Title

Shiba: A unified computational method for robust identification of differential RNA splicing across platforms

Authors and affiliations

Naoto Kubota[1,2], Liang Chen[3], Sika Zheng[1,2,4]

[1]Division of Biomedical Sciences, School of Medicine, University of California, Riverside, CA 92521, USA

[2]Center for RNA Biology and Medicine, University of California, Riverside, CA 92521, USA

[3]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA 90089, USA

[4]Lead contact: sikaz@ucr.edu

Abstract

Alternative pre-mRNA splicing (AS) is a fundamental regulatory process that generates transcript diversity and cell type variation. We developed Shiba, a robust method integrating transcript assembly, splicing event identification, read counting, and statistical analysis, to efficiently quantify exon splicing levels across various types of RNA-seq datasets. Compared to existing pipelines, Shiba excels in capturing both annotated and unannotated or cryptic differential splicing events with superior accuracy, sensitivity, and reproducibility. Furthermore, Shiba's unique consideration of junction read imbalance and exon-body read coverage reduces false positives, essential for downstream functional analyses. We have further developed scShiba for single-cell/nucleus (sc/sn) RNA-seq data, enabling the exploration of splicing variations in heterogeneous cell populations. Both simulated and real data demonstrate Shiba's robustness across multiple sample sizes, including n=1 datasets and individual cell clusters from scRNA-seq. Application of Shiba on single replicates of RNA-seq identified new AS-NMD targets, and scShiba on snRNA-seq

revealed intricate temporal AS regulation in dopaminergic neurons. Both Shiba and scShiba are provided in Docker/Singularity containers and Snakemake pipeline, enhancing accessibility and reproducibility. The comprehensive capabilities of Shiba and scShiba allow systematic and robust quantification of alternative splicing events, laying a solid foundation for mechanistic exploration of functional complexity in RNA splicing.

Introduction

High-throughput RNA sequencing (RNA-seq) has been instrumental in unraveling the complexity of transcriptomes generated by alternative pre-mRNA splicing and processing. Regulatory functional splicing events play a pivotal role in shaping proteome diversity, cell type transitions, and cellular responses (1–3). The accurate detection and quantification of differential alternative splicing events from RNA-seq are fundamental prerequisites for downstream dissection of regulatory mechanisms and mechanistic characterization of functional splicing.

Several computational tools have been developed to analyze alternative splicing events from RNA-seq data, categorized into exon-centric and intron-centric tools. Exon-centric tools, such as rMATS (4, 5), SUPPA2 (6), Whippet (7), MISO (8), Vast-tools (9), JUM (10), and Quantas (11), examine differential splicing patterns at the exon level for changes in exon inclusion or exclusion across biological conditions. Intron-centric tools, such as LeafCutter (12) and MAJIQ (13, 14), delve into cross-intron splice site usages. Intron-centric tools excel in detecting complex splicing events often overlooked by exon-centric analyses. On the other hand, differential splicing (ΔPSI) derived from exon-centric analyses offers a tangible path to interpret the functional outcomes of alternative splicing regulation. The choice between exon-centric and intron-centric tools depends on the specific research question and the complexity of the alternative splicing events under investigation.

Despite their utility, results from these existing tools often have limited overlaps (15), creating confusion for researchers who focus on downstream functional analyses of alternative splicing events. The

partial intersection of analysis results from these pipelines indicates their intrinsic limitations and presents an ongoing challenge for the high accuracy and reliability of alternative splicing analyses. High confidence list is particularly important for the increasing number of investigators outside of the RNA biology fields who are interested in exploring the functional aspects of specific alternative splicing regulation. Second, existing pipelines apply statistical methods requiring many biological replicates, sometimes significantly higher than experimental practicality, to claim differential splicing events, otherwise resulting in relatively low sensitivity and high false positives. Many public bulk RNA-seq data with limited or lacking replicates are therefore shunned from rigorous splicing analyses to derive high-confidence information. Third, these statistical approaches are not applicable to single-cell/nucleus RNA-seq (sc/snRNA-seq) data. Fourth, existing tools also neglect junction read imbalances, leading to a further increase in false positives and inaccurate interpretation of splicing events. Furthermore, some existing tools, particularly those following an exon-centric approach, rely predominantly on annotated splicing events. This limitation results in the oversight of unannotated and cryptic splicing events.

The increasing amounts of sc/snRNA-seq data present a gap in developing a robust statistical framework to extract differential splicing events as reliably as possible. Ideally, high-resolution sc/snRNA-seq with sufficient sequencing depth allows exploration of splicing events at a single-cell level, offering insights into the heterogeneity of splicing patterns within a cell population. However, popular scRNA-seq platforms, such as the droplet-based 10x technology, exhibit biased and low read coverage. The second best would be splicing analyses at the cell cluster level by aggregating reads to increase sequencing depth. This is feasible and directly addresses the issue of low coverage but presents a n=1 problem for statistical analyses of differential splicing because each cell cluster is defined only once.

In this study, we present Shiba, a novel computational method specifically designed to fill the gaps in current splicing analysis tools and enhance the precision of calling differential alternative splicing using either bulk or sc/snRNA-seq data. Shiba incorporates advanced methodologies for reference-guided transcript assembly, accurate identification of both annotated and unannotated splicing events, and a meticulous statistical framework for differential splicing analysis. By leveraging these improvements,

Shiba provides researchers with a comprehensive tool for unraveling the complexity of alternative splicing dynamics in diverse biological contexts with high confidence. Applicable to both bulk RNA-seq and scRNA-seq, Shiba is a unified approach for cross-platform analyses. Through rigorous performance evaluations and comparisons with established tools, we herein demonstrate the superior efficiency and accuracy of Shiba as a valuable tool for the investigation of RNA splicing regulation.

Materials and methods

Method overview

Shiba comprises four steps: transcript assembly, detection of alternative splicing events, extraction of splice junction reads, and identification of differentially spliced events (DSEs). During the first step, mapped RNA-seq reads are assembled to build transcript structures using StringTie2 (version 2.2.1) (16). The "-G reference.annotation.gtf --conservative" option is used for each file, and the transcripts are merged with the "-G reference.annotation.gtf --merge" option. This step helps discover unannotated splicing events not described in the reference gene annotation file. In the second step, alternative splicing events are extracted from the newly assembled transcript structures and classified into five classical splicing events: skipped exon (SE), alternative five prime splice site (FIVE), alternative three prime splice site (THREE), mutually exclusive exons (MXE), and retained intron (RI). In the third step, the per-sample number of exon-exon and exon-intron junction reads are counted using RegTools (version 1.0.0) (17) and featureCounts implemented in the SubRead package (version 2.0.3) (18), respectively. In the fourth step, junction reads of replicated samples for each group are pooled together, and two-by-two tables of the count of inclusion and exclusion reads for target exons in two groups are created. Two two-by-two tables are made for SE and RI for both sides of exon-exon/exon-intron junctions. One table is made for FIVE and THREE. Four tables are made for MXE. For a given SE or RI event with enough read coverage (Either >20 on inclusion or >10 on exclusion junction in both groups by default), odds ratio (OR) and Percent Spliced In (PSI) are calculated as follows:

$$OR^{up} = \frac{J_{inc}^{up}(group\ 1)\ /\ J_{exc}(group\ 1)}{J_{inc}^{up}(group\ 2)\ /\ J_{exc}(group\ 2)}$$

$$OR^{down} = \frac{J_{inc}^{down}(group\ 1)\ /\ J_{exc}(group\ 1)}{J_{inc}^{down}(group\ 2)\ /\ J_{exc}(group\ 2)}$$

$$PSI = \frac{(J_{inc}^{up}\ +\ J_{inc}^{down})\ /\ 2}{(J_{inc}^{up}\ +\ J_{inc}^{down})\ /\ 2\ +\ J_{exc}} \times\ 100$$

where J_{inc}^{up} and J_{inc}^{down} are the number of inclusion junction reads for upstream and downstream, respectively, and J_{exc} is the number of exclusion junction reads.

For a given FIVE or THREE event with enough read coverage, OR and PSI are calculated as follows:

$$OR = \frac{J_{short}(group\ 1)\ /\ J_{long}(group\ 1)}{J_{short}(group\ 2)\ /\ J_{long}(group\ 2)}$$

$$PSI = \frac{J_{short}}{J_{short}\ +\ J_{long}} \times\ 100$$

Where J_{short} is the number of junction reads spanning the shorter intron, and J_{long} is the number of junction reads spanning the longer intron.

For a given MXE event with enough read coverage, OR and PSI are calculated as follows:

$$OR^{up-up} = \frac{J_{inc1}^{up}(group\ 1)\ /\ J_{inc2}^{up}(group\ 1)}{J_{inc1}^{up}(group\ 2)\ /\ J_{inc2}^{up}(group\ 2)}$$

$$OR^{up-down} = \frac{J_{inc1}^{up}(group\ 1)\ /\ J_{inc2}^{down}(group\ 1)}{J_{inc1}^{up}(group\ 2)\ /\ J_{inc2}^{down}(group\ 2)}$$

$$OR^{down-up} = \frac{J_{inc1}^{down}(group\ 1)\ /\ J_{inc2}^{up}(group\ 1)}{J_{inc1}^{down}(group\ 2)\ /\ J_{inc2}^{up}(group\ 2)}$$

$$OR^{down-down} = \frac{J_{inc1}^{down}(group\ 1)\ /\ J_{inc2}^{down}(group\ 1)}{J_{inc1}^{down}(group\ 2)\ /\ J_{inc2}^{down}(group\ 2)}$$

$$PSI = \frac{J_{inc1}^{up}\ +\ J_{inc1}^{down}}{J_{inc1}^{up}\ +\ J_{inc1}^{down}\ +\ J_{inc2}^{up}\ +\ J_{inc2}^{down}} \times\ 100$$

where J_{inc1}^{up} and J_{inc1}^{down} are the number of upstream/downstream inclusion junction reads for the upstream exon, and J_{inc2}^{up} and J_{inc2}^{down} are the number of upstream/downstream inclusion junction reads for the downstream exon.

Fisher's exact test is performed for each two-by-two table. The *P*-values were adjusted based on the false discovery rate using the Benjamini–Hochberg method (19). DSEs are defined as those that satisfy all of the following criteria:

- All OR > 3/2 or All OR < 2/3
- Adjusted Fisher's exact test P (q) < 0.05
- |ΔPSI| > 10

Users have the flexibility to configure parameters, such as the minimum read coverage, *q*-value cutoff, and ΔPSI, according to their preferences. In cases with multiple replicates for two groups, Shiba conducts Welch's t-test for PSI values, providing a comprehensive assessment to identify DSEs across multiple metrics. In addition to splicing analysis, Shiba incorporates standard gene expression analysis by quantifying mapped reads per gene using featureCounts. This is followed by normalization and differential expression analysis using DESeq2 (version 1.34.0) (20) and simultaneous calculation of TPM and CPM values, offering additional metrics for diverse downstream analyses.

Shiba has an alternative version for splicing analysis using sc/snRNA-seq data, named scShiba. The processing steps are similar to those of Shiba: Detection of alternative splicing events, pooling splice junction reads for each cell type, and identification of DSEs. First, alternative splicing events are extracted from the gene annotation file and classified into four classical alternative splicing events (SE, FIVE, THREE, and MXE). Second, taking the junction read count matrix produced by STARsolo (21) and a table of cell barcodes and the corresponding cell types as inputs, reads mapped to each junction are pooled for each cell type so each cell type can be treated as a biological replicate in the following statistical analysis. Finally, the PSI of each splicing event is calculated, and differential splicing analysis is performed in the same way as in Shiba.

