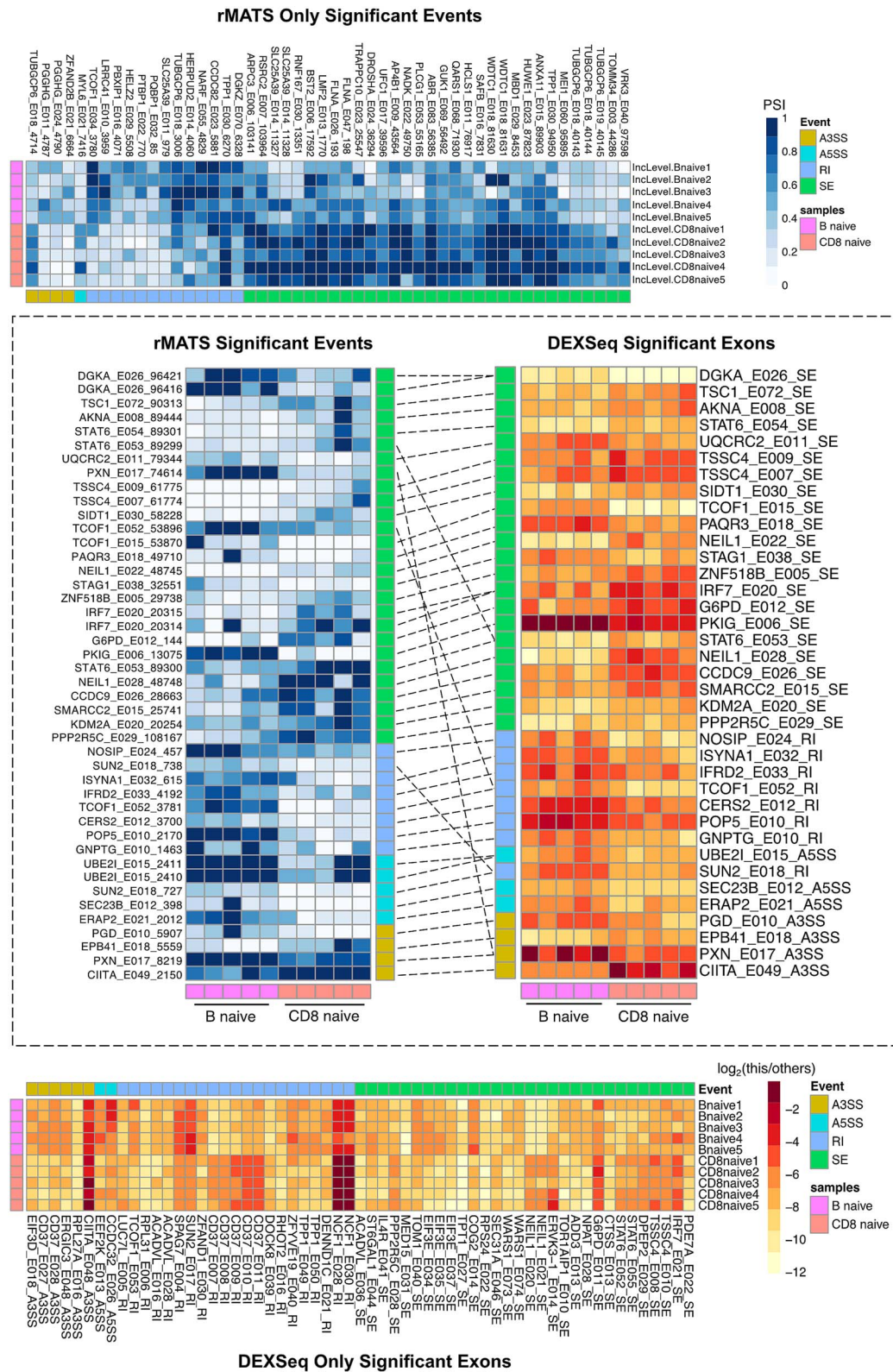


Figure 11. Heatmaps of significant genes detected by rMATS and DEXSeq on B-naive and CD8<sup>+</sup>-naive cells. The columns of the heatmap represent the genes grouped by the four AS events: A3SS, A5SS, RI, SE. The rows represent the samples: B-naive cells, CD8<sup>+</sup>-naive cells. The top heatmap displays the inclusion levels (PSI) of the significant genes (FDR  $\leq 0.05$ ) only detected by rMATS. Annotated on the columns are the gene symbols followed by the DEXSeq exon and then the unique rMATS ID. The bottom heatmap displays the this/others counts  $\log_2$ -transformed of the significant genes (adjusted P-value  $\leq 0.05$ ) only detected by DEXSeq. The gene symbols are followed by the DEXSeq exon and then the corresponding AS event. The middle heatmaps represent significant genes that are shared between rMATS (left) and DEXSeq (right) connected by dashed lines.



