Systematic evaluation of spliced alignment programs for RNA-seq data

Pär G Engström[1,13], Tamara Steijger[1], Botond Sipos[1], Gregory R Grant[2,3], André Kahles[4,5], The RGASP Consortium[6], Gunnar Rätsch[4,5], Nick Goldman[1], Tim J Hubbard[7], Jennifer Harrow[7], Roderic Guigó[8,9] & Paul Bertone[1,10–12]

High-throughput RNA sequencing is an increasingly accessible method for studying gene structure and activity on a genome-wide scale. A critical step in RNA-seq data analysis is the alignment of partial transcript reads to a reference genome sequence. To assess the performance of current mapping software, we invited developers of RNA-seq aligners to process four large human and mouse RNA-seq data sets. In total, we compared 26 mapping protocols based on 11 programs and pipelines and found major performance differences between methods on numerous benchmarks, including alignment yield, basewise accuracy, mismatch and gap placement, exon junction discovery and suitability of alignments for transcript reconstruction. We observed concordant results on real and simulated RNA-seq data, confirming the relevance of the metrics employed. Future developments in RNA-seq alignment methods would benefit from improved placement of multimapped reads, balanced utilization of existing gene annotation and a reduced false discovery rate for splice junctions.

Programs for aligning transcript reads to a reference genome address the challenging task of placing spliced reads across introns and correctly determining exon-intron boundaries. The advent of RNA-seq prompted the development of a new generation of spliced-alignment software, with several advances over earlier programs such as the BLAST-like alignment tool (BLAT)[1,2]. The tools GEM[3], GSTRUCT, MapSplice[4] and TopHat[5,6] implement a two-step approach in which initial read alignments are analyzed to discover exon junctions; these junctions are then used to guide final alignment. Several programs can also use existing gene annotation to inform spliced-read placement[5–9]. Most RNA-seq aligners can further increase accuracy by prioritizing alignments in which read pairs map in a consistent fashion[3,5–7,9,10]. To place reads that match multiple genomic sequences, GSTRUCT examines the density of independent reads at those loci. Many algorithms also consider base-call quality scores and use sophisticated indexing schemes to decrease runtime.

Here we assess the performance of 26 RNA-seq alignment protocols on real and simulated human and mouse transcriptomes. We adopted a competitive evaluation model applied in other areas of bioinformatics[11–14]. Developers were invited to run their software and submit results for evaluation as part of the RNA-seq Genome Annotation Assessment Project (RGASP). Programs included six spliced aligners GSNAP[7], MapSplice[4], PALMapper[8], ReadsMap, STAR[9] and TopHat[5,6]) and four alignment pipelines (GEM[3], PASS[15], GSTRUCT and BAGET). GSTRUCT is based on GSNAP, whereas BAGET uses a contiguous DNA aligner to map reads to the genome as well as to exon junction sequences derived from reference gene annotation. For comparison, the contiguous aligner SMALT was also tested. SMALT can map reads in a split manner, but it lacks several features of dedicated spliced aligners, such as precise determination of exon-intron boundaries. We demonstrate that choice of alignment software is critical for accurate interpretation of RNA-seq data, and we identify aspects of the spliced-alignment problem in need of further attention.

RESULTS

Alignment protocols were evaluated on Illumina 76-nucleotide (nt) paired-end RNA-seq data from the human leukemia cell line K562 (1.3×10^9 reads), mouse brain (1.1×10^8 reads) and two simulated human transcriptomes (8.0×10^7 reads each; **Supplementary Table 1**). Nine development teams contributed alignments for evaluation. We additionally included two versions of the widely used RNA-seq aligner TopHat[5,6]. Most development teams provided results from several alignment protocols, corresponding to different parameter choices and pipeline configurations (**Fig. 1** and **Supplementary Note**).

[1]European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge, UK. [2]Department of Genetics, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [3]Penn Center for Bioinformatics, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [4]Computational Biology Center, Sloan-Kettering Institute, New York, New York, USA. [5]Friedrich Miescher Laboratory of the Max Planck Society, Tübingen, Germany. [6]Full lists of members and affiliations appear at the end of the paper. [7]Wellcome Trust Sanger Institute, Cambridge, UK. [8]Centre for Genomic Regulation, Barcelona, Spain. [9]Universitat Pompeu Fabra, Barcelona, Spain. [10]Genome Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany. [11]Developmental Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany. [12]Wellcome Trust–Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Cambridge, UK. [13]Present address: Department of Biochemistry and Biophysics, Science for Life Laboratory, Stockholm University, Stockholm, Sweden. Correspondence should be addressed to P.B. (bertone@ebi.ac.uk).