All the necessary software and scripts required to run the Shiba/scShiba pipeline are provided as a single Docker/Singularity container, ensuring a highly reproducible analysis environment. Furthermore, we offer the SnakeShiba pipeline, based on Snakemake workflow (22), facilitating accelerated execution through optimized parallelization.

Simulated RNA-seq data

We executed Polyester (version 1.22.0) (23) to generate three control and three perturbed RNA-seq samples (each with 101 bp paired-end reads) in a Singularity container (naotokubota/polyester:1.0 in Docker Hub). Different base read coverages (x30, x20, x10) were set. We randomly selected 3,000 mouse protein-coding genes that have alternative splicing events (1750, 500, 500, 50, and 200 for SE, FIVE, THREE, MXE, and RI, respectively) and divided them into three groups: (i) 1,000 genes were unaltered, (ii) 1,000 genes had inclusion isoforms with 5-fold higher expression, and (iii) 1,000 genes had inclusion isoforms with 5-fold lower expression in the perturbed samples. These groups were used as the ground truth to evaluate the performance of various computational tools for event-based splicing analysis. The simulated reads were generated in FASTA format and converted to FASTQ format using seqtk seq function (version 1.4) (https://github.com/lh3/seqtk) with the "-F I" options. These reads were mapped to the mouse reference genome (mm10) and the Ensembl 102 transcript annotations using STAR (version 2.7.10b) (24) with the "--outFilterMultimapNmax 1" option. To evaluate the Shiba performance at detection of unannotated alternative splicing events, we customized the reference gene annotation file by removing records of inclusion isoforms. For the performance evaluation of detecting alternative splicing events that show junction read imbalance, reads mapped on the left side of exon-exon junctions for skipped exons (upstream/downstream for genes encoded in the plus/minus strand, respectively) were discarded from 587 and 588 genes that show the same and different expression levels of inclusion isoforms between the two groups using bedtools (version 2.30.0) (25). The mapped reads were converted back to FASTQ format using samtools (version 1.18) (26) and used for remapping to the mouse genome.

Real RNA-seq data

For the performance evaluation, RNA-seq samples from cytoplasmic (Cyp) fraction, nucleoplasmic (Nup), and chromatin (Chr) of mouse embryonic stem cells (mESC) with siRNA-mediated double knockdown (KD) of *Ptbp1* and *Ptbp2* were downloaded from GEO (GSE159993) (27). We performed quality control using fastp (version 0.23.2) (28) with the "-3 -q 30" options to discard low-quality reads. We aligned the remaining reads to the mouse reference genome (mm10) and the Ensembl 102 transcript annotations using STAR with the "--outFilterMultimapNmax 1" option.

For the identification of alternative exons associated with nonsense-mediated mRNA decay, RNA-seq samples of HEK293 Flp-In T-Rex cells with siRNA-mediated double KD of *XRN1* and *UPF1* were downloaded from GEO (GSE57433) (29). We aligned the high-quality reads to the human reference genome (hg38) and Ensembl 106 transcript annotations using STAR with the "--outFilterMultimapNmax 1" option. The mapped reads were analyzed by Shiba with the default settings to identify DSEs between the control and the KD sample.

For the identification of neuronal cell class-specific alternative splicing events, RNA-seq of tagged-ribosomal affinity purification (RiboTRAP) samples from genetically-defined neurons were downloaded from GEO (GSE133291) (30). The RiboTRAP samples were generated using conditionally HA-tagged endogenous ribosomal protein Rpl22 in glutamatergic (excitatory) neurons (CamK2-cre and Scnn1a-cre), and GABAergic (inhibitory) interneurons (SST-cre, PV-cre and VIP-cre) in cortex. Within the hippocampus, excitatory neurons (CamK2-cre and Grik4-cre) and inhibitory interneurons (SST-cre) were targeted. Quality control and mapping to the mouse reference genome (mm10) were performed using fastp and STAR with the same option as the other dataset. The mapped reads were analyzed by Shiba with the default settings to calculate PSI for each sample.

For the identification of alternative splicing events under the perturbation of neuronal RNA binding proteins (RBPs), we used RNA-seq of Emx1-cre *Ptbp2* knockout (KO) mouse cortex (postnatal day 1) (GSE84803) (31), Nestin-cre *Mbnl1* and *Mbnl2* double KO mouse adult cortex (SRP142522) (32), and *Mbnl2* KO mouse adult hippocampus (2–3 months old) (SRP013558) (33). Quality control and mapping to

the mouse reference genome (mm10) were performed using fastp and STAR with the same option as the other dataset. The mapped reads were analyzed by Shiba with the default settings to identify DSEs between the control and the RBP knockout samples.

Single-cell RNA-seq data

We downloaded the single-cell RNA-seq (scRNA-seq) data of the mouse midbrain dopaminergic neurons (embryonic day 13.5, 15.5, 18.5, postnatal day 1, 7, and 90) from GEO (GSE116138) (34). The raw sequencing reads were aligned to the mouse reference genome (mm10) using STAR in STARsolo mode (21) to obtain the gene expression matrix (GEM) across genes and individual cells and the splice junction matrix (SJM) across junctions and cells. The GEM was processed by Scanpy (version 1.9.5) (35). Possible outlier cells expressing non-neuronal gene markers were removed, and 987 cells were kept for downstream analysis. The table of cell barcodes was used for subsequent scShiba analysis with the SJM to identify DSEs between time points.

Single-nucleus RNA-seq data

We downloaded the single-nucleus RNA-seq (snRNA-seq) data of the mouse primary visual cortex (postnatal day 28) from GEO (GSE190940) (36). The raw sequencing reads were aligned to the mouse reference genome (mm10) using STAR with the STARsolo mode (21) to obtain the GEM across genes and single nuclei and the SJM across junctions and nuclei. GEM was processed by Scanpy (version 1.9.5) (35). We used Scrublet (37) for cell doublets detection and Harmony (38) for batch-effect correction. Cell types were manually assigned based on the expression level of marker genes. The table of cell barcodes and assigned cell types was used for the subsequent scShiba analysis with the SJM to identify DSEs between cell types of interest.

Performance evaluation

RNA-seq reads were analyzed by Shiba (version 0.1), rMATS (version 4.1.2) (4, 5), SUPPA2 (version 2.3) (6), and Whippet (version 1.6.1) (7) to predict DSEs. Shiba, rMATS, and SUPPA2 were executed in Singularity containers (naotokubota/shiba:0.1, xinglab/rmats:v4.1.2-no-entrypoint, and naotokubota/suppa:2.3 in Docker Hub). We ran tools supporting multi-core usage on 16 CPUs, and others on 1 CPU.

Shiba:

For Shiba, mapped reads and the Ensembl 102 transcript annotation file were used with the default parameters ($q < 0.05$ and $|\Delta PSI| > 10$). Minimum read coverage for differential splicing analysis was set to 10.

rMATS:

For rMATS, mapped reads and the Ensembl 102 transcript annotation file were used with the default parameters ($P < 0.05$ and $|\Delta PSI| > 10$). The options "--novelSS --mil 0 --mel 10000" were set to enable detection of unannotated splicing sites.

SUPPA2:

Salmon (version 1.10.1) (39) was executed to estimate transcript abundance in a Singularity container (combinelab/salmon:1.10.1 in Docker Hub). SUPPA2 was executed with the default parameters ($P < 0.05$ and $|\Delta PSI| > 10$) using the estimated transcript abundance.

Whippet:

According to the authors' guidelines, incorporating bam files when creating an index enables the detection of unannotated splicing in Whippet. However, errors occurred during index creation in multiple samples, making analysis impossible. The authors acknowledged this issue as a significant bug (referenced here: https://github.com/timbitz/Whippet.jl/issues/79). Consequently, we opted for standard index creation and

restricted our analysis solely to annotated events in this study. We performed analyses using default parameters (Probability > 0.95 and |ΔPSI| > 10) with the FASTQ file as input. Minimum read coverage for differential splicing analysis was set to 10.

The execution time of each tool was measured using the Linux time command (/usr/bin/time). We ran each tool for RNA-seq samples from the cytoplasmic fraction of mESC with siRNA-mediated double KD of *Ptbp1* and *Ptbp2* (GSE159993). The control and KD groups have three replicates with 40-60 million reads each.

To assess how well each tool predicted DSEs from simulated data, we used the 2,000 (differentially spliced genes) and 1,000 (unaltered) gene sets as the ground truth. We measured performance using accuracy, precision, recall, F1, false positive rate (FPR), false discovery rate (FDR), false negative rate (FNR), and Matthews correlation coefficient (MCC). To evaluate performance in detecting unannotated splicing events, we customized the reference gene annotation file by removing records of inclusion isoforms and used it for each computational tool so that alternative splicing events were treated as unannotated events in each analysis. The receiver operating characteristic (ROC) and precision-recall (PR) curves were created by changing the statistical metric (*P*-value for Shiba, rMATS, and SUPPA2; Probability for Whippet) from 0 to 1 at increments of 0.01. It should be noted that other criteria (such as ΔPSI or OR) were not considered, so the classification process of differentially spliced exons as usually performed by each program may be strictly different.

Comparison of multiple genomic sequences

We obtained multiple species' genomic sequences from the UCSC Genome Browser (40) to confirm the evolutionary conservation of the *THRAP3* cryptic exon. We used the following species' genomic sequences (assembly): Human (GRCh38/hg38), Chimp (CSAC 2.1.4/panTro4), Gorilla (gorGor3.1/gorGor3), Rhesus (BGI CR_1.0/rheMac3), Mouse (GRCm38/mm10), Rat (RGSC 6.0/rn6), Rabbit (Broad/oryCun2), Pig (SGSC Sscrofa10.2/susScr3), Cow (Bos_taurus_UMD_3.1.1/bosTau8), Cat (ICGSC Felis_catus

8.0/felCat8), Dog (Broad CanFam3.1/canFam3), Elephant (Broad/loxAfr3), Armadillo (Baylor/dasNov3), Chicken (ICGSC Gallus_gallus-4.0/galGal4).

Data visualization

We used the UCSC Genome Browser to visualize RNA-seq read coverage in regions of interest. All graphs were generated using the pandas (version 1.5.0), matplotlib (version 3.6.1), and seaborn (version 0.12.0) packages in the Python environment (version 3.10.6) in a Singularity container (jupyter/scipy-notebook:notebook-6.4.12 in Docker Hub).

Data Availability

The raw real RNA-seq data used in this study are available at GEO or SRA with the following IDs: GSE159993, GSE57433, GSE133291, GSE84803, SRP142522, and SRP013558. The sc/snRNA-seq data used in this study is available at GSE116138 and GSE190940.

Code Availability

The code of Shiba and scShiba is available at https://github.com/NaotoKubota/Shiba and https://figshare.com/articles/software/Shiba/25884322. The docker container is available at https://hub.docker.com/r/naotokubota/shiba.

