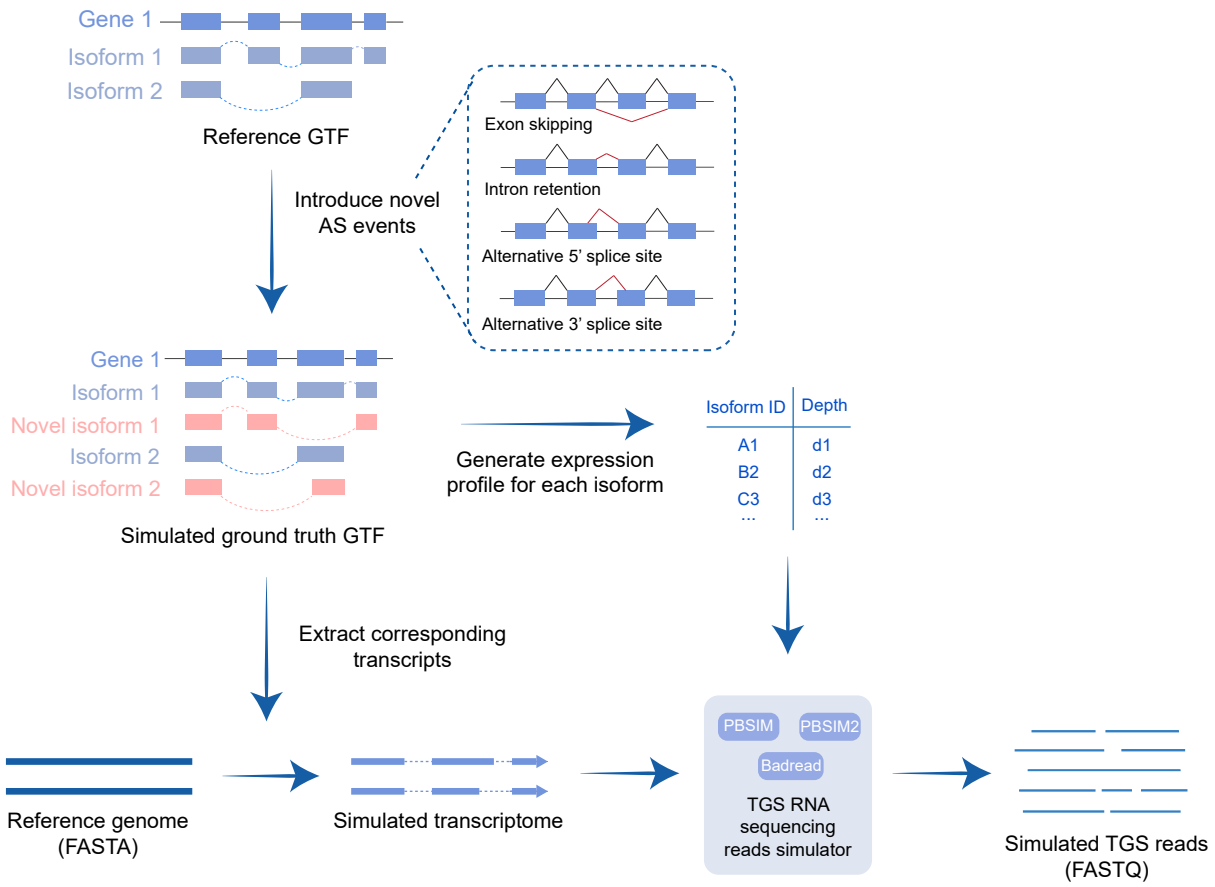
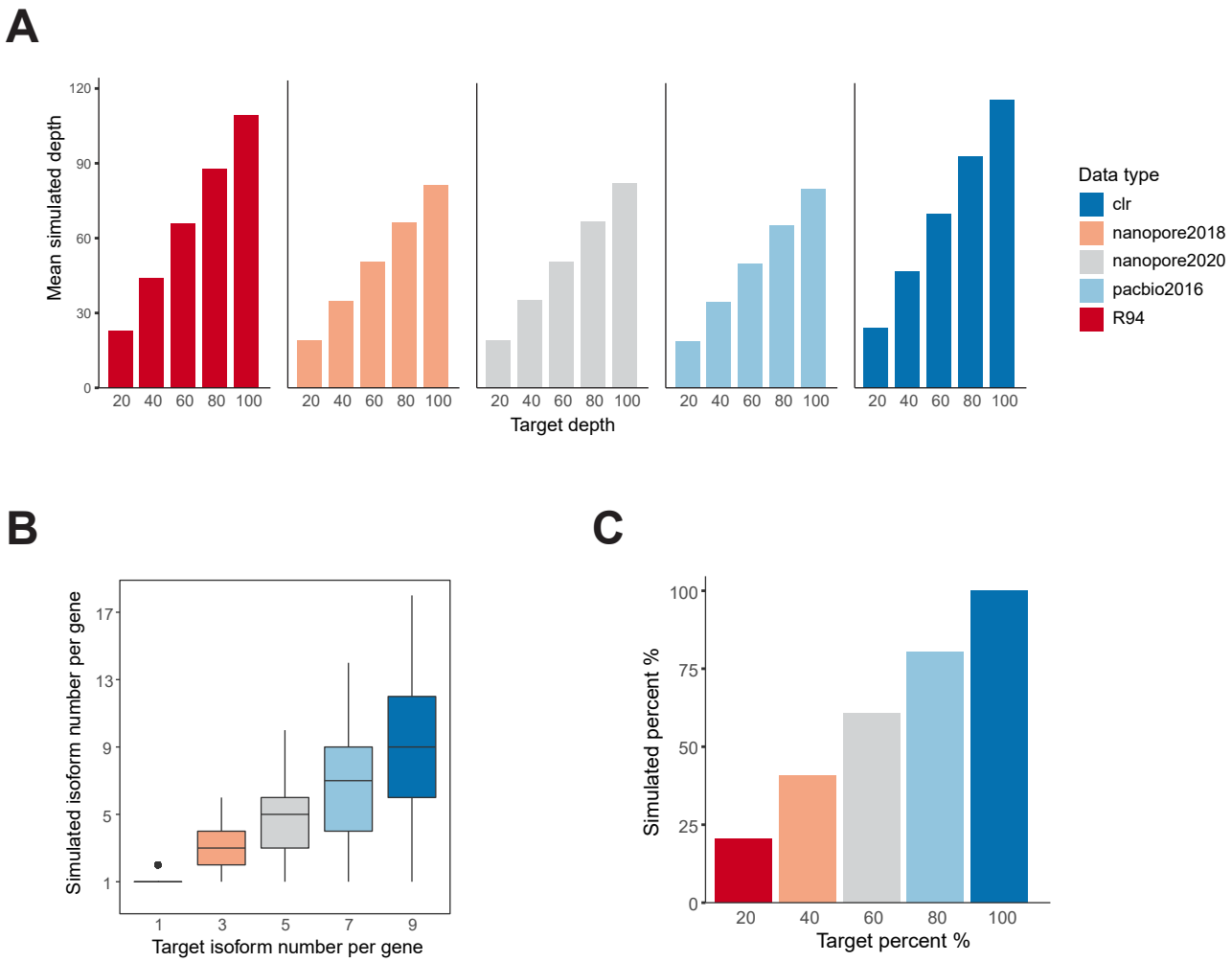
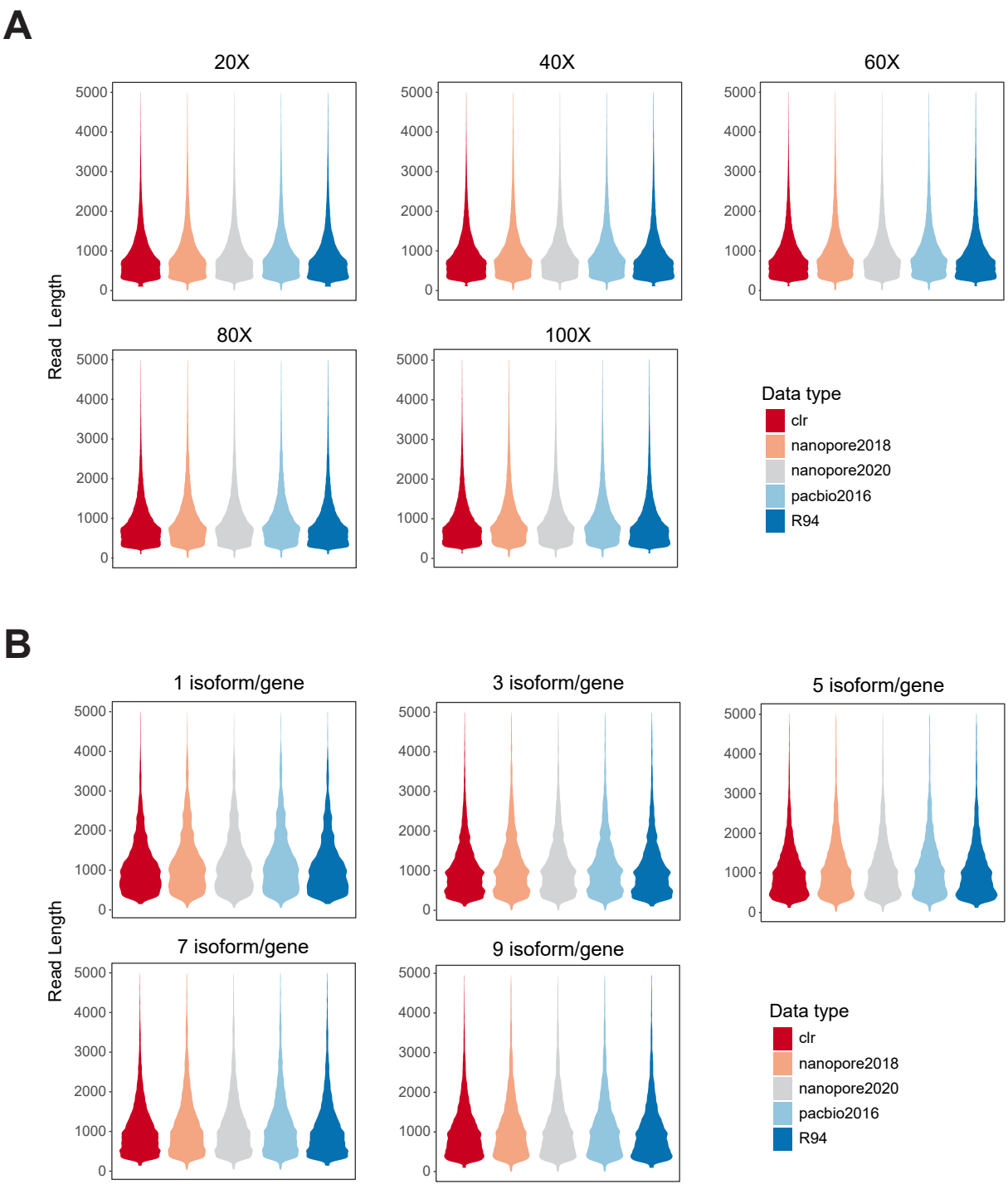
Supplementary Figures



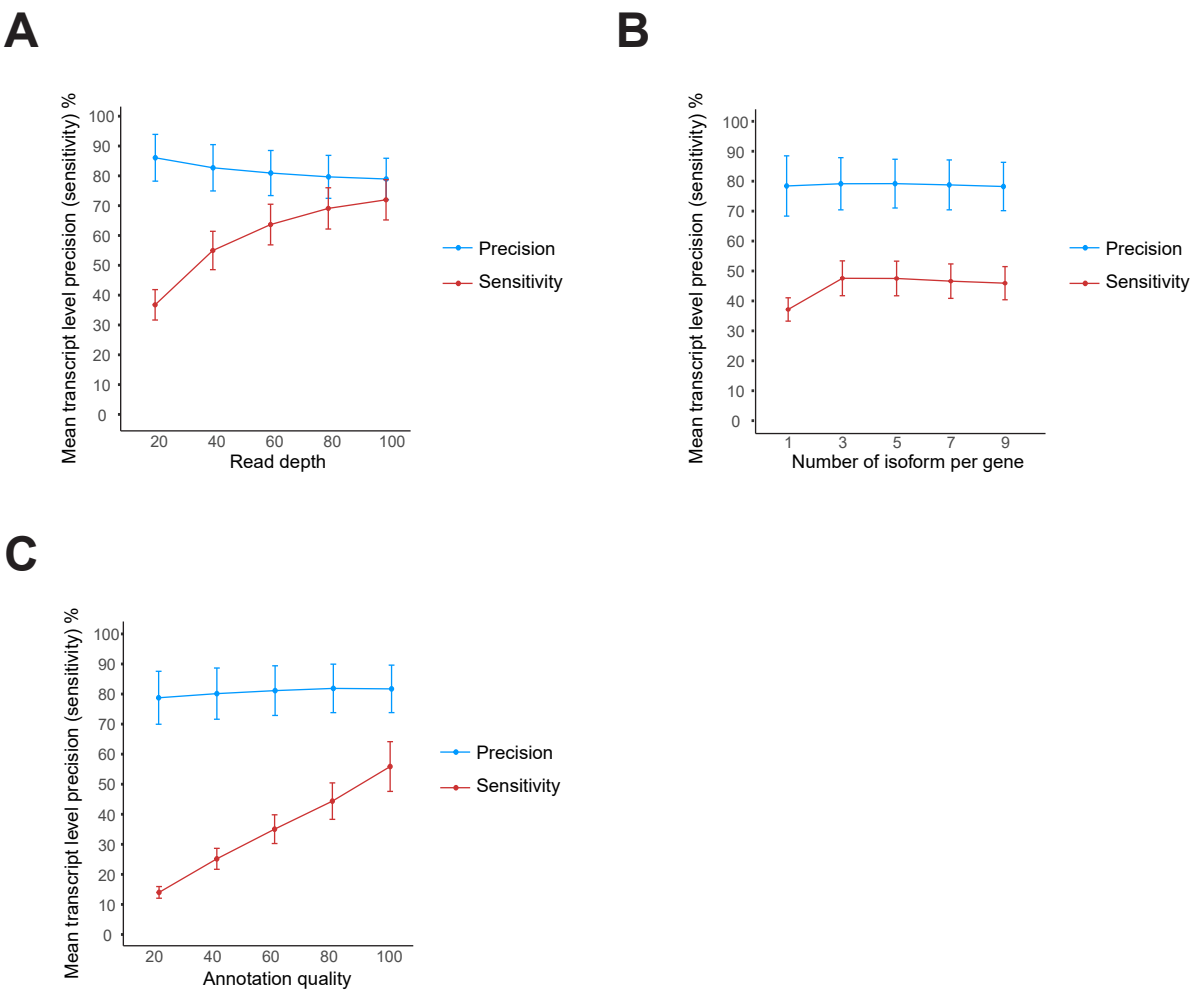
Supplementary figure 1. An overview of the workflow of YASIM.
First, four kinds of novel AS events are introduced into the original reference annotation GTF, and then the read depth is generated for each simulated isoform according to uniform distribution. The simulated transcriptome is also extracted from the given reference genome based on the simulated ground truth GTF. Next, both the simulated transcriptome and the generated expression profile are transferred to the previously published third-generation sequencing reads generator LLRG, which transforms the whole transcriptome into sequencing reads and simulates sequencing errors according to certain user-defined error model.



Supplementary figure 2. Statistics of the simulated data.
A. Mean read depths of each simulated dataset generated with different target expression profiles.
B. Number of isoforms per gene of each groung truth annotation generated with different sizes of AS events.
C. The percentage of completeness of the reference annotation provided for guided methods.



Supplementary figure 3. Read length distributions of the simulated datasets.
The distribution of read lengths of datasets simulated with different **A.** read depths and **B.** number of isoforms per gene. Those reads with length over 5000 bp were filtered out.



Supplementary figure 4. Results of FLAMES on simulated datasets with support read threshold equals 10. Mean transcript level precision and sensitivity obtained from simulated data with different **A.** read depths (n=25). **B.** number of isoforms per gene (n=25). **C.** annotation quality (n=5). All values are denoted as mean +/- SD.