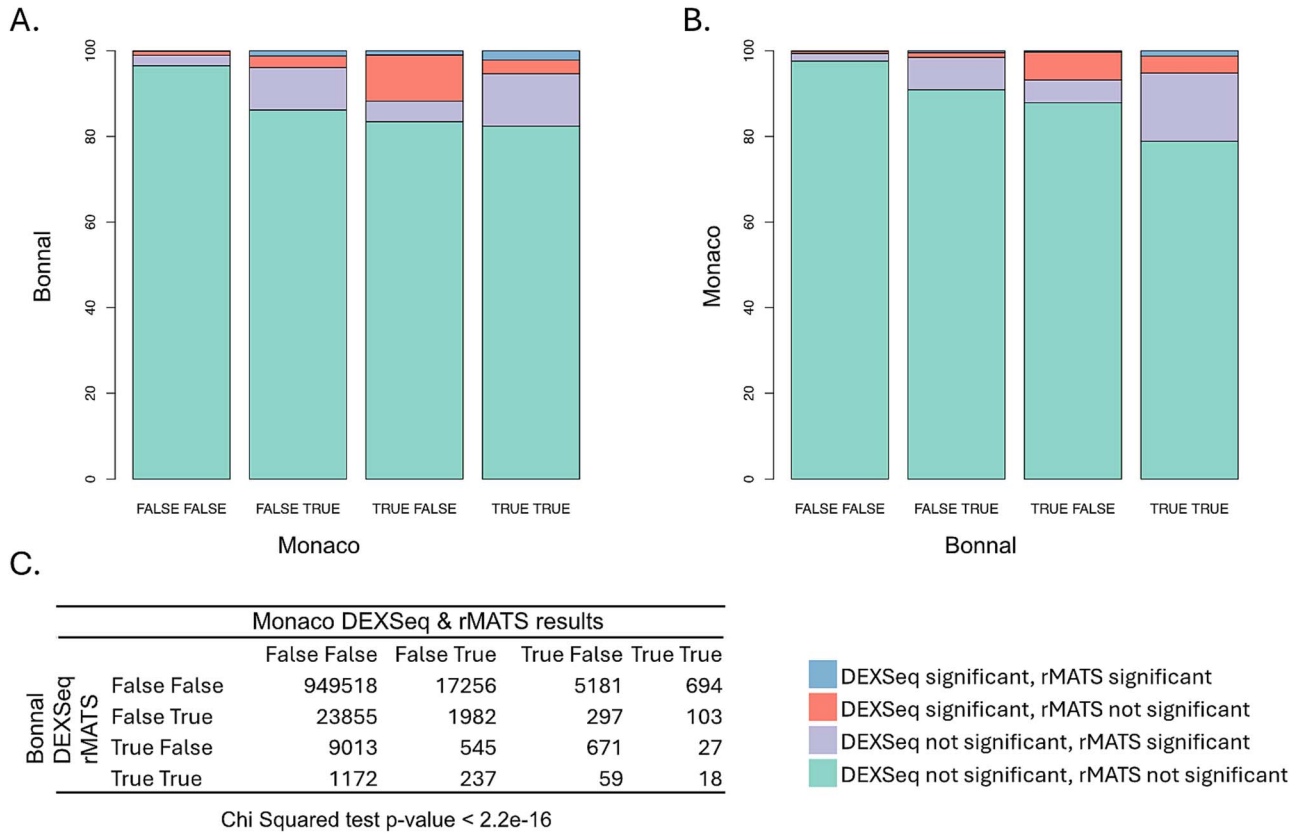


Figure 12. The proportion of exonic parts/events that are detected to be significantly different between naive B cells versus naive CD8+ T cells in two independent studies, Bonnal *et al.* and Monaco *et al.*, by two different software rMATS and DEXSeq. The first false and true indicate the significance by DEXSeq, and the second false and true indicate the significance by rMATS. (A) The proportion of significant exonic parts/events in Bonnal *et al.* based on the detection in Monaco *et al.* (B) The proportion of significant exonic parts/events in Monaco *et al.* based on the detection in Bonnal *et al.* (C) The contingency table listing the number of events in each category for each study.

On the other hand, the strengths of DEXSeq and the weaknesses of rMATS became apparent where it exhibited higher specificity in scenarios with larger variance. Although the simulation only considered dispersion as small as 1/30 of the transcript's mean, further reduction of dispersion would likely have highlighted DEXSeq's strength in specificity even more prominently. Additionally, we observed the weaknesses of rMATS when there were a small number of reads mapping to junctions as expected. This was evidenced by the prevalence of zero-sum counts among the events missed by rMATS but detected by DEXSeq. It was also seen in the incorrect estimation of PSI-level change when the exclusion counts were zero.