Results

The Shiba method for alternative splicing analysis

To enable robust and reproducible detection of alternative splicing events using RNA-seq data, we constructed the four-step pipeline illustrated in Figure 1A-B. In the first step, using aligned RNA-seq data and reference gene annotation in GTF format, reference-guided transcript assembly is performed by StringTie2 (16). This step identifies novel transcripts not yet described in the reference gene annotation file,

enabling high-confidence detection of annotated and unannotated splicing events. Therefore, by utilizing exon body read coverage, Shiba is more sensitive in detecting unannotated events compared to existing methods that rely solely on junction read information.

In the second step, based on the information of transcript structures built in the first step, all alternative splicing events are listed and divided into five classic categories: skipped exon (SE), alternative five prime splice site (FIVE), alternative three prime splice site (THREE), mutually exclusive exons (MXE), and retained intron (RI). Each event is labeled as "annotated" or "unannotated". This label can be useful for the functional characterization of cryptic splicing in disease conditions.

In the third step, the counting of reads mapped to each exon-exon and exon-intron junction is performed using RegTools (17) and featureCounts (18). For exon-exon junctions, users can set the anchor length of reads and the maximum and minimum length of the intron to be analyzed. In the final step of the pipeline, the Percent Spliced In (PSI) is calculated and statistical analysis is performed to identify differentially spliced events (DSEs) between the two groups.

To allow the method to be agnostic to sample sizes, including n=1 for unique or hard-to-replicate biological samples, or individual cell clusters from scRNA-seq, or a single replicate of RNA-seq data (e.g., those published during the early age of RNA-seq), Fisher's exact test is performed for two-by-two tables of inclusive and exclusive junction read counts between the reference and alternative groups (Supplementary Figure 1). The odds ratio (OR) and P-value are then calculated for each alternative splicing event. Junction reads of replicated samples for each group are pooled to gain read coverage.

These values are combined with the difference in PSI between the two groups (ΔPSI) to define differential splicing events. In the default setting, events meeting all the following criteria will be classified as DSEs: (i) OR is greater than 3/2 or smaller than 2/3, (ii) FDR-adjusted P (q) is less than 0.05, and (iii) |ΔPSI| is greater than 10. In the case of SE and RI events, two-by-two tables are created separately for the upstream and downstream inclusive exon-exon/exon-intron junctions, resulting in two ORs and two P-values (Supplementary Figure 1A). In this scenario, an event is claimed as a DSE only if both ORs exceed the threshold and the larger q-value is below 0.05. MXE events have four inclusive exon-exon junctions

connecting each of the two flanking constitutive exons with each of the mutually exclusive exons. Four two-by-two tables are created for the calculation of four ORs and *P*-values (Supplementary Figure 1B). In this scenario, an event is detected as a DSE only if all four ORs exceed the threshold and the largest *q*-value is below 0.05. For the FIVE and THREE events, a two-by-two table is created for the two exon-exon junctions corresponding to each alternative splice site to calculate the OR and *P*-value to claim a DSE (Supplementary Figure 1C).

Performance evaluation

To assess Shiba, we performed comprehensive comparisons with other highly cited computational tools, including rMATS (4, 5), SUPPA2 (6), and Whippet (7), using both simulated and real RNA-seq datasets. We applied the computational tools to analyze simulated RNA-seq data. Specifically, we simulated RNA-seq reads to represent 2,000 alternative splicing events showing differential splicing and 1,000 alternative splicing events showing unaltered splicing between two groups in the mouse genome (Supplementary Figure S2). We set three different read coverages (x30, x20, and x10) and analyzed each to evaluate the differences in DS detection power as a function of read coverage. In addition to annotated splicing events, we included unannotated splicing events in the simulation by removing records of inclusion isoforms from the reference gene annotation file. Furthermore, we set variable biological replicates (3 vs. 3, 2 vs. 2, and 1 vs. 1) to examine each method's sensitivity to replicate numbers. Because rMATS and SUPPA do not support DS detection using one replicate data, they are excluded from the 1 vs. 1 comparison.

First, we evaluated the performance across methods of detecting DS for annotated events listed in the gene annotation file, using metrics of accuracy, precision, recall, F1, false positive rate (FPR), false discovery rate (FDR), false negative rate (FNR), and Matthews correlation coefficient (MCC). As shown in Figure 2A, Shiba and SUPPA2 demonstrate high accuracy and recall, as well as a low FNR, compared to rMATS and Whippet across all read coverages. Additionally, Shiba and SUPPA2 achieve higher F1 and MCC statistics than rMATS and Whippet. While SUPPA cannot handle data with a single replicate, Shiba

lacks this limitation; its performance persists across different coverages and biological replicates (Supplementary Figure S3A-B, S4, and S5).

Second, we assessed their performance in detecting unannotated events missing in the gene annotation file (Figure 2B). For the unannotated exons, Shiba showed the highest score in accuracy, recall, F1, and MCC, followed by rMATS across different coverages and biological replicates (Figure 2B, Supplementary Figure S3C-D, S6, and S7). SUPPA2 and Whippet showed high FNR as expected since these tools rely only on the reference gene annotation file. Notably, Shiba's FNR scores were around 0.2 to 0.3 at 30x and 20x coverage, making it the only software with FNR scores below 0.5. Overall, our findings demonstrated that Shiba achieves higher performance in the detection of annotated and unannotated DSEs than rMATS, SUPPA2, and Whippet.

Evaluation of robustness from sampling biological replicates

A robust computational tool, when applied to detect DSEs, should show greater consistency when sampling subsets of biological replicates. To compare different tools, we assessed their performance using real dataset: RNA-seq samples from cytoplasmic, nucleoplasmic, and chromatin fraction of mouse embryonic stem cells (mESC) with siRNA-mediated double knockdown (KD) of *Ptbp1* and *Ptbp2* (27). This dataset consists of control and perturbed samples with three biological replicates, making it a representative dataset. *Ptbp1* and *Ptbp2* are RNA-binding proteins (RBPs) known to regulate alternative splicing (31, 41). The three biological replicates allow us to scramble the replicate combinations to assess reproducibility, employing the reproducibility ratio (RR) metric (14). In essence, RR quantifies the proportion of DSEs identified by algorithm A in dataset D that are consistently identified using a similar replicate dataset D'. Our RR analysis subsampled two out of three biological replicates for each condition (Control and dKD of *Ptbp1* and *Ptbp2*), resulting in nine instances of DS analysis.

Shiba exhibited the highest RR in detecting significant DSEs across all fractions (Figure 3A). This trend persisted when events were ranked by a statistical metric, highlighting Shiba's superior reproducibility for the top 1,000 events compared to other algorithms (Figure 3B). To further validate these findings, we

performed RR analysis by subsampling one out of three replicates, specifically focusing on assessing the reproducibility of Shiba and Whippet. In all comparisons, Shiba consistently outperformed Whippet regarding reproducibility (Figure 3C-D). In summary, this comprehensive RR analysis underscores Shiba's robustness in consistently detecting DSEs across various cellular fractions.

Shiba considers junction read imbalance to reduce false positives

For SE, MXE, and RI events, Shiba employs Fisher's exact test on both sides of junctions. This careful approach allows Shiba to accurately assess and account for junction read imbalance, minimizing false positives. Our analysis of the Yeom et al. dataset revealed that 23.37% of alternative skipped exons exhibit a >4-fold difference in read counts between upstream and downstream inclusion junctions (Figure 4A-B), demonstrating the prevalence of junction read imbalance in real RNA-seq data. We hypothesized that the higher number of false positives identified by other computational tools is due to inadequate consideration of junction read imbalance. To test this hypothesis, we compared Shiba's performance against rMATS, anticipating that the careful approach of Shiba in addressing junction read imbalance would set it apart in terms of accuracy and effectiveness in detecting SE, MXE, and RI events, ultimately reducing the likelihood of false positives in the obtained results.

We compared the DSEs detected by Shiba vs rMATS in the cytoplasmic fraction of mESC upon siRNA-mediated double KD of *Ptbp1* and *Ptbp2* (27). For SE, this comparison revealed 218 events detected by both tools, with Shiba identifying an additional 223 events and rMATS detecting 318 events exclusively (Figure 4C). Our initial analysis revealed that the rMATS-exclusive events exhibited a lower total count of supporting junction reads compared to those exclusively identified by Shiba; notably 45 rMATS events have less than 10 junction reads (Figure 4D). Remarkably, 25.70% of the rMATS-identified events contained no reads at on one side of the junctions, contrasting with events exclusively identified by Shiba where all instances had at least one read on both sides of the junctions (Figure 4E). The absence of junction reads in some of these rMATS-identified events is associated with erroneous junction site determination, illustrated by a DSE claimed by rMATS in *Mark4* gene (Figure 4F). In this case, the rMATS exon and

Shiba exon share the common upstream 3' splice site but differ in their downstream 5' splice site, where no junction reads support the 5' splice site of the rMATS exon. Genome browser tracks of the RNA-seq data clearly agree with the Shiba results. These kinds of "imbalanced" alternative skip exons were commonly observed in the rMATS output (Supplementary Figure S8). The identification of DSEs with imbalanced junction reads by rMATS is not limited to cassette exons but is also observed in all other types of alternative splicing events (Supplementary Figure S9A-C).

Next, we quantitatively evaluate junction read imbalance by computing the percentage of junction reads mapped on either side (Figure 4G). Our analysis unveiled a bimodal distribution in the metric for events exclusively identified by rMATS, indicating that, in the majority of these events, the number of junction reads on one side greatly differed from those on the other side. In contrast, the metric for events exclusively identified by Shiba exhibited a normal distribution with an average score of nearly 50%, indicating a balanced distribution of junction reads on both sides of the exons for these events (Figure 4H). This result suggests that rMATS tends to identify events with a notable imbalance in the mapping of junction reads between the two sides. The aggregation plots of read coverage also demonstrate that events exclusively identified by Shiba show an even distribution of read coverage on the meta exonic region, while rMATS-identified events exhibit uneven read distribution despite higher read coverage (Figure 4I, Supplementary Figure S9D). This underscores the importance of considering read coverage patterns in addition to coverage depth for accurate event identification and highlights Shiba's ability to provide more reliable quantification by separately considering the two junctions.

We benchmarked the computational tools for sensitivity in handling junction read imbalance using simulated RNA-seq datasets where randomly selected SE events have no reads at all on one side of the junctions (Supplementary Figure S10A). This simulation revealed that Shiba outperformed rMATS, SUPPA2, and Whippet in all metrics (Supplementary Figure S10B-D), showcasing its superior sensitivity in identifying events with pronounced junction read imbalances. Taken together, these results underscore the robustness and effectiveness of Shiba in addressing junction read imbalance for accurate and reliable identification of alternative splicing events.

Evaluation of runtime

We evaluated the runtime of each tool using real RNA-seq data from Yeom et al (27). We allocated 16 CPUs for Shiba, rMATS, STAR, and Salmon. SUPPA2 and Whippet do not support multi-core usage, so they were executed on a single CPU. Additionally, we evaluated the runtime of SnakeShiba, which is a Snakemake-based and faster version of the Shiba pipeline. We found Shiba, SnakeShiba, and rMATS, utilizing the mapped files as input, took 68, 51, and 43 minutes, respectively. Shiba implements DESeq2 to perform differential gene expression analysis that rMATS and SUPPA2 do not do (Figure 1A), resulting in a slightly longer runtime. SUPPA2, including the expression quantification step using Salmon, required 22 minutes. Whippet, which internally performs pseudo-mapping, took 182 minutes (Supplementary Figure S11). Overall, the time difference is negligible against the backdrop of time needed for data acquisition. In summary, Shiba is capable of evaluating transcriptome data for both DEG and DSE in a multifaceted and speedy manner.