RECEIVED 31 MARCH; ACCEPTED 10 SEPTEMBER; PUBLISHED ONLINE 3 NOVEMBER 2013; DOI:10.1038/NMETH.2722

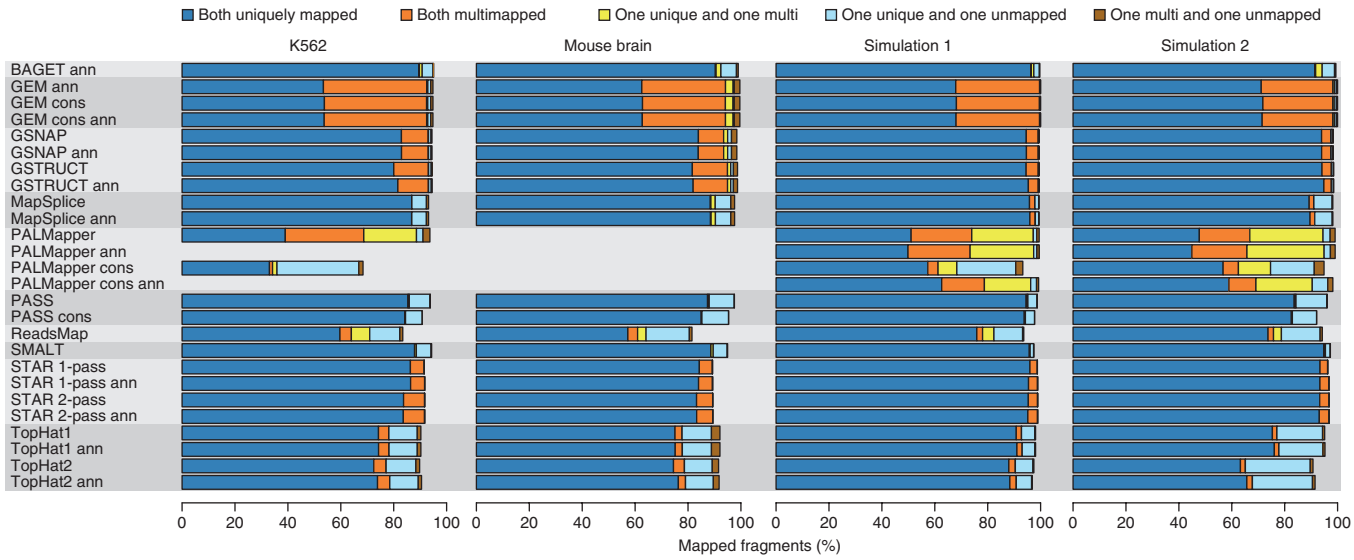


Figure 1 | Alignment yield. Shown is the percentage of sequenced or simulated read pairs (fragments) mapped by each protocol. Protocols are grouped by the underlying alignment program (gray shading). Protocol names contain the suffix "ann" if annotation was used. The suffix "cons" distinguishes more conservative protocols from others based on the same aligner. The K562 data set comprises six samples, and the metrics presented here were averaged over them.

Alignment yield

There were major differences among protocols in the alignment yield (68.4–95.1% of K562 read pairs; mean = 91.5%, s.d. = 5.4), extent to which both reads from a pair were mapped, and frequency of ambiguous mappings (reads with several reported alignments) (**Fig. 1** and **Supplementary Tables 2** and **3**). These trends were similar across data sets (**Fig. 1**). The fraction of pairs with only one read aligned was typically highest for TopHat, ReadsMap and PASS, whereas PALMapper output exhibited more complex discrepancies within read pairs. GEM results consistently included many ambiguous mappings (37% of sequenced reads per data set on average). Mapping ambiguities were also common with PALMapper, although these were reduced with the more conservative protocols that involve stringent filtering of alignments (**Fig. 1** and **Supplementary Fig. 1**). To avoid introducing bias at later evaluation stages due to differences in the number of alignments per read, we instructed developer teams to assign a preferred (primary) alignment for each read mapped in their program output. The following results are based on these primary alignments unless otherwise noted.

Mismatches and basewise accuracy

Compared to the other aligners, GSNAP, GSTRUCT, MapSplice, PASS, SMALT and STAR reported more primary alignments devoid of mismatches (**Fig. 2a**), partly because these methods can truncate read ends and thus output an incomplete alignment when they are unable to map an entire sequence (**Fig. 2b**). PASS and SMALT performed extensive truncation, suggesting that these programs often report alignments shorter than is optimal. MapSplice, PASS and TopHat displayed a low tolerance for mismatches (**Fig. 2a**). Consequently, a large proportion of reads with low base-call quality scores were not mapped by these methods (**Supplementary Fig. 2**). The mapping yield of TopHat was particularly low (mean yield of 84% on K562 data, compared to 90% for MapSplice; **Fig. 2a** and **Supplementary Tables 2** and **3**), likely owing to a lack of read truncation (**Fig. 2b**). Note that many aligners have options to increase mismatch tolerance beyond the settings used here, but this approach may negatively affect other performance aspects.

Polymorphisms and accumulated mutations distinguish the cancer cell line K562 from the human reference assembly, which itself is a consensus based on several individuals[16]. Conversely, mouse RNA samples were obtained from strain C57BL/6NJ, the genome of which is nearly identical to the mouse reference assembly[17]. Accordingly, high-quality reads from mouse were mapped at a greater rate and with fewer mismatches than those from K562 (**Supplementary Fig. 3**). Even so, differences among aligners in mismatch and truncation frequencies were consistent across data sets (**Fig. 2** and **Supplementary Fig. 4**). Mapping properties are thus largely dependent on software algorithms even when the genome and transcriptome are virtually identical.

Consistent with real RNA-seq data, GSNAP, GSTRUCT, MapSplice and STAR outperformed other methods for basewise accuracy on simulated data (**Supplementary Table 2**). As expected, error rates were substantially lower for uniquely mapped reads than for primary alignments of multimapped reads (**Supplementary Table 4**). Notably, despite the many ambiguous mappings reported by GEM and PALMapper, the primary alignments were usually correct (**Supplementary Table 4**).

Differences among methods were most apparent for spliced reads (**Supplementary Tables 5–7**). On the first simulated data set, GSNAP, GSTRUCT, MapSplice and STAR mapped 96.3–98.4% of spliced reads to the correct locations and 0.9–2.9% to alternative locations (**Fig. 3** and **Supplementary Table 6**). Although these mappers assigned nearly all spliced reads to the correct locus, the frequency of reads for which they aligned all bases correctly was substantially lower (60.3–89.3% of spliced reads from simulation 1; **Fig. 3**). In contrast, ReadsMap and the annotation-based TopHat2 protocol produced high rates of perfect spliced alignments and few partially correct ones (**Fig. 3** and **Supplementary Table 6**), a behavior consistent with the aforementioned lack of read truncation. However, ReadsMap also