We also showed that by integrating the results of the two approaches, we can expand the sensitivity, or focus on the high-confidence events that have been detected by both approaches. By comparing the events detected by both software or either software in one study (Bonnal *et al.*) against the events detected from an independent dataset of a different study (Monaco *et al.*), we observed that the events and exonic parts detected by both methods are indeed more likely to be replicated by at least one software in another study. This supports the idea that it will be advantageous to prioritize the common events detected by both software for further analysis and experiments.

One limitation of this study is that we relied on the detection of events by rMATS as the starting point of our mapping between events and exons. We proposed a method to detect events from

complex bubbles as well, but most of our downstream analyses were based on the simple and unambiguous events defined by rMATS. In a complex bubble, where there are more than two paths possible, we must determine the partitioning of the set of alternative paths that will define the focal exonic parts that are different between the partitions. Our proposed method only examines a subset of those comparisons in a greedy manner. Nested bubbles can lead to uncertainty on which event is responsible for the exon usage change. On the other hand, with local splicing variants [8, 29], we give up on the information connecting the source and sink that defines the event and aim for a modest resolution at either of the junctions rather than capturing the full event. This is an inherent limitation of the short reads that will only be solved with integration of the long-read sequencing data.

In summary, we have proposed a method for bridging two methodologies in AS analysis. Through performance benchmarking, we identified strengths and weaknesses in each approach. The large discrepancy observed between the two most popular software used in the field can be understood as the large difference in the models of what they are designed to detect. It also underscores the fact that AS detection based on short read sequencing remains a challenging problem that is far from being solved. The weaknesses identified through this comparison highlights areas for improvement and will aid in improving the detection of AS utilizing short reads.

### Key Points

- The new splicing graph, GrASE, allows the mapping between local alternative splicing events and exonic parts for the first time.
- Comparing the detection of mapped events and exonic parts revealed differences in the results of splice junction-based approaches and exon-based approaches on simulated and empirical data.
- Integrating the results generated a comprehensive list of both common and unique alternative splicing events detected by both approaches.

## Supplementary data

Supplementary data are available at *Briefings in Bioinformatics* online.

## Author contributions

J.A., A.A., and M.H. conceived the experiments; J.A., D.W., S.P., and J.H. conducted the experiments; J.A., D.W., and M.H. analyzed the results. J.A., D.W., J.H., and M.H. wrote and S.P., and A.A. reviewed the manuscript.

## Funding

This work was supported by the National Institutes of Health [R16GM149510 to M.H.].

## Data availability

The codes are available online at <https://github.com/HanLabUNLV/GrASE>.