Shiba analysis of one replicate RNA-seq data reveals novel alternative splicing events coupled with nonsense-mediated decay

A notable strength of Shiba is its proficiency in identifying both annotated and unannotated DSEs with high accuracy, even when utilizing RNA-seq data without biological/technical replicates. In the early days of RNA-seq, biological samples often lacked replicates. Even nowadays, while three biological replicates represent the experimental gold standard, they are inherently insufficient for regression statistics. Shiba's robust capability to extract high-confidence DSEs from single-replicate datasets underscores its utility in mining existing resources for novel biological insights.

We applied the Shiba method to RNA-seq data derived from HEK293 Flp-In T-Rex cells, featuring siRNA-mediated double KD of *XRN1* and *UPF1*, notable for its absence of replicates (29) (Figure 5A). XRN1, a cytoplasmic 5′–3′ exonuclease, plays a pivotal role in the swift removal of decay intermediates (42), and UPF1 serves as a key component in nonsense-mediated decay (NMD) (43). Therefore, this dataset

holds significant promise for the identification of alternative splicing events coupled with NMD (AS-NMD). Shiba identified 1,861 DSEs ($q < 0.05$, $|\Delta PSI| > 10$) (Figure 5B). Up-regulation of the exon inclusion under the NMD inhibition condition suggests that the exon-inclusion isoform is an NMD target. For example, we identified an up-regulated SE within the *THRAP3* gene, encoding thyroid hormone receptor-associated protein 3 or TRAP150, a spliceosomal component (44) (Figure 5C). The ΔPSI of the SE was 45.2 (PSI in *XRN1*&*UPF1* dKD = 47.7, PSI in control = 2.45, $q = 4.56 \times 10^{-88}$). This exon was previously unannotated and would not be identified by some common algorithms. We found it contains an evolutionarily conserved TAA codon across vertebrates (Figure 5D), suggesting that the poison exon is subject to evolution selection and functional regulation. Since this event has not been documented in the reference annotation file and there is only one replicate in the dataset, this evolutionarily conserved poison exon cannot be identified by other algorithms but Shiba. These results demonstrate the value of Shiba in unveiling biologically relevant alternative splicing events with robust statistical metrics.

Detection of alternative splicing events in single-cell/nucleus RNA-seq data by scShiba

Shiba application on one-replicate data presents an invaluable tool to derive differential splicing from sc/snRNA-seq data with robust statistical metrics. Toward this goal, we developed scShiba, an alternative pipeline of Shiba for sc/snRNA-seq data that extends the capabilities of Shiba to sc/snRNA-seq data, allowing for a comprehensive analysis of alternative splicing dynamics between heterogeneous cell clusters. Due to the limited read coverage for individual cells, a fundamental strategy employed by the Shiba method involves aggregating reads for samples of interest, e.g., the same cell type or a cell cluster. This pooling approach for DSE detection enhances our ability to derive novel insights from sc/snRNA-seq data. scShiba takes several output files of STARsolo (21), a text file of user-defined cell type annotations, and a gene annotation file (Figure 6A). The following scShiba steps are identical to those of Shiba: pooling junction reads for each cell type, detecting alternative splicing events, calculating PSI, and performing statistical tests.

To test its performance, we applied scShiba to scRNA-seq datasets of dopaminergic neurons to uncover splicing regulation during their differentiation. Specifically, we analyzed Smart-seq2 data of mouse Pitx3+ midbrain dopaminergic neurons at embryonic day 13.5 (E13.5), 15.5 (E15.5), 18.5 (E18.5), postnatal day 1 (P1), 7 (P7), and 90 (P90) (34), aiming to identify DSEs during neuronal maturation (Figure 6B). Our scShiba analysis revealed significant differential splicing across each time point, identifying a total of 17,056 DSEs ($q < 0.05$, $|\Delta PSI| > 10$) (Figure 6C). For example, among these events was a skipped exon in the *Dbn1* gene (*Dbn1* E11), encoding drebrin, a post-synaptic actin-binding protein (45) (Figure 6D). The E11+ isoform of *Dbn1* is predominant in the adult brain (45). scShiba shows the PSI values for this exon were below 15 during early neuronal maturation stages (E13.5 ~ P7), rose to around 80 at the later stage (P90) (Figure 6E), resulting in a statistically significant difference of PSIs in comparison of E13.5 vs. P7, E15.5 vs. P7, and P90 vs. others (Figure 6F).

To further understand the developmental control of *Dbn1* E11, we explored possible upstream RBP regulators. RNA binding protein PTBP2 has been shown to regulate alternative splicing during early neuronal maturation (46–48). In later stages (adult neurons), other RBPs like MBNL1 and MBNL2 play significant roles in alternative splicing regulation of synaptic proteins (32). In line with the previous reports, analysis of dopaminergic neuron scRNA-seq data revealed high expression of *Ptbp2* during early stages (Figure 6G), and increased expression of *Mbnl1* and *Mbnl2* in later stages (Figure 6H-I). We applied Shiba to RNA-seq samples from Emx1-cre *Ptbp2* knockout (KO) mouse cortex (postnatal day 1) (31), Nestin-cre *Mbnl1* and *Mbnl2* double KO mouse adult cortex (32), and *Mbnl2* KO mouse adult hippocampus (2–3 months old) (33) to investigate splicing events regulated by these RBPs. Overlapping DSEs regulated by PTBP2 and MBNL1/2, Our Shiba analysis revealed that the largest overlapping group is those up-regulated in *Ptbp2* KO and down-regulated in Mbnl1/2 dKO (Group 2 containing 44 DSEs in Supplementary Figure S12A). Among them, 22 were DSEs between E13.5 vs. P90 in our scShiba analysis (Supplementary Figure S12B), including *Dbn1* E11. E11 is up-regulated in *Ptbp2* KO (PSI in KO = 12.54, PSI in control = 2.10, $\Delta PSI = 10.44$, $q = 1.71 \times 10^{-151}$) and down-regulated in *Mbnl1* and *Mbnl2* double KO (PSI in KO = 50.27, PSI in control = 80.38, $\Delta PSI = 30.11$, $q = 1.97 \times 10^{-61}$) and *Mbnl2* KO (PSI in KO = 62.05, PSI in control =

89.70, ΔPSI = 27.65, $q = 2.95 \times 10^{-33}$) (Figure 6J). These findings suggest a potential regulatory mechanism for alternative splicing of synaptic protein *Dbn1* during neuronal maturation: PTBP2 repression during early development and MBNL1/2 promotion of inclusion during later stages (Figure 6K).

Next, we applied scShiba to 10x snRNA-seq data from the mouse primary visual cortex on postnatal day 28 (36) to investigate DSEs between distinct cell types (Supplementary Figure S13A). Despite the low sequencing depth for each cell and highly biased 3'end sequencing of this dataset, scShiba identified a total of 84 DSEs ($q < 0.05$, $|\Delta PSI| > 10$) (Supplementary Figure S13B). We identified a skipped exon in the *Ppp1r12a* gene (*Ppp1r12a* E14) that exhibited differential splicing between excitatory and inhibitory neurons (Supplementary Figure S13C). The splicing levels of this exon were significantly different in excitatory neurons (PSI = 35.8) vs in inhibitory neurons (PSI = 90.9), resulting in a ΔPSI of 55.2 ($q = 7.03 \times 10^{-16}$) (Supplementary Figure S13D-E). This observation aligns with scShiba analyses of prior bulk RNA-seq data of neuronal subtypes by Furlanis et al. (30), in which the exon is more often skipped in excitatory neurons than in inhibitory neurons (Supplementary Figure S13F).

To infer its underlying regulation, we examined MBNL expression previously suggested to regulate differential splicing between glutamatergic and GABAergic neurons (49). Analysis of 10x snRNA-seq data from the visual cortex unveiled higher expression levels of *Mbnl2* in excitatory neurons compared to inhibitory neurons, while *Mbnl1* showed no disparity between neuronal subtypes (Supplementary Figure S13G-H). Our Shiba analyses of RNA-seq samples from *Mbnl1/2* KO brains showed MBNL repression of Ppp1r12a E14. Its inclusion is up-regulated in *Mbnl1* and *Mbnl2* double KO (PSI in KO = 99.57, PSI in control = 61.24, ΔPSI = 38.33, $q = 3.31 \times 10^{-54}$) and *Mbnl2* single KO (PSI in KO = 92.75, PSI in control = 53.15, ΔPSI = 39.60, $q = 1.11 \times 10^{-4}$) (Supplementary Figure S13I). These findings suggest a potential regulatory mechanism for alternative splicing of *Ppp1r12a* E14 in two distinct neuronal subtypes and implicate MBNL2 in repressing its inclusion in excitatory neurons (Supplementary Figure S13J).

These results, taken together, underscore the utility of scShiba in detecting alternative splicing events using scRNA-seq data from both the Smart-seq and droplet-based (e.g. 10x) technologies. Shiba and scShiba extend the values of scRNA-seq data for robust splicing analyses and allow cross-platform

corroboration with bulk RNA-seq data to derive high-confidence splicing regulation, presenting a valuable tool for unraveling the complexity of alternative splicing in diverse biological systems.

Discussion

In this study, we developed and conducted an extensive evaluation of the Shiba method for detecting differential alternative splicing from bulk and single-cell/nucleus RNA-seq datasets. The comprehensive comparison with commonly used computational tools, including rMATS, SUPPA2, and Whippet, reveals Shiba's superior performance in multiple metrics across various scenarios and read coverages. Besides its high efficiency, Shiba's strength lies in its accurate detection, even with very limited biological replicates, of both annotated and unannotated differential splicing events; the latter are commonly overlooked by other popular software. The reliability of Shiba is particularly evident thanks to its meticulous consideration of junction read imbalance for SE, MXE, and RI events, reducing false positives and providing more high-confidence DSEs. These qualities position Shiba as a comprehensive, robust tool for researchers aiming to unravel the intricate landscape of alternative splicing events with high precision for follow-up functional analyses.

Shiba employs robust splicing analysis agnostic to sample sizes, alleviating a common challenge for experimenters to produce a larger number of biological replicates in adopting existing computational pipelines. This does not mean we advocate a low number of biological replicates. It only showcases the practical value of Shiba in common experimental settings, where many biological replicates are infeasible. Increasing the sample size only increases the statistical power employed by Shiba. While Shiba is not built to analyze complex splicing events championed by intron-centric tools like LeafCutter and MAJIQ, Shiba’s exon-centric outputs provide immediate interpretable results for functional implication and experimental follow-up.

Recent revolutionary advances in high-throughput sequencing technology dissect the intricacies of gene regulation dynamics at a cellular resolution. The integration of scShiba further enhances our ability to

uncover subtle variations in splicing events, opening new avenues for exploring the complex landscape of alternative splicing in diverse cellular contexts. This granularity is particularly advantageous in unraveling cell-specific variations and rare splicing events that might be obscured in bulk analyses. Given that cell clusters generated from scRNA-seq data lack biological replicates, reliance on methods like scShiba for splicing analyses becomes imperative. Leveraging the wealth of scRNA-seq data from international consortia such as the Human Cell Atlas (50) and Allen Brain Atlas (51), we expect to explore cell-type-specific alternative splicing events more rigorously with scShiba (49).