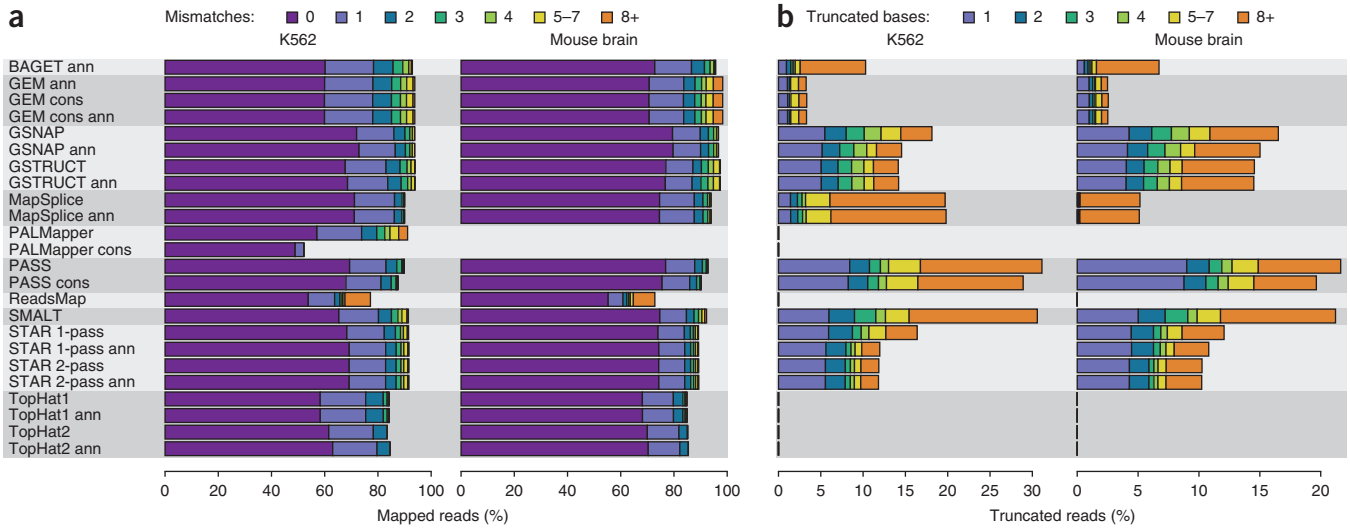


Figure 2 | Mismatch and truncation frequencies. (**a**) Percentage of sequenced reads mapped with the indicated number of mismatches. (**b**) Percentage of sequenced reads truncated at either or both ends. Bar colors indicate the number of bases removed.

assigned an exceptionally high proportion of bases to the wrong genomic positions, largely owing to a programmatic error that placed reads a few bases from their correct locations (**Fig. 3** and **Supplementary Table 5**).

The second simulated data set was designed to be more challenging, with higher frequencies of insertions and deletions (indels), base-calling errors and novel transcript isoforms. MapSplice, PASS and TopHat showed a reduction in performance on this data set relative to the other methods (**Fig. 3** and **Supplementary Tables 5**–**7**), results consistent with the low mismatch tolerance of these protocols (**Fig. 2a**).

Indel frequency and accuracy

GEM and PALMapper output included more indels than any other method (up to 115 indels per 1,000 K562 reads; **Fig. 4a** and **Supplementary Fig. 5**), but GEM preferentially reported insertions, and PALMapper, mostly deletions. Long deletions were most common with GSNAP and GSTRUCT, whereas TopHat2 called numerous long insertions. In contrast, PASS, ReadsMap and TopHat1 reported few long indels, and the conservative PALMapper protocols allowed only single-nucleotide indels.

These results were corroborated by analysis of indel accuracy on simulated data (**Fig. 4b**), which demonstrated that GEM and PALMapper report many false indels (indel precision < 37% for all protocols except PALMapper cons; simulation 1), that GSNAP and GSTRUCT exhibit high sensitivity for deletions largely independent of size (recall > 68% for each length interval depicted in **Fig. 4b**), and that the annotation-based TopHat2 protocol is the most sensitive method for long insertions (recall = 87% for insertions ≥5 bp; simulation 1). The ability of GSNAP, GSTRUCT and TopHat2 to detect long indels was accompanied by high false discovery rates, however, and MapSplice achieved a better balance between precision and recall for long deletions than GSNAP (**Fig. 4b**; this

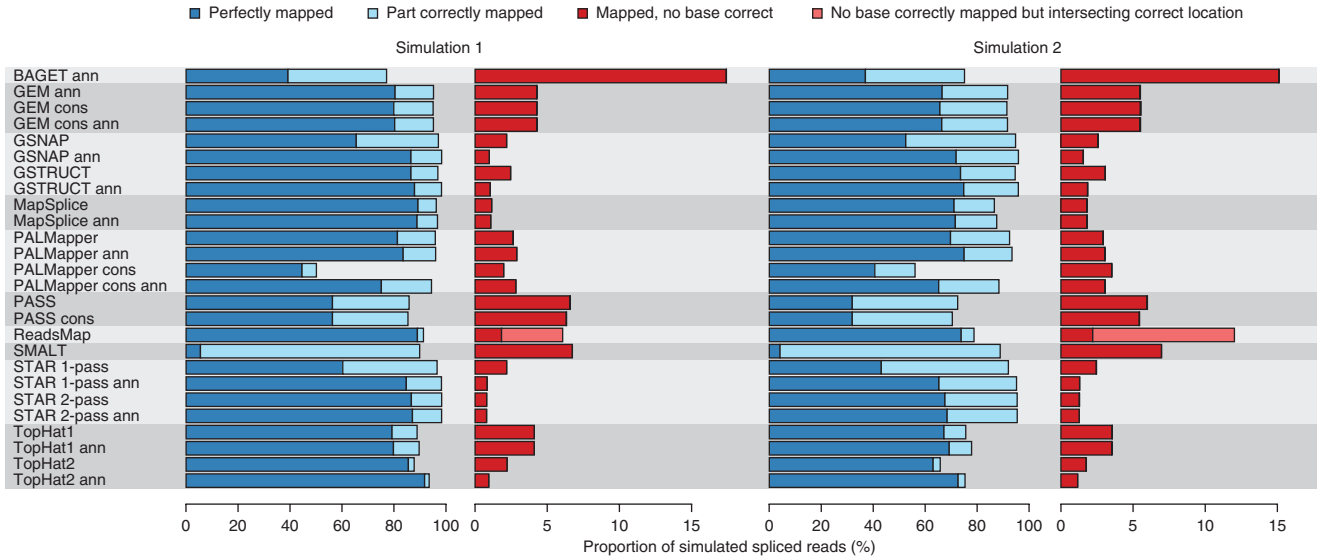


Figure 3 | Read placement accuracy for simulated spliced reads.