## References

1. Su C-H, Dhananjaya D, Tarn W-Y. Alternative splicing in neurogenesis and brain development. *Front Mol Biosci* 2018;**5**:12. <https://doi.org/10.3389/fmolb.2018.00012>
2. Liu H, He L, Tang L. Alternative splicing regulation and cell lineage differentiation. *Curr Stem Cell Res Ther* 2012;**7**:400–6. <https://doi.org/10.2174/157488812804484666>
3. Evsyukova I, Somarelli JA, Gregory SG. et al. Alternative splicing in multiple sclerosis and other autoimmune diseases. *RNA Biol* 2010;**7**:462–73. <https://doi.org/10.4161/rna.7.4.12301>
4. Li H, Wang Z, Ma T. et al. Alternative splicing in aging and age-related diseases. *Transl Med Aging* 2017;**1**:32–40. <https://doi.org/10.1016/j.tma.2017.09.005>
5. Bergsma AJ, van der Wal E, Broeders M. et al. Alternative splicing in genetic diseases: Improved diagnosis and novel treatment options. In: Loos F (ed.), *International Review of Cell and Molecular Biology*, Vol. **335**, pp. 85–141. Cambridge, MA, USA: Academic Press, 2018.
6. Shen S, Park JW, Lu Z. et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci* 2014;**111**:E5593–601. <https://doi.org/10.1073/pnas.1419161111>
7. Trincado JL, Entizne JC, Hysenaj G. et al. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol* 2018;**19**:40. <https://doi.org/10.1186/s13059-018-1417-1>
8. Vaquero-Garcia J, Barrera A, Gazzara MR. et al., Eds.. A new view of transcriptome complexity and regulation through the lens of local splicing variations Valcárcel J (ed.). *elife* 2016;**5**:e11752. <https://doi.org/10.7554/eLife.11752>
9. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Research*, 2012;**22**:2008–17. <https://doi.org/10.1101/gr.133744.111>
10. Hartley SW, Mullikin JC. Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. *Nucleic Acids Res* 2016;**44**:e127. <https://doi.org/10.1093/nar/gkw501>
11. Soneson C, Matthes KL, Nowicka M. et al. Isoform prefiltering improves performance of count-based methods for analysis of differential transcript usage. *Genome Biol* 2016;**17**:12. <https://doi.org/10.1186/s13059-015-0862-3>
12. Jiang M, Zhang S, Yin H. et al. A comprehensive benchmarking of differential splicing tools for RNA-seq analysis at the event level. *Brief Bioinform* 2023;**24**:bbad121. <https://doi.org/10.1093/bib/bbad121>
13. Olofsson D, Preußner M, Kowar A. et al. One pipeline to predict them all? On the prediction of alternative splicing from RNA-Seq data. *Biochem Biophys Res Commun* 2023;**653**:31–7. <https://doi.org/10.1016/j.bbrc.2023.02.053>
14. Ding L, Rath E, Bai Y. Comparison of alternative splicing junction detection tools using RNA-Seq data. *Curr Genomics* 2017;**18**:268–77. <https://doi.org/10.2174/1389202918666170215125048>
15. Liu R, Loraine AE, Dickerson JA. Comparisons of computational methods for differential alternative splicing detection using RNA-seq in plant systems. *BMC Bioinformatics* 2014;**15**:364. <https://doi.org/10.1186/s12859-014-0364-4>
16. Mehmood A, Laiho A, Venäläinen MS. et al. Systematic evaluation of differential splicing tools for RNA-seq studies. *Brief Bioinform* 2020;**21**:2052–65. <https://doi.org/10.1093/bib/bbz126>
17. Sterne-Weiler T, Weatheritt RJ, Best AJ. et al. Efficient and accurate quantitative profiling of alternative splicing patterns of any complexity on a laptop. *Mol Cell* 2018;**72**:187–200.e6. <https://doi.org/10.1016/j.molcel.2018.08.018>
18. Sammeth M. Complete alternative splicing events are bubbles in splicing graphs. *J Comput Biol* 2009;**16**:1117–40. <https://doi.org/10.1089/cmb.2009.0108>
19. Bindreither D, Carlson M, Morgan M. et al. SplicingGraphs: create, manipulate, visualize splicing graphs, and assign RNA-seq reads to them. R package version 1.48.0. 2022. <https://doi.org/10.18129/B9.bioc.SplicingGraphs>.
20. Frazee AC, Jaffe AE, Langmead B. et al. Polyester: simulating RNA-seq datasets with differential transcript expression. *Bioinformatics* 2015;**31**:2778–84. <https://doi.org/10.1093/bioinformatics/btv272>
21. Ergun A, Doran G, Costello JC. et al. Differential splicing across immune system lineages. *Proc Natl Acad Sci* 2013;**110**:14324–9. <https://doi.org/10.1073/pnas.1311839110>
22. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res* 2012;**22**:2008–17. <https://doi.org/10.1101/gr.133744.111>
23. Reyes A, Huber W. Alternative start and termination sites of transcription drive most transcript isoform differences across human tissues. *Nucleic Acids Res* 2018;**46**:582–92. <https://doi.org/10.1093/nar/gkx1165>

24. Byrne A, Beaudin AE, Olsen HE. *et al.* Nanopore long-read RNAseq reveals widespread transcriptional variation among the surface receptors of individual B cells. *Nat Commun* 2017;**8**:16027. <https://doi.org/10.1038/ncomms16027>
25. Fruman DA, Snapper SB, Yballe CM. *et al.* Impaired B cell development and proliferation in absence of phosphoinositide 3-kinase p85 $\alpha$ . *Science* 1999;**283**:393–7. <https://doi.org/10.1126/science.283.5400.393>
26. Ip JY, Tong A, Pan Q. *et al.* Global analysis of alternative splicing during T-cell activation. *RNA N Y N* 2007;**13**:563–72. <https://doi.org/10.1261/rna.457207>
27. Monaco G, Lee B, Xu W. *et al.* RNA-Seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep* 2019;**26**:1627–1640.e7. <https://doi.org/10.1016/j.celrep.2019.01.041>
28. Inamo J, Suzuki A, Ueda MT. *et al.* Long-read sequencing for 29 immune cell subsets reveals disease-linked isoforms. *Nat Commun* 2024;**15**:4285. <https://doi.org/10.1038/s41467-024-48615-4>
29. Vaquero-Garcia J, Aicher JK, Jewell S. *et al.* RNA splicing analysis using heterogeneous and large RNA-seq datasets. *Nat Commun* 2023;**14**:1230. <https://doi.org/10.1038/s41467-023-36585-y>