In conclusion, Shiba is a powerful and versatile computational method for alternative splicing analysis, offering reliable detection of both annotated and unannotated differential splicing events. Its efficient runtime, reproducibility, consideration of junction read imbalance, and applicability to single-replicate and single-cell/nucleus data make it a valuable tool for researchers investigating transcriptome dynamics in diverse biological contexts. Shiba can make a significant contribution to the field of alternative splicing analysis and enhance our understanding of transcriptomic regulation.

Acknowledgements

We thank all members of Sika Zheng's laboratory for their valuable discussions. Manuscript writing was partially supported by ChatGPT to improve readability.

Funding

This work was supported by National Institutes of Health [R01NS15276 to S.Z., R01GM137428 to L.C.]. N.K. was partially supported by the Uehara Memorial Foundation Postdoctoral Fellowship.

Figure legends

Figure 1: Overview of Shiba. (A) Visualization of Shiba's statistical analysis process for detecting differential splicing events. (B) Schematic representation of the Shiba pipeline illustrating the four key steps:

reference-guided transcript assembly, identification of alternative splicing events, read counting, and statistical analysis.

Figure 2: Performance evaluation on detecting differentially spliced events in simulated RNA-seq data. (A) Performance metrics, including accuracy, precision, recall, F1 score, false positive rate (FPR), false discovery rate (FDR), false negative rate (FNR), and Matthews correlation coefficient (MCC), for annotated splicing events detection by four computational tools (Shiba, rMATS, SUPPA2, and Whippet) in different read coverages (x30, x20, and x10). (B) The corresponding metrics for unannotated splicing events detection.

Figure 3: Performance evaluation on reproducibility in real RNA-seq data. (A) The reproducibility ratio (RR) metric when subsampling two out of three biological replicates in RNA-seq samples from cytoplasmic (Cyp), nucleoplasmic (Nup), and chromatin (Chr) fraction of mouse embryonic stem cells (mESC) with siRNA-mediated double knockdown of *Ptbp1* and *Ptbp2* (27). (B) RR plots when ranking events by a statistical metric in the three fractions with a 95% confidence interval. The X-axis represents the ranked number of events reported, while the Y-axis indicates the fraction of those events reproduced within the same number of top-ranking events when repeating the analysis using different two replicates of samples. (C) The RR metric when subsampling one out of three biological replicates. (D) RR plots when ranking events by a statistical metric in the three fractions when subsampling one out of three biological replicates.

Figure 4: Shiba considers junction read imbalance to reduce false positives. (A) Schematic representation of junction read ratio in a skipped exon (SE). (B) Number of SE events for each bin of junction read ratio in RNA-seq data of cytoplasmic fraction of mESCs (27). (C) Comparison of detected SE events in the cytoplasmic fraction of mESC with siRNA-mediated double knockdown of *Ptbp1* and *Ptbp2* (27) between Shiba and rMATS. Shiba identified 218 events common to both tools, 223 additional events, while rMATS detected 318 events exclusively. (D) Total count of supporting junction reads for events exclusively

identified by Shiba and rMATS. (E) Percentage of events with no reads on one side of junctions for Shiba and rMATS exclusively identified events. (F) Examination of a specific SE event in the *Mark4* gene exclusively identified by rMATS, showing potential mis-annotation with a discrepancy between read coverage and the detected alternative exons. The exons exclusively identified by rMATS and Shiba are shown in the orange and green colored boxes, respectively. (G) Schematic representation of assessment of junction read imbalance. (H) Distribution of the percentage of junction reads mapped on the left side for events exclusively identified by Shiba and rMATS. (I) Aggregation plots of read coverage on the meta-exon, demonstrating even distribution for Shiba-identified events and uneven distribution for rMATS-identified events.

Figure 5: Shiba identifies novel nonsense-mediated mRNA decay-associated unannotated exons using RNA-seq data without biological replicates. (A) Experimental design of RNA-seq for HEK293 Flp-In T-Rex cells with siRNA-mediated double knockdown of *XRN1* and *UPF1*, lacking replicates (29). (B) Detection of 1,861 differentially spliced events with statistical significance ($q < 0.05$, $|\Delta PSI| > 10$) under NMD inhibition conditions by Shiba. (C) UCSC Genome Browser view of RNA-seq read coverage for the control (gray) and perturbed sample (purple) in the *THRAP3* gene with the PhastCons conservation score. The region highlighted in blue is the up-regulated cryptic skipped exon that contains a conserved premature termination codon. (D) Multiple alignments of the genomic sequence around evolutionary conserved TAA codon (highlighted in red) in the cryptic exon across vertebrates.

Figure 6: scShiba identifies neuronal-stage-specific alternative splicing events using single-cell RNA-seq data. (A) Workflow of scShiba, an extension of Shiba designed for single-cell/nucleus (sc/sn) RNA-seq data. (B) Smart-seq2 scRNA-seq data from mouse Pitx3+ midbrain dopaminergic neurons (34), plotted by the first two principal components of the normalized gene expression matrix. (C) A heatmap of a number of differentially spliced events (DSEs) detected in scShiba. The number in each box represents up-regulated DSEs ($q < 0.05$, $\Delta PSI > 10$) in the alternative group compared to the reference group. (D) Genome browser

track of *Dbn1* gene locus. (E) PSI values of *Dbn1* E11 in the samples of different stages from the Smart-seq2 data. (F) ΔPSI values of *Dbn1* E11 in all combinations of stages. Magenta boxes represent statistical significance ($q < 0.05$, $|\Delta PSI| > 10$). (G-I) Expression levels of *Ptbp2* (G), *Mbnl1* (H), and *Mbnl2* (I) in the dopaminergic neurons. Color represents normalized gene expression levels in each cell. (J) PSI values of *Dbn1* E11 in control and knockout (KO) samples from Emx1-cre *Ptbp2* KO mouse cortex (postnatal day 1) (31), Nestin-cre *Mbnl1* and *Mbnl2* double KO mouse adult cortex (32), and *Mbnl2* KO mouse adult hippocampus (2–3 months old) (33). (K) A model of alternative splicing program of *Dbn1* E11 in early and late stages of neuronal maturation.

References

1. Vuong,C.K., Black,D.L. and Zheng,S. (2016) The neurogenetics of alternative splicing. *Nat Rev Neurosci*, **17**, 265–281.

2. Marasco,L.E. and Kornblihtt,A.R. (2023) The physiology of alternative splicing. *Nat. Rev. Mol. Cell Biol.*, **24**, 242–254.

3. Nikom,D. and Zheng,S. (2023) Alternative splicing in neurodegenerative disease and the promise of RNA therapies. *Nat. Rev. Neurosci.*, 10.1038/s41583-023-00717-6.

4. Shen,S., Park,J.W., Lu,Z., Lin,L., Henry,M.D., Wu,Y.N., Zhou,Q. and Xing,Y. (2014) rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc National Acad Sci*, **111**, E5593–E5601.

5. Wang,Y., Xie,Z., Kutschera,E., Adams,J.I., Kadash-Edmondson,K.E. and Xing,Y. (2024) rMATS-turbo: an efficient and flexible computational tool for alternative splicing analysis of large-scale RNA-seq data. *Nat. Protoc.*, 10.1038/s41596-023-00944-2.

6. Trincado,J.L., Entizne,J.C., Hysenaj,G., Singh,B., Skalic,M., Elliott,D.J. and Eyras,E. (2018) SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol*, **19**, 40.

7. Sterne-Weiler,T., Weatheritt,R.J., Best,A.J., Ha,K.C.H. and Blencowe,B.J. (2018) Efficient and Accurate Quantitative Profiling of Alternative Splicing Patterns of Any Complexity on a Laptop. *Mol Cell*, **72**, 187-200.e6.

8. Katz,Y., Wang,E.T., Airoldi,E.M. and Burge,C.B. (2010) Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods*, **7**, 1009.

9. Tapial,J., Ha,K.C.H., Sterne-Weiler,T., Gohr,A., Braunschweig,U., Hermoso-Pulido,A., Quesnel-Vallières,M., Permanyer,J., Sodaei,R., Marquez,Y., *et al.* (2017) An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res*, **27**, 1759–1768.

10. Wang,Q. and Rio,D.C. (2018) JUM is a computational method for comprehensive annotation-free analysis of alternative pre-mRNA splicing patterns. *Proc National Acad Sci*, **115**, E8181–E8190.

11. Yan,Q., Weyn-Vanhentenryck,S.M., Wu,J., Sloan,S.A., Zhang,Y., Chen,K., Wu,J.Q., Barres,B.A. and Zhang,C. (2015) Systematic discovery of regulated and conserved alternative exons in the mammalian brain reveals NMD modulating chromatin regulators. *Proc National Acad Sci*, **112**, 3445–3450.

12. Li,Y.I., Knowles,D.A., Humphrey,J., Barbeira,A.N., Dickinson,S.P., Im,H. and Pritchard,J.K. (2018) Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics*, **50**, 151–158.

13. Vaquero-Garcia,J., Barrera,A., Gazzara,M.R., González-Vallinas,J., Lahens,N.F., Hogenesch,J.B., Lynch,K.W. and Barash,Y. (2016) A new view of transcriptome complexity and regulation through the lens of local splicing variations. *Elife*, **5**, e11752.

14. Vaquero-Garcia,J., Aicher,J.K., Jewell,S., Gazzara,M.R., Radens,C.M., Jha,A., Norton,S.S., Lahens,N.F., Grant,G.R. and Barash,Y. (2023) RNA splicing analysis using heterogeneous and large RNA-seq datasets. *Nat Commun*, **14**, 1230.

15. Mehmood,A., Laiho,A., Venäläinen,M.S., McGlinchey,A.J., Wang,N. and Elo,L.L. (2019) Systematic evaluation of differential splicing tools for RNA-seq studies. *Brief Bioinform*, **21**, 2052–2065.

16. Kovaka,S., Zimin,A.V., Pertea,G.M., Razaghi,R., Salzberg,S.L. and Pertea,M. (2019) Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol*, **20**, 278.

17. Cotto,K.C., Feng,Y.-Y., Ramu,A., Richters,M., Freshour,S.L., Skidmore,Z.L., Xia,H., McMichael,J.F., Kunisaki,J., Campbell,K.M., *et al.* (2023) Integrated analysis of genomic and transcriptomic data for the discovery of splice-associated variants in cancer. *Nat Commun*, **14**, 1589.

18. Liao,Y., Smyth,G.K. and Shi,W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, **30**, 923–930.

19. Benjamini,Y. and Hochberg,Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J Royal Statistical Soc Ser B Methodol*, **57**, 289–300.

20. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

21. Kaminow,B., Yunusov,D. and Dobin,A. (2021) STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. *bioRxiv*, 10.1101/2021.05.05.442755.

22. Mölder,F., Jablonski,K.P., Letcher,B., Hall,M.B., Tomkins-Tinch,C.H., Sochat,V., Forster,J., Lee,S., Twardziok,S.O., Kanitz,A., *et al.* (2021) Sustainable data analysis with Snakemake. *F1000Research*, **10**, 33.

23. Frazee,A.C., Jaffe,A.E., Langmead,B. and Leek,J.T. (2015) Polyester: simulating RNA-seq datasets with differential transcript expression. *Bioinformatics*, **31**, 2778–2784.

24. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

25. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

26. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and Subgroup,1000 Genome Project Data Processing (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

27. Yeom,K.-H., Pan,Z., Lin,C.-H., Lim,H.Y., Xiao,W., Xing,Y. and Black,D.L. (2021) Tracking pre-mRNA maturation across subcellular compartments identifies developmental gene regulation through intron retention and nuclear anchoring. *Genome Res*, **31**, 1106–1119.