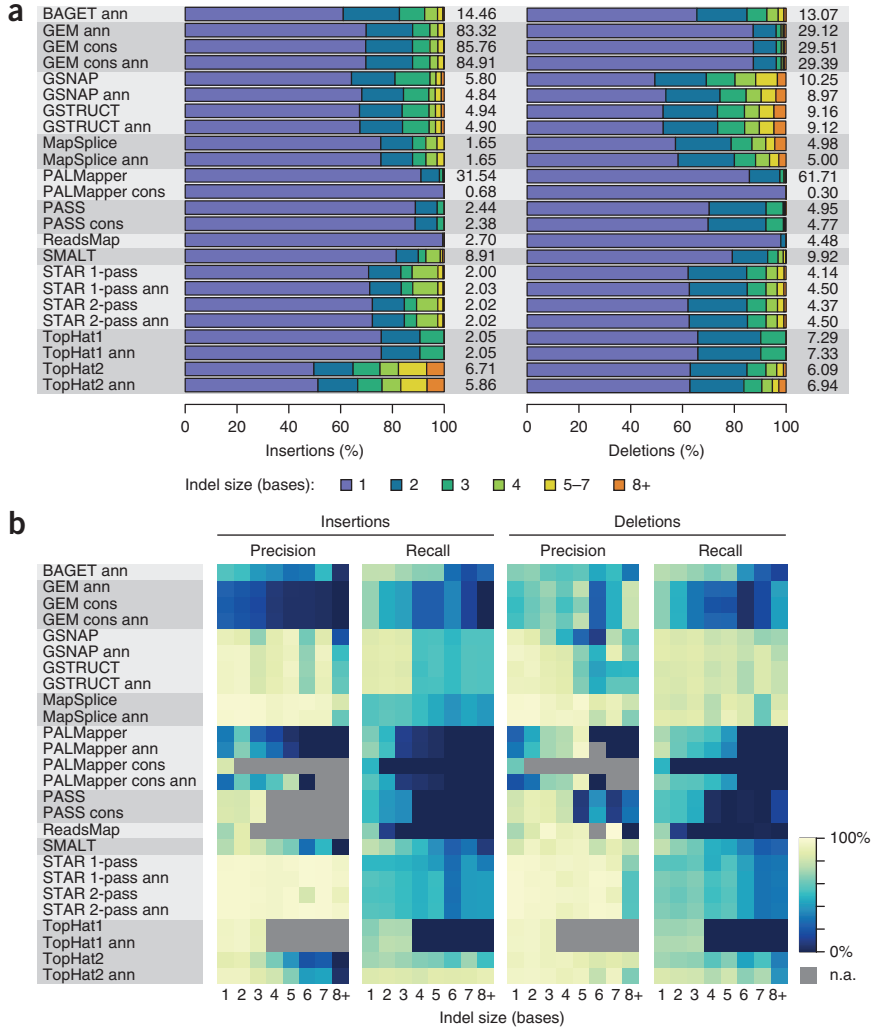


Figure 4 | Indel frequency and accuracy. (**a**) Bars show the size distribution of indels for the human K562 data set. Indel frequencies are tabulated (number of indels per 1,000 sequenced reads). (**b**) Precision and recall, stratified by indel size, for human simulated data set 1.

balance can be quantified using the *F*-score, which for deletions ≥5 bp was 87% for MapSplice and 36% for GSNAP on simulation 1 when these programs were executed without provision of gene annotation). **Supplementary Figure 6** illustrates alignments of two simulated reads that each contain a small insertion, resulting in erroneous mappings by several protocols.

Positioning of mismatches and gaps in reads

We determined the spatial distribution of mismatches, indels and introns over read sequences (**Supplementary Fig. 7**). All methods except MapSplice and PASS consistently reported an increasing frequency of mismatches along reads, in agreement with base-call quality-score distributions (**Supplementary Figs. 2** and **8**). BAGET, GEM, MapSplice, PALMapper and TopHat produced an excess of mismatches at read termini, whereas other methods avoided such a bias by truncating reads (**Fig. 2b**). Indels were preferentially placed near ends of reads by some methods, such as PALMapper and TopHat; others, such as MapSplice and STAR, tended to place them internally. GSTRUCT produced the most uniform distribution of indel frequency over the K562 data (coefficient of variation (CV) = 0.32), and TopHat produced the most variable (CV = 1.5 and 1.1 for TopHat1 and TopHat2, respectively). The positioning of splice junctions was generally more even, although several methods did not call junctions near read termini (**Supplementary Fig. 7**).

Coverage of annotated genes

We assessed how RNA-seq reads were placed in relation to annotated gene structures from the Ensembl database (**Supplementary Note**). Given the extensive annotation of the human and mouse genomes, the majority of reads would be expected to originate from known exons. Experimental data will also contain an unknown fraction of sequencing reads from unannotated transcripts and heterogeneous nuclear RNA. The simulated data sets were generated to recapitulate these features (Online Methods). Mapping trends were typically very similar between real and simulated data, a result indicating that simulation results reflect alignment performance in real RNA-seq experiments (**Supplementary Figs. 9–11**). The number of reads mapped to annotated exons were highest for GSNAP and GSTRUCT, on both real and simulated data, and close to the true number for the latter (**Supplementary Figs. 9–12**). However, all methods dispersed reads across too many genes: whereas reads from the first simulation should map to 16,554 Ensembl genes, all protocols reported primary alignments for more than 17,800 genes. This effect was largely due to the placement of reads at pseudogenes and was most severe for SMALT, BAGET and GEM (**Supplementary Figs. 9–11**).