28. Chen,S., Zhou,Y., Chen,Y. and Gu,J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, **34**, i884–i890.

29. Lykke-Andersen,S., Chen,Y., Ardal,B.R., Lilje,B., Waage,J., Sandelin,A. and Jensen,T.H. (2014) Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Gene Dev*, **28**, 2498–2517.

30. Furlanis,E., Traunmüller,L., Fucile,G. and Scheiffele,P. (2019) Landscape of ribosome-engaged transcript isoforms reveals extensive neuronal cell class-specific alternative splicing programs. *Nat Neurosci*, **22**, 1709–1717.

31. Vuong,J.K., Lin,C.-H., Zhang,M., Chen,L., Black,D.L. and Zheng,S. (2016) PTBP1 and PTBP2 Serve Both Specific and Redundant Functions in Neuronal Pre-mRNA Splicing. *Cell Reports*, **17**, 2766–2775.

32. Weyn-Vanhentenryck,S.M., Feng,H., Ustianenko,D., Duffié,R., Yan,Q., Jacko,M., Martinez,J.C., Goodwin,M., Zhang,X., Hengst,U., *et al.* (2018) Precise temporal regulation of alternative splicing during neural development. *Nat Commun*, **9**, 2189.

33. Charizanis,K., Lee,K.-Y., Batra,R., Goodwin,M., Zhang,C., Yuan,Y., Shiue,L., Cline,M., Scotti,M.M., Xia,G., *et al.* (2012) Muscleblind-like 2-Mediated Alternative Splicing in the Developing Brain and Dysregulation in Myotonic Dystrophy. *Neuron*, **75**, 437–450.

34. Tiklová,K., Björklund,Å.K., Lahti,L., Fiorenzano,A., Nolbrant,S., Gillberg,L., Volakakis,N., Yokota,C., Hilscher,M.M., Hauling,T., *et al.* (2019) Single-cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. *Nat. Commun.*, **10**, 581.

35. Wolf,F.A., Angerer,P. and Theis,F.J. (2018) SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.*, **19**, 15.

36. Cheng,S., Butrus,S., Tan,L., Xu,R., Sagireddy,S., Trachtenberg,J.T., Shekhar,K. and Zipursky,S.L. (2022) Vision-dependent specification of cell types and function in the developing cortex. *Cell*, **185**, 311-327.e24.

37. Wolock,S.L., Lopez,R. and Klein,A.M. (2019) Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst.*, **8**, 281-291.e9.

38. Korsunsky,I., Millard,N., Fan,J., Slowikowski,K., Zhang,F., Wei,K., Baglaenko,Y., Brenner,M., Loh,P. and Raychaudhuri,S. (2019) Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*, **16**, 1289–1296.

39. Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*, **14**, 417–419.

40. Nassar,L.R., Barber,G.P., Benet-Pagès,A., Casper,J., Clawson,H., Diekhans,M., Fischer,C., Gonzalez,J.N., Hinrichs,A.S., Lee,B.T., *et al.* (2022) The UCSC Genome Browser database: 2023 update. *Nucleic Acids Res.*, **51**, D1188–D1195.

41. Keppetipola,N., Sharma,S., Li,Q. and Black,D.L. (2012) Neuronal regulation of pre-mRNA splicing by polypyrimidine tract binding proteins, PTBP1 and PTBP2. *Crit. Rev. Biochem. Mol. Biol.*, **47**, 360–378.

42. Nagarajan,V.K., Jones,C.I., Newbury,S.F. and Green,P.J. (2013) XRN 5′ →3′ exoribonucleases: Structure, mechanisms and functions. *Biochim. Biophys. Acta (BBA) - Gene Regul. Mech.*, **1829**, 590–603.

43. Kim,Y.K. and Maquat,L.E. (2019) UPFront and center in RNA decay: UPF1 in nonsense-mediated mRNA decay and beyond. *RNA*, **25**, 407–422.

44. Lee,K.-M., Hsu,I.-W. and Tarn,W.-Y. (2010) TRAP150 activates pre-mRNA splicing and promotes nuclear mRNA degradation. *Nucleic Acids Res.*, **38**, 3340–3350.

45. Koganezawa,N., Hanamura,K., Sekino,Y. and Shirao,T. (2017) The role of drebrin in dendritic spines. *Mol. Cell. Neurosci.*, **84**, 85–92.

46. Li,Q., Zheng,S., Han,A., Lin,C.-H., Stoilov,P., Fu,X.-D. and Black,D.L. (2014) The splicing regulator PTBP2 controls a program of embryonic splicing required for neuronal maturation. *eLife*, **3**, e01201.

47. Zhang,M., Ergin,V., Lin,L., Stork,C., Chen,L. and Zheng,S. (2019) Axonogenesis Is Coordinated by Neuron-Specific Alternative Splicing Programming and Splicing Regulator PTBP2. *Neuron*, **101**, 690-706.e10.

48. Zheng,S., Gray,E.E., Chawla,G., Porse,B.T., O'Dell,T.J. and Black,D.L. (2012) Psd-95 is post-transcriptionally repressed during early neural development by PTBP1 and PTBP2. *Nat Neurosci*, **15**, 381-S1.

49. Feng,H., Moakley,D.F., Chen,S., McKenzie,M.G., Menon,V. and Zhang,C. (2021) Complexity and graded regulation of neuronal cell-type–specific alternative splicing revealed by single-cell RNA sequencing. *Proc National Acad Sci*, **118**, e2013056118.

50. Regev,A., Teichmann,S.A., Lander,E.S., Amit,I., Benoist,C., Birney,E., Bodenmiller,B., Campbell,P., Carninci,P., Clatworthy,M., *et al.* (2017) The Human Cell Atlas. *eLife*, **6**, e27041.

51. Yao,Z., Velthoven,C.T.J. van, Kunst,M., Zhang,M., McMillen,D., Lee,C., Jung,W., Goldy,J., Abdelhak,A., Aitken,M., *et al.* (2023) A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, **624**, 317–332.

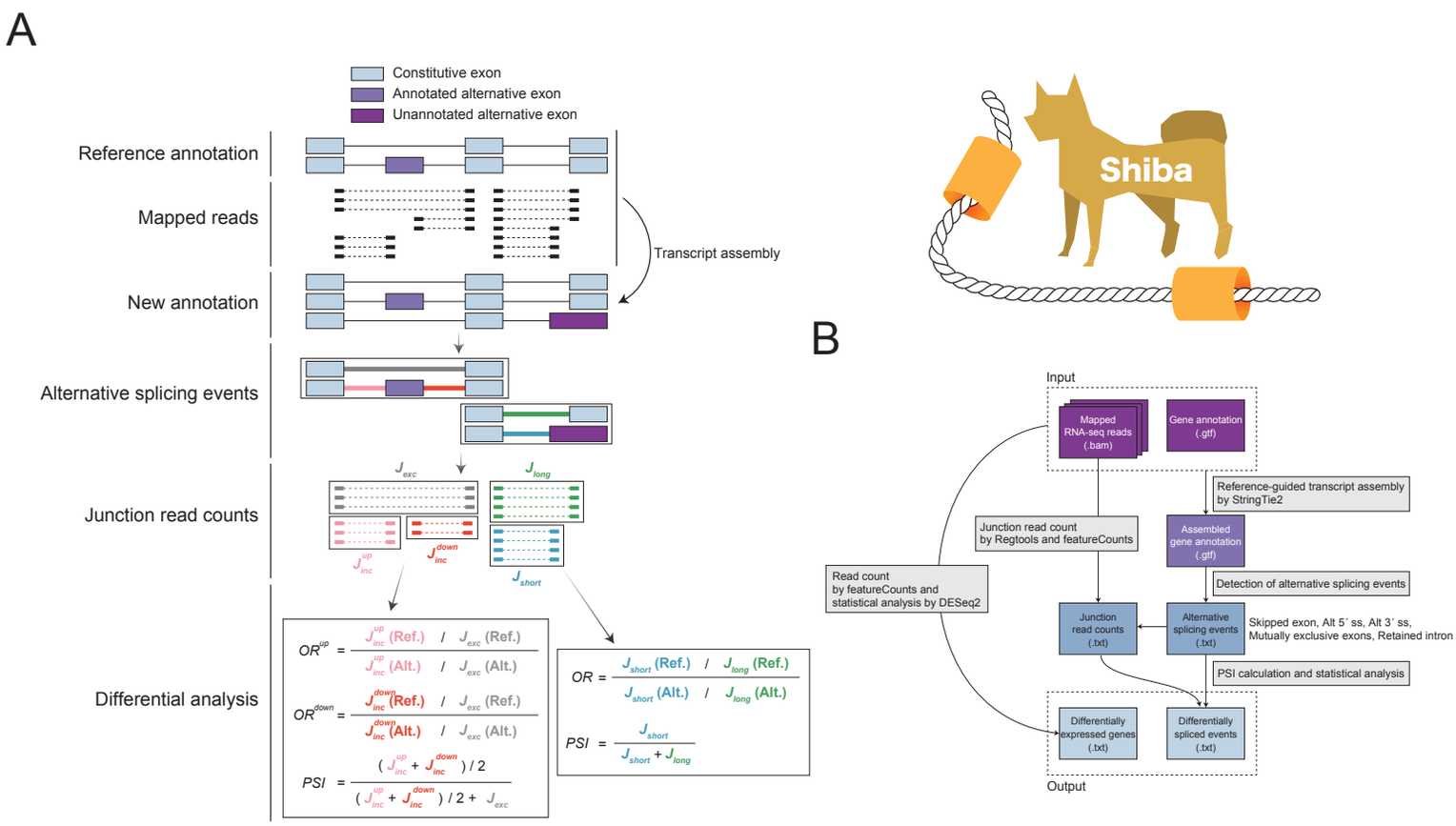

Figure 1: Overview of Shiba. (A) Visualization of Shiba's statistical analysis process for detecting differential splicing events. (B) Schematic representation of the Shiba pipeline illustrating the four key steps: reference-guided transcript assembly, identification of alternative splicing events, read counting, and statistical analysis.

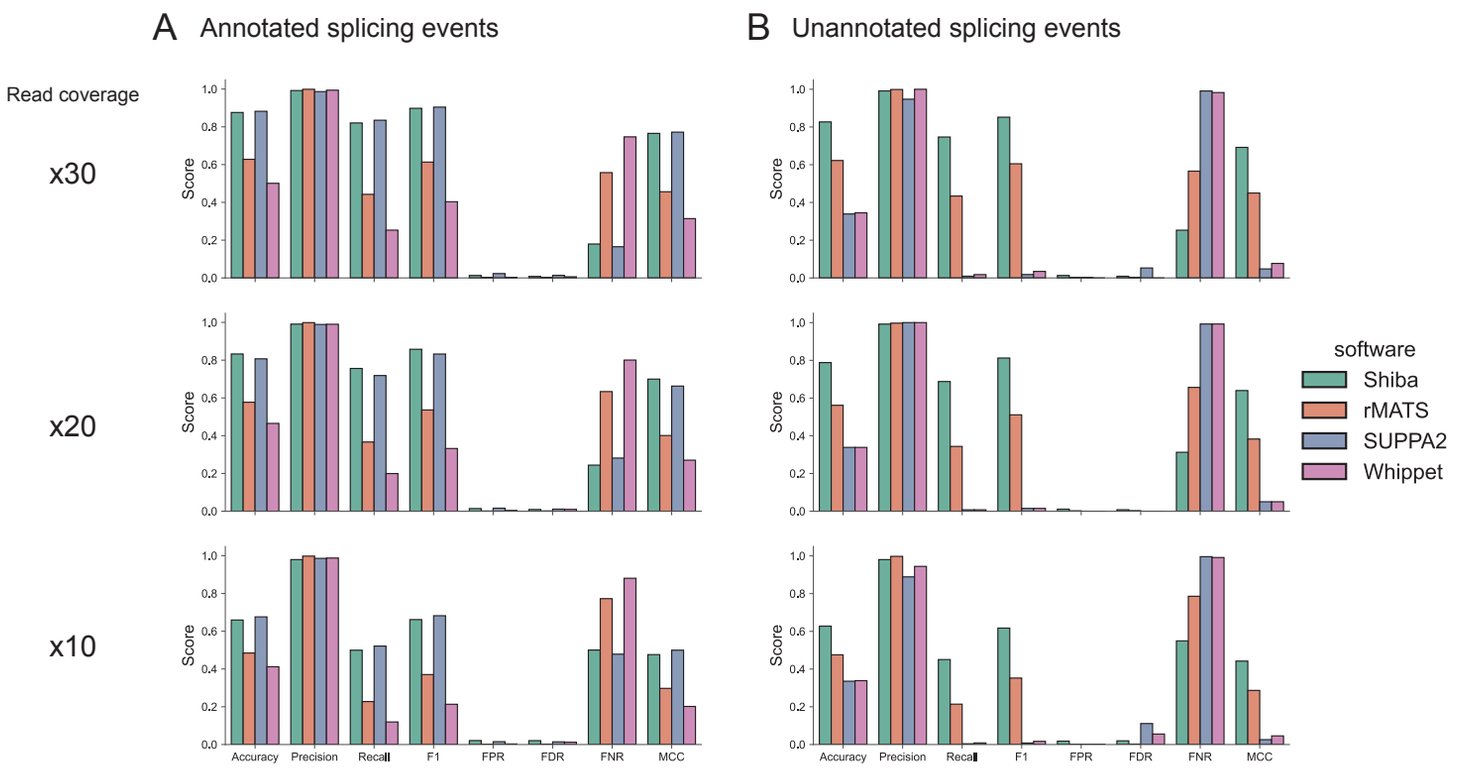

Figure 2: Performance evaluation on detecting differentially spliced events in simulated RNA-seq data. (A) Performance metrics, including accuracy, precision, recall, F1 score, false positive rate (FPR), false discovery rate (FDR), false negative rate (FNR), and Matthews correlation coefficient (MCC), for annotated splicing events detection by four computational tools (Shiba, rMATS, SUPPA2, and Whippet) in different read coverages (x30, x20, and x10). (B) The corresponding metrics for unannotated splicing events detection.

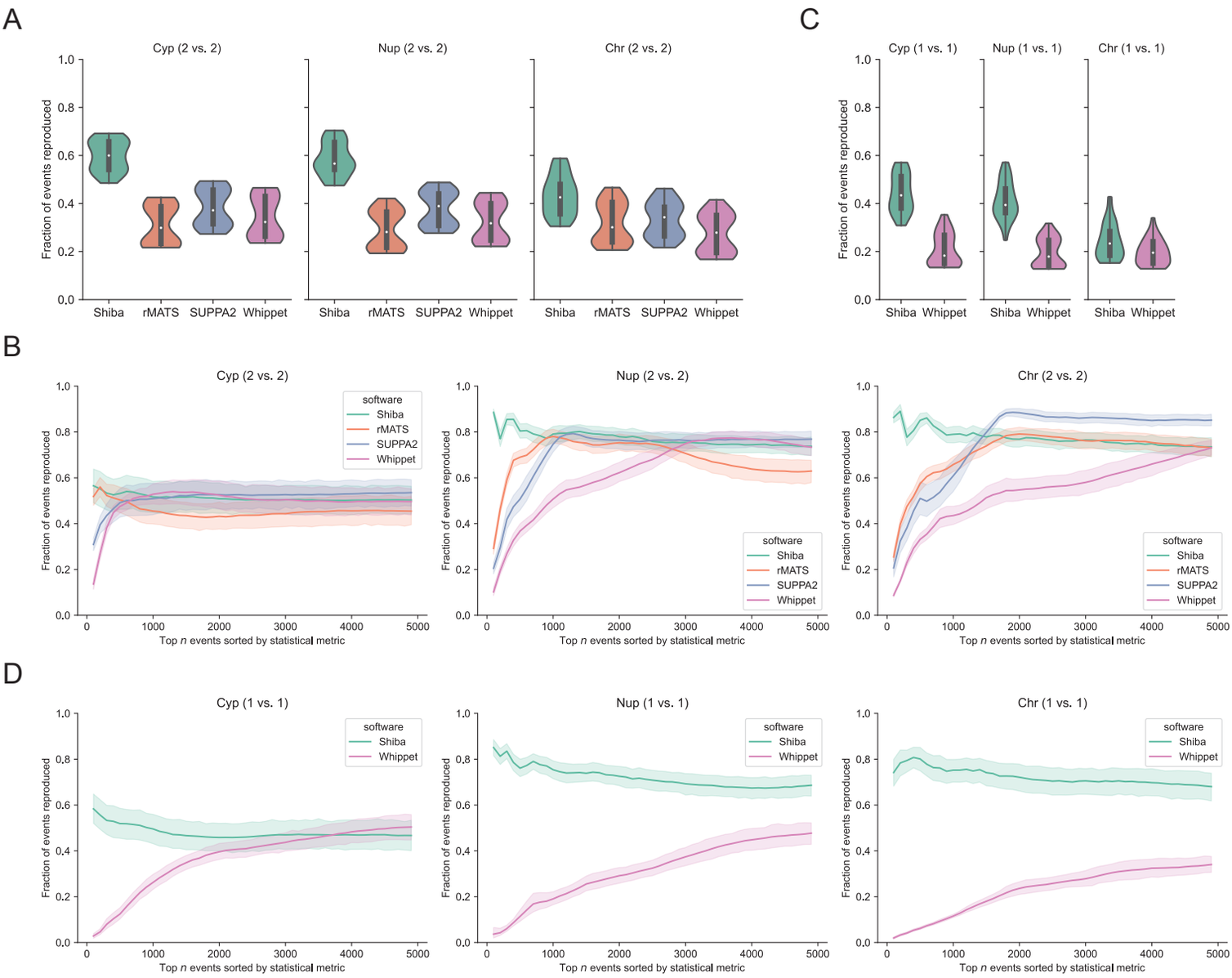

Figure 3: Performance evaluation on reproducibility in real RNA-seq data. (A) The reproducibility ratio (RR) metric when subsampling two out of three biological replicates in RNA-seq samples from cytoplasmic (Cyp), nucleoplasmic (Nup), and chromatin (Chr) fraction of mouse embryonic stem cells (mESC) with siRNA-mediated double knockdown of Ptbp1 and Ptbp2 (27). (B) RR plots when ranking events by a statistical metric in the three fractions with a 95% confidence interval. The X-axis represents the ranked number of events reported, while the Y-axis indicates the fraction of those events reproduced within the same number of top-ranking events when repeating the analysis using different two replicates of samples. (C) The RR metric when subsampling one out of three biological replicates. (D) RR plots when ranking events by a statistical metric in the three fractions when subsampling one out of three biological replicates.

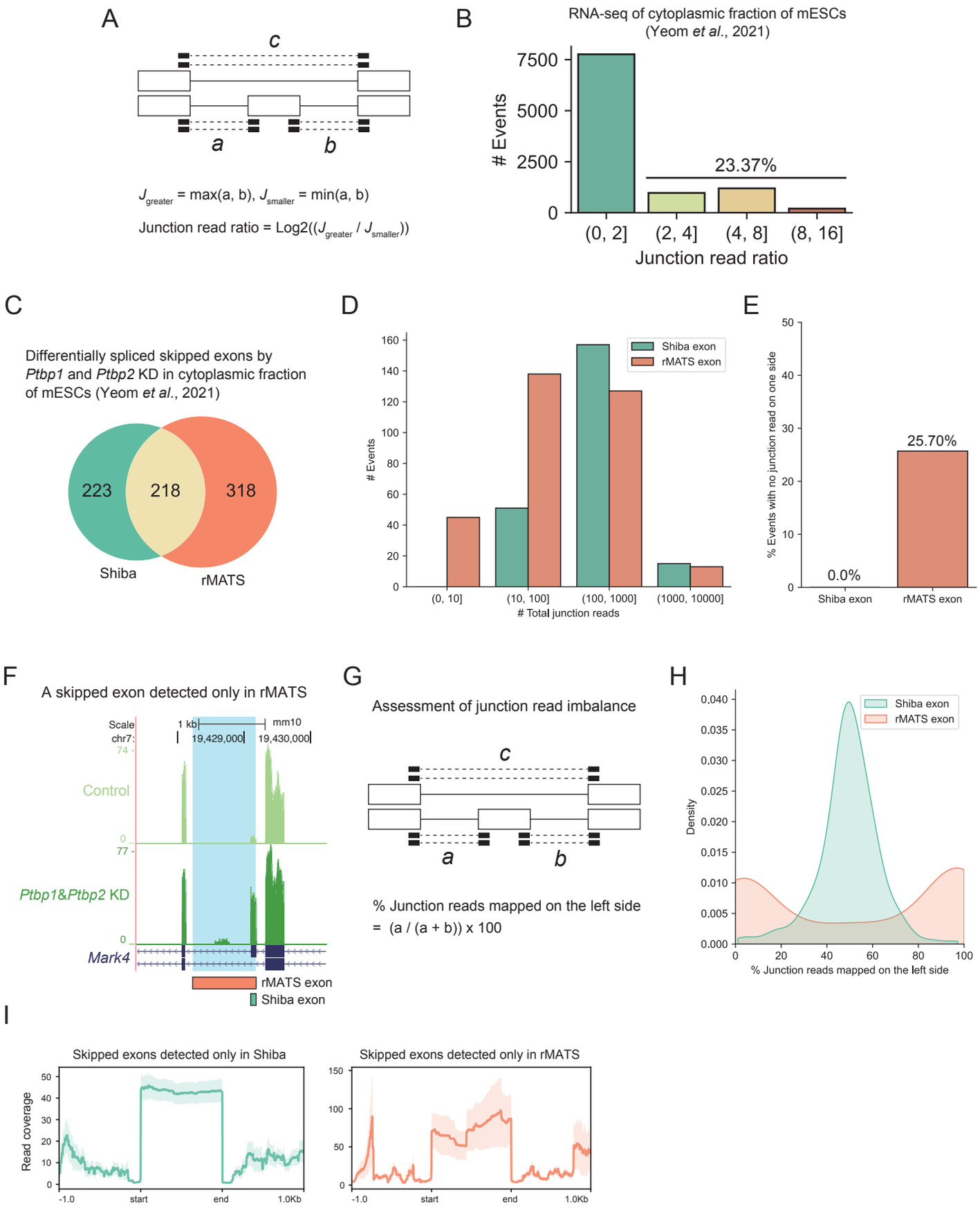

Figure 4, Kubota *et al.*

Figure 4: Shiba considers junction read imbalance to reduce false positives. (A) Schematic representation of junction read ratio in a skipped exon (SE). (B) Number of SE events for each bin of junction read ratio in RNA-seq data of cytoplasmic fraction of mESCs (27). (C) Comparison of detected SE events in the cytoplasmic fraction of mESC with siRNA-mediated double knockdown of Ptbp1 and Ptbp2 (27) between Shiba and rMATS. Shiba identified 218 events common to both tools, 223 additional events, while rMATS detected 318 events exclusively. (D) Total count of supporting junction reads for events exclusively identified by Shiba and rMATS. (E) Percentage of events with no reads on one side of junctions for Shiba and rMATS exclusively identified events. (F) Examination of a specific SE event in the Mark4 gene exclusively identified by rMATS, showing potential mis-annotation with a discrepancy between read coverage and the detected alternative exons. The exons exclusively identified by rMATS and Shiba are shown in the orange and green colored boxes, respectively. (G) Schematic representation of assessment of junction read imbalance. (H) Distribution of the percentage of junction reads mapped on the left side for events exclusively identified by Shiba and rMATS. (I) Aggregation plots of read coverage on the meta-exon, demonstrating even distribution for Shiba-identified events and uneven distribution for rMATS-identified events.