Spliced alignment

In assessing spliced-alignment performance, we distinguish between detection of splices in individual reads and detection of unique splice junctions on the genomic sequence. The latter are often supported by multiple splices depending on expression level and sequencing depth. In general, GSNAP, GSTRUCT, ReadsMap, STAR and TopHat2 reported more (predicted) splices than other aligners (**Fig. 5a** and **Supplementary Table 2**). However, these results differed among protocol variants, such that GSNAP, STAR 1-pass and TopHat2 produced substantially fewer spliced mappings unless alignment was guided by known splice sites. SMALT, BAGET, PASS and the conservative PALMapper protocols inferred the fewest splices from the data (**Fig. 5a** and **Supplementary Fig. 13**). Several methods reported numerous splices not corresponding to known introns, particularly ReadsMap and PALMapper, and, to a lesser extent, SMALT, GSTRUCT and STAR 2-pass (**Fig. 5a**). These novel splice junctions were typically supported by few alignments, and many featured noncanonical splice signals, which suggests that they may be incorrect (**Fig. 5b** and **Supplementary Figs. 14** and **15**).

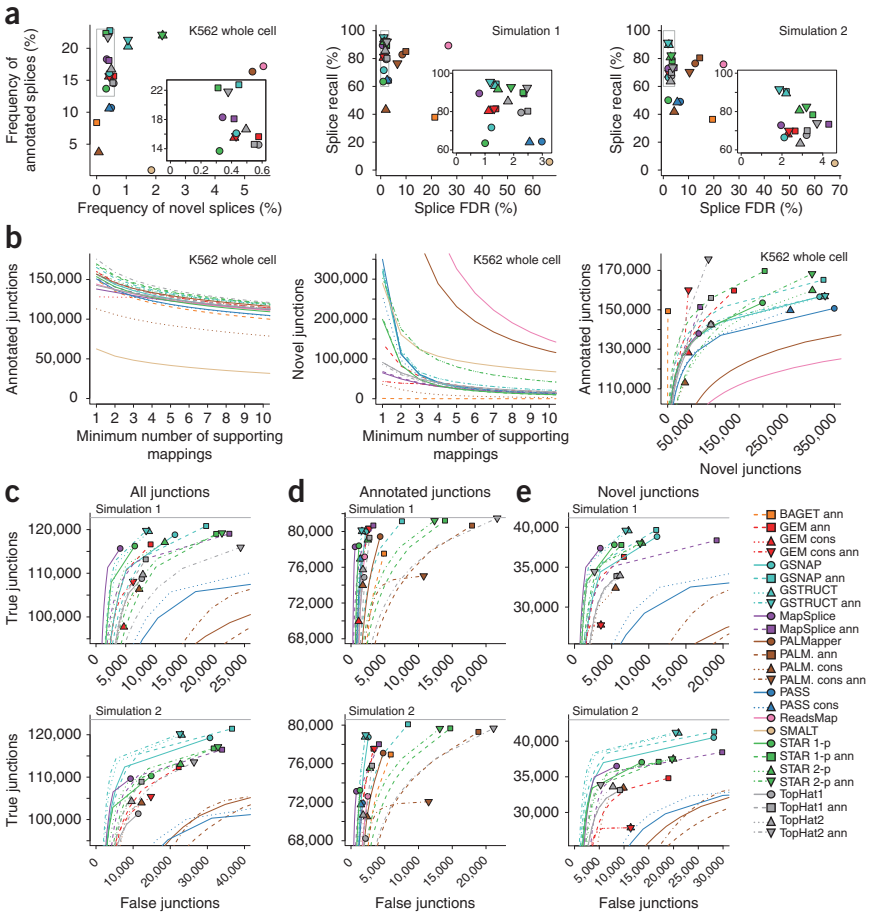


Figure 5 | Spliced alignment performance. (**a**) Frequency and accuracy of splices in primary alignments. Splice frequency was defined as the number of reported splices divided by the number of sequenced reads. For simulated data (center and right), splice recall and false discovery rate (FDR) is presented. Insets show details of the dense upper-left areas (gray rectangles). (**b**) Number of annotated and novel junctions reported at different thresholds for the number of supporting mappings. In the rightmost plot, filled symbols depict the number of junctions with at least one supporting mapping, and lines demonstrate the result of thresholding. (**c**) Junction discovery accuracy for simulated data set 1 (top) and 2 (bottom). Counts of true and false junctions were computed at increasing thresholds for the number of supporting mappings, and results were depicted as in **b** to obtain receiver operating characteristic–like curves. Gray horizontal lines indicate the number of junctions supported by true simulated alignments. (**d**) Accuracy for the subset of junctions contained in the Ensembl annotation. (**e**) Accuracy for junctions absent from the Ensembl annotation.

A substantial proportion were exclusive to particular methods. For example, 52–54% of the novel junctions reported by GSNAP/GSTRUCT on K562 whole-cell RNA were absent from the output of all other mappers (**Supplementary Table 8**).

Analysis of splice-detection performance on simulated data confirmed a substantial false discovery rate for ReadsMap, PALMapper and SMALT, whereas the highest accuracy was achieved by protocols based on GSNAP, GSTRUCT, MapSplice and STAR (**Fig. 5a**). Splices near the ends of reads can be particularly difficult to align, as a minimum amount of sequence is needed to confidently identify exon boundaries. Accuracy improved when the assessment was restricted to splices located between positions 20 and 57 in the 76-nt reads, but the same four methods still performed best (**Supplementary Fig. 16**). The use of simulated data further allowed us to measure the rate at which splices were detected in individual reads as a function of true coverage at corresponding junctions. Most protocols displayed decreased sensitivity at junctions covered by <5 reads (**Supplementary Fig. 17**). This reflects the reliance on junction coverage by alignment algorithms to increase precision. Accordingly, the trend was absent for methods that align each read independently (BAGET, GSNAP, PASS, SMALT and STAR 1-pass). Notably, the annotation-based GSNAP protocol achieved high sensitivity irrespective of junction coverage (**Supplementary Fig. 17**).

The number of false junction calls was considerable for most protocols but was greatly reduced if junctions were filtered by supporting alignment counts (**Fig. 5c**). At a threshold of two alignments, GSTRUCT outperformed most other methods on both simulated data sets when assessed by numbers of true and false junction calls (**Fig. 5c** and **Supplementary Tables 2** and **9**). MapSplice displayed similar performance on the first simulated data set, but only if used without annotation.

The simulated transcriptomes contain a subset of splice junctions in the Ensembl annotation as well as junctions from other gene catalogs and those created by simulating alternate isoforms of known genes. This corresponds to a realistic scenario wherein a subset of known transcripts are expressed in the assayed sample and knowledge of the transcriptome is incomplete. Protocols using annotation recovered nearly all of the known junctions in expressed transcripts, but most of these protocols also aligned reads at thousands of annotated junctions that were not expressed the simulated transcriptomes (**Fig. 5d**). This effect was particularly severe for TopHat2, PALMapper and STAR. For novel-junction discovery, GSTRUCT and MapSplice outperformed other methods (**Fig. 5e**).

Most programs could detect three or more splices per read, but PASS and PALMapper rarely reported more than two, and BAGET and SMALT never reported more than one (**Supplementary Fig. 18** and **Supplementary Table 10**). In general, ReadsMap, STAR and the annotation-based TopHat2 protocol produced the most primary alignments with at least three splices. The last protocol was also the most sensitive for recovering such multi-intron alignments from the simulated reads (recall = 79.3% for simulation 1; **Supplementary Table 11**). Among the protocols run without annotation, ReadsMap exhibited the best recall for alignments spanning three or more introns (72.1%), followed by the 2-pass version of STAR (70.7%) and GSTRUCT (65.8%).

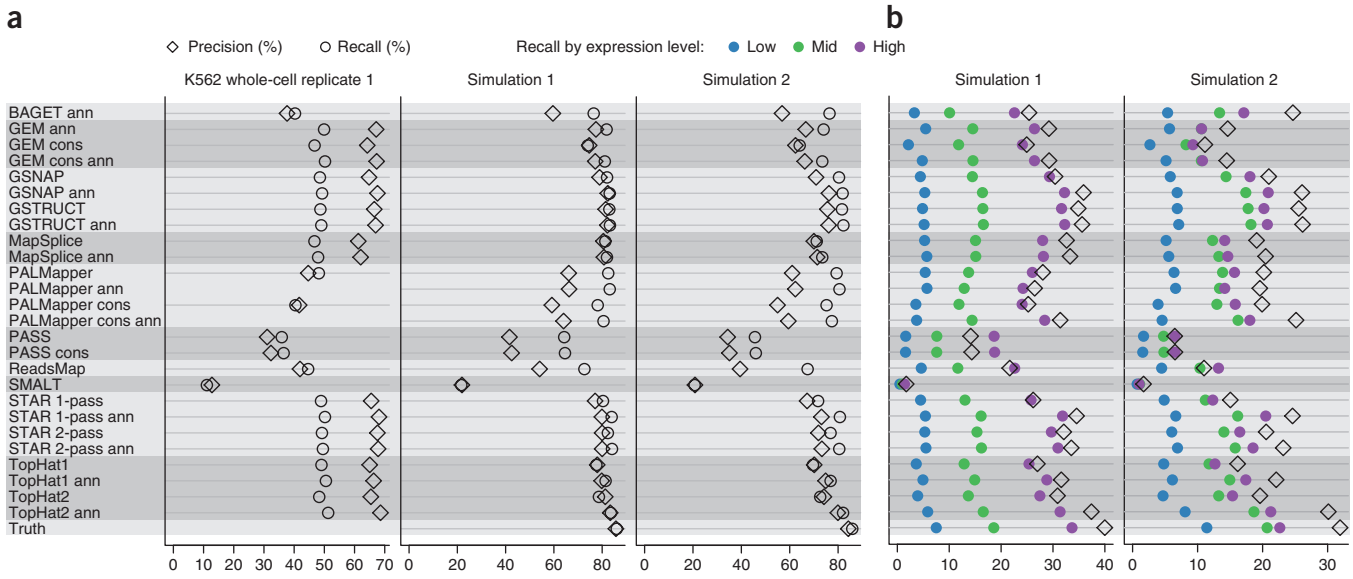


Figure 6 | Aligner influence on transcript assembly. (**a**,**b**) Cufflinks performance was assessed by measuring precision and recall for individual exons (**a**) and spliced transcripts (**b**). For K562 data, precision was defined as the fraction of predicted exons matching Ensembl annotation, and recall as the fraction of annotated protein-coding gene exons that were predicted.

However, ReadsMap also exhibited exceptionally low precision for such alignments (**Supplementary Table 11**).

Influence of aligners on transcript reconstruction

To assess the impact of alignment methodology on exon discovery and transcript reconstruction, we applied the transcript assembly program Cufflinks to the alignments. Exon detection results based on K562 data were similar for GEM, GSNAP, GSTRUCT, MapSplice, STAR and TopHat (**Fig. 6a**). With the K562 whole-cell RNA primary alignments from these methods, up to 69% of the exons reported by Cufflinks matched Ensembl annotation, and up to 51% of all exons from annotated protein-coding genes were recovered. Performance was substantially lower with output from the other alignment programs (**Fig. 6a**). Inclusion of secondary alignments negatively affected transcript reconstruction for methods that reported numerous such alignments (GEM and PALMapper) but typically had a small effect for other methods (**Supplementary Fig. 19**).