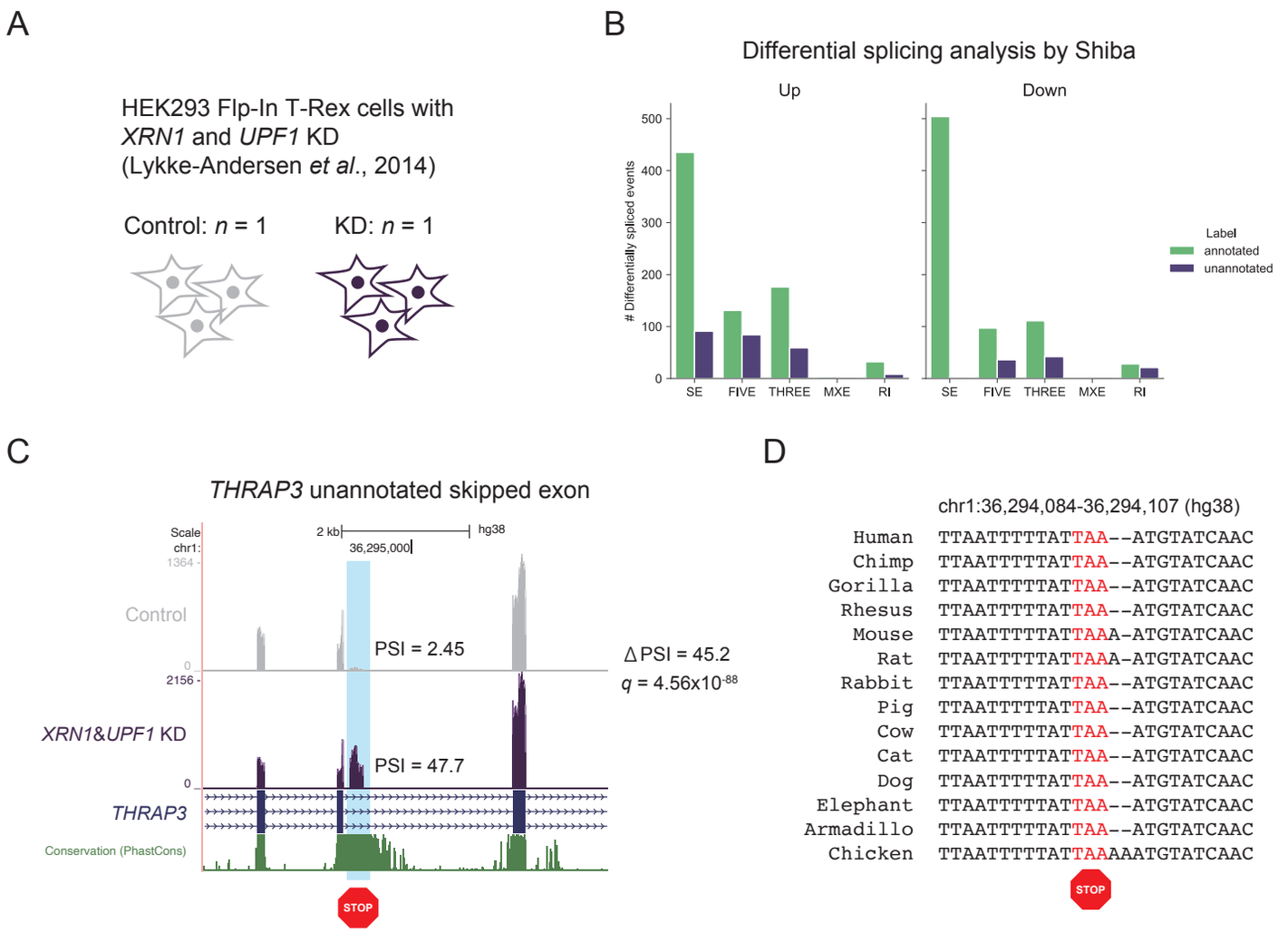

Figure 5: Shiba identifies novel nonsense-mediated mRNA decay-associated unannotated exons using RNA-seq data without biological replicates. (A) Experimental design of RNA-seq for HEK293 Flp-In T-Rex cells with siRNA-mediated double knockdown of XRN1 and UPF1, lacking replicates (29). (B) Detection of 1,861 differentially spliced events with statistical significance ($q < 0.05$, $|\Delta PSI| > 10$) under NMD inhibition conditions by Shiba. (C) UCSC Genome Browser view of RNA-seq read coverage for the control (gray) and perturbed sample (purple) in the THRAP3 gene with the PhastCons conservation score. The region highlighted in blue is the up-regulated cryptic skipped exon that contains a conserved premature termination codon. (D) Multiple alignments of the genomic sequence around evolutionary conserved TAA codon (highlighted in red) in the cryptic exon across vertebrates.

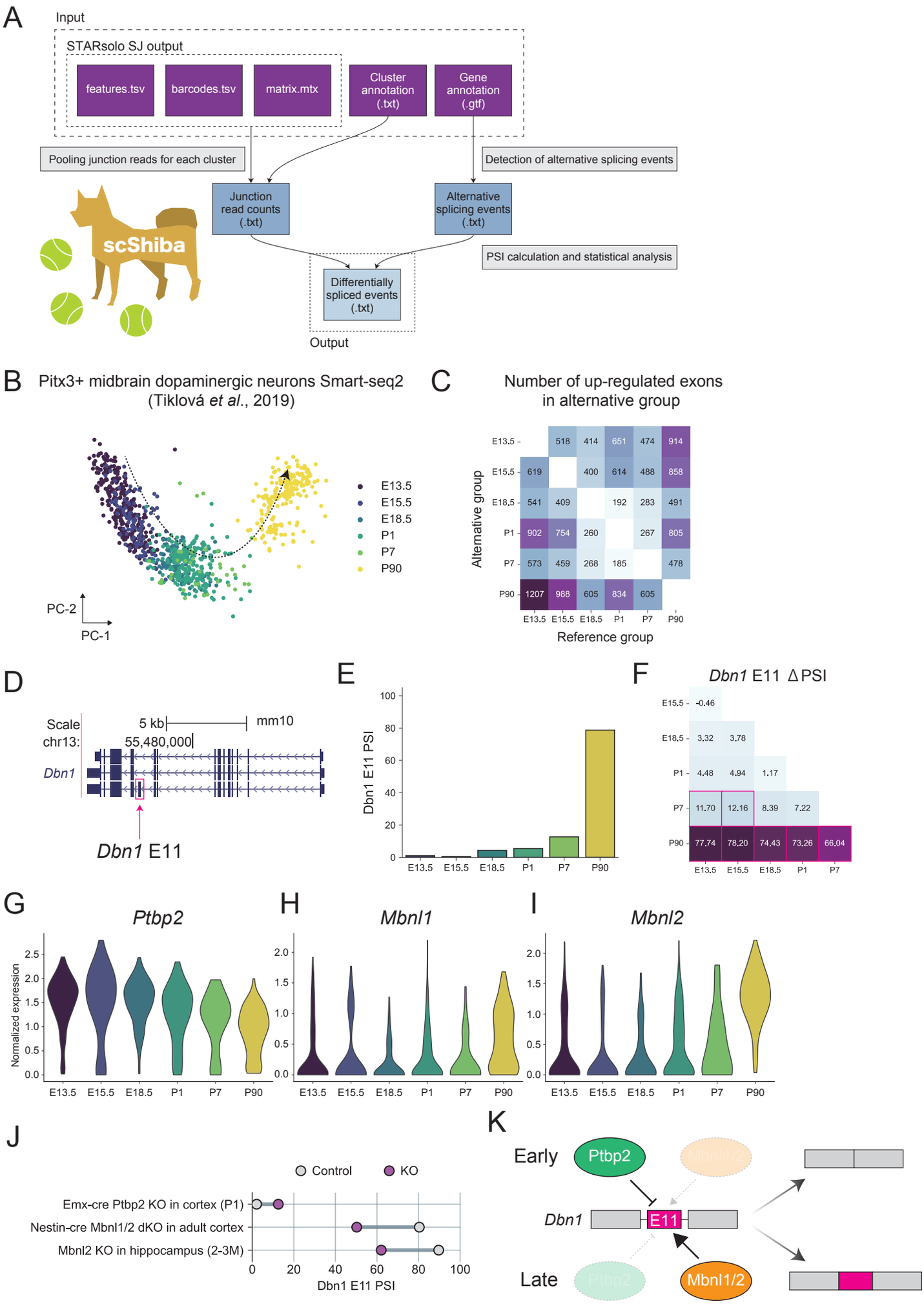

Figure 6, Kubota *et al.*

Figure 6: scShiba identifies neuronal-stage-specific alternative splicing events using single-cell RNA-seq data. (A) Workflow of scShiba, an extension of Shiba designed for single-cell/nucleus (sc/sn) RNA-seq data. (B) Smart-seq2 scRNA-seq data from mouse Pitx3+ midbrain dopaminergic neurons (34), plotted by the first two principal components of the normalized gene expression matrix. (C) A heatmap of a number of differentially spliced events (DSEs) detected in scShiba. The number in each box represents up-regulated DSEs ($q < 0.05$, Δ PSI > 10) in the alternative group compared to the reference group. (D) Genome browser track of Dbn1 gene locus. (E) PSI values of Dbn1 E11 in the samples of different stages from the Smart-seq2 data. (F) ΔPSI values of Dbn1 E11 in all combinations of stages. Magenta boxes represent statistical significance ($q < 0.05$, $|\Delta PSI| > 10$). (G-I) Expression levels of Ptbp2 (G), Mbnl1 (H), and Mbnl2 (I) in the dopaminergic neurons. Color represents normalized gene expression levels in each cell. (J) PSI values of Dbn1 E11 in control and knockout (KO) samples from Emx1-cre Ptbp2 KO mouse cortex (postnatal day 1) (31), Nestin-cre Mbnl1 and Mbnl2 double KO mouse adult cortex (32), and Mbnl2 KO mouse adult hippocampus (2–3 months old) (33). (K) A model of alternative splicing program of Dbn1 E11 in early and late stages of neuronal maturation.