The six aligners noted above also enabled highly accurate exon detection on the first simulated data set, with recall reaching 84% and precision 83% (**Fig. 6a**). On the second, more challenging simulated data set, the TopHat2 protocol using annotation outperformed other methods, followed by GSNAP (with annotation) and GSTRUCT (with or without annotation) (**Fig. 6a**). The same protocols gave the best Cufflinks accuracy for the more complex task of reconstructing spliced transcripts (**Fig. 6b**).

It should be noted that the advantage of the annotation-based TopHat2 protocol was apparent only for reconstruction of exons and transcripts present in the annotation provided to aligners (**Supplementary Table 12**). This observation is consistent with the unique approach of TopHat2 involving read alignment to full-length annotated transcript sequences. It may seem paradoxical that several methods exhibiting relatively poor precision for junction alignments (**Fig. 5c–e**) produced high-quality input for transcript reconstruction. However, the Cufflinks algorithm is able to discard erroneous exon junctions in the input data at a high rate. For example, on the data from the first simulation, 71% of true junctions identified by the annotation-based TopHat2 protocol were incorporated into transcripts by Cufflinks, compared to 5% of false junctions (**Supplementary Table 13**).

DISCUSSION

In general, GSNAP, GSTRUCT, MapSplice and STAR compared favorably to the other methods, consistent with an earlier evaluation that included a subset of these tools[18]. Our assessment shows MapSplice to be a conservative aligner with respect to mismatch frequency, indel and exon junction calls. Conversely, the most significant issue with GSNAP, GSTRUCT and STAR is the presence of many false exon junctions in the output. This can be ameliorated by filtering junctions on the number of supporting alignments. It should be noted that both GSNAP and GSTRUCT require considerable computing time when parameterized for sensitive spliced alignment[7], and the GSTRUCT pipeline has not yet been released. A recent runtime comparison found GSNAP and MapSplice to perform similarly, whereas TopHat2 and STAR were about 3 and 180 times faster, respectively[9].

RNA-seq aligners use gene annotation to achieve better placement of spliced reads, and the resulting improvement was apparent on several metrics, particularly for GSNAP and the 1-pass version of STAR. Notably, these programs align each read independently, and the effect of using annotation was generally less pronounced for tools that carry out splice-junction discovery before final alignment, such as GEM, MapSplice, GSTRUCT and STAR 2-pass. TopHat also belongs to this class of programs, but provision of annotation still had a major effect on TopHat2 results, most likely because of the unique strategy whereby reads are aligned directly against annotated transcripts. This approach is clearly effective in several respects but may be suitable only for genomes with near-complete annotation.

Remaining challenges include exploiting gene annotation without introducing bias, correctly placing multimapped reads, achieving optimal yet fast alignment around gaps and mismatches, and

reducing the number of false exon junctions reported. Ongoing developments in sequencing technology will demand efficient processing of longer reads with higher error rates and will require more extensive spliced alignment as reads span multiple exon junctions. We expect performance of the aligners evaluated here to improve as current shortfalls are addressed. Differential treatment of these issues will enhance and expand the range of RNA-seq aligners suited to varied computational methodologies and analysis aims.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by the European Molecular Biology Laboratory, US National Institutes of Health/NHGRI grants U54HG004555 and U54HG004557, Wellcome Trust grant WT09805, and grants BIO2011-26205 and CSD2007-00050 from the Ministerio de Educación y Ciencia.

AUTHOR CONTRIBUTIONS

P.B., R.G., J.H., T.J.H. and N.G. conceived of and organized the study. G.R.G. and B.S. created the simulated RNA-seq data. Consortium members provided alignments for evaluation. P.G.E., T.S., B.S. and G.R.G. analyzed the data. P.G.E. and P.B. coordinated the analysis and wrote the paper with input from the aforementioned authors. A.K. and G.R. carried out preliminary analysis and metric development based on earlier RNA-seq and alignment data but did not evaluate the alignments described herein.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Kent, W.J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
2. Fonseca, N.A., Rung, J., Brazma, A. & Marioni, J.C. Tools for mapping high-throughput sequencing data. *Bioinformatics* **28**, 3169–3177 (2012).
3. Marco-Sola, S., Sammeth, M., Guigó, R. & Ribeca, P. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat. Methods* **9**, 1185–1188 (2012).
4. Wang, K. *et al.* MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* **38**, e178 (2010).
5. Trapnell, C., Pachter, L. & Salzberg, S.L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
6. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).
7. Wu, T.D. & Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873–881 (2010).
8. Jean, G., Kahles, A., Sreedharan, V.T., De Bona, F. & Rätsch, G. RNA-Seq read alignments with PALMapper. *Curr. Protoc. Bioinformatics* **32**, 11.6 (2010).
9. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
10. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
11. Guigó, R. *et al.* EGASP: the human ENCODE Genome Annotation Assessment Project. *Genome Biol.* **7** (suppl. 1), S2 (2006).
12. Moult, J., Fidelis, K., Kryshtafovych, A. & Tramontano, A. Critical assessment of methods of protein structure prediction (CASP)—round IX. *Proteins* **79** (suppl. 10), 1–5 (2011).
13. Earl, D. *et al.* Assemblathon 1: a competitive assessment of *de novo* short read assembly methods. *Genome Res.* **21**, 2224–2241 (2011).
14. Marbach, D. *et al.* Wisdom of crowds for robust gene network inference. *Nat. Methods* **9**, 796–804 (2012).
15. Campagna, D. *et al.* PASS: a program to align short sequences. *Bioinformatics* **25**, 967–968 (2009).
16. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931–945 (2004).
17. Keane, T.M. *et al.* Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* **477**, 289–294 (2011).
18. Grant, G.R. *et al.* Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics* **27**, 2518–2528 (2011).

RNA-seq Genome Annotation Assessment Project (RGASP) Consortium

Tyler Alioto[14], Jonas Behr[4,5], Paul Bertone[1,10–12], Regina Bohnert[5], Davide Campagna[15], Carrie A Davis[16], Alexander Dobin[16], Pär G Engström[1,13], Thomas R Gingeras[16], Nick Goldman[1], Gregory R Grant[2,3], Roderic Guigó[8,9], Jennifer Harrow[7], Tim J Hubbard[7], Géraldine Jean[5], André Kahles[4,5], Peter Kosarev[17], Sheng Li[18], Jinze Liu[19], Christopher E Mason[18], Vladimir Molodtsov[17], Zemin Ning[7], Hannes Ponstingl[7], Jan F Prins[20], Gunnar Rätsch[4,5], Paolo Ribeca[14], Igor Seledtsov[17], Botond Sipos[1], Victor Solovyev[21], Tamara Steijger[1], Giorgio Valle[15], Nicola Vitulo[15], Kai Wang[19], Thomas D Wu[22] & Georg Zeller[5]

[14]Centro Nacional de Análisis Genómico, Barcelona, Spain. [15]CRIBI Biotechnology Centre, Università di Padova, Padova, Italy. [16]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, USA. [17]Softberry Inc., Mount Kisco, New York, USA. [18]Department of Physiology and Biophysics, Weill Cornell Medical College, New York, New York, USA. [19]Department of Computer Science, University of Kentucky, Lexington, Kentucky, USA. [20]Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [21]Department of Computer Science, Royal Holloway, University of London, London, UK. [22]Department of Bioinformatics and Computational Biology, Genentech, San Francisco, California, USA.

ONLINE METHODS

RNA-seq data. The human K562 data used here correspond to the K562 poly(A)$^+$ RNA samples produced at Cold Spring Harbor Laboratory for the ENCODE project[19] and can be accessed at http://www.encodeproject.org/. RNA-seq libraries were sequenced using a strand-specific protocol and comprise two biological replicates each of whole-cell, cytoplasmic and nuclear RNA. The mouse RNA-seq data set was produced at the Wellcome Trust Sanger Institute as part of the Mouse Genomes Project using brain tissue from adult mice of strain C57BL/6NJ. The library was constructed using the standard Illumina protocol that does not retain strand information. These data have been previously described[20] and are available from the European Nucleotide Archive (http://www.ebi.ac.uk/ena/) under accessions ERR033015 and ERR033016. All of the data used in this study have been consolidated as a single experimental record in the ArrayExpress repository (http://www.ebi.ac.uk/arrayexpress/) under accession E-MTAB-1728.

Simulated RNA-seq data were generated using the BEERS toolkit (http://cbil.upenn.edu/BEERS/), and additional modeling of base-call errors and quality scores was done with simNGS (http://www.ebi.ac.uk/goldman-srv/simNGS/). BEERS has been previously described[18]. Briefly, the simulator takes as input a database of transcript models and a quantification file that specifies expression levels for each transcript and intron in the database. A transcriptome is simulated by sampling a specified number of transcript models from the database at random and creating additional alternative splice forms from each model. Polymorphisms (indels and substitutions) are introduced into the exons according to independent rates. Reads are then produced from the transcriptome in an iterative manner. In each iteration, a transcript is chosen with probability proportional to its expression level in the quantification file. An intron may be left in, with probability based on the intronic expression levels in the quantification file. A fragment of normally distributed length is sampled from the transcript, and the L bases from each end of this fragment are reported, where L is the read length.

Here, the simulator was executed using the transcript database and quantification file previously described[18]. This database comprises 538,991 transcript models merged from 11 annotation tracks available from the UCSC Genome Browser (AceView, Ensembl, Geneid, Genscan, NSCAN, Other RefSeq, RefSeq, SGP, Transcriptome, UCSC and Vega), and expression levels were derived from a human retina RNA-seq data set. In each of the two simulations, 25,000 transcripts were randomly chosen from the database, and two additional alternative isoforms were generated for each sampled transcript. The proportion of signal originating from novel isoforms was 20% and 35% for simulation 1 and 2, respectively. Substitution variants were introduced into exons at rates of 0.001 (simulation 1) and 0.005 (simulation 2) events per base pair, and indel polymorphisms at rates of 0.0005 (simulation 1) and 0.0025 (simulation 2). The simulated transcriptomes included 136,226 (simulation 1) and 134,717 (simulation 2) unique splice junctions, of which 90% and 92%, respectively, were represented in the simulated reads (**Supplementary Table 9**).

The option to simulate sequencing errors was disabled. Instead, the program simNGS was used to add noise to the simulated reads. simNGS recreates observations from Illumina sequencing machines using the statistical models underlying the AYB base-calling software[21]. Here, base-call errors and quality scores were simulated by applying simNGS version 1.5 with a paired-end simulation model. The model was trained on intensity data released by Illumina from a sequencing run on the HiSeq 2000 instrument using TruSeq chemistry. The resulting quality-score distributions are shown in **Supplementary Figure 8**, and the correct alignments of simulated data have been deposited in ArrayExpress under accession E-MTAB-1728.

Alignment protocols making use of gene annotation were provided with annotation from Ensembl only (**Supplementary Note**), whereas the simulated transcriptomes were based on Ensembl as well as several additional gene catalogs. In addition, novel transcript isoforms and retained introns were simulated, as detailed above. This reflects a realistic scenario where knowledge of the transcriptome is incomplete even for well-studied organisms, and a proportion of transcripts captured by RNA-seq correspond to pre-spliced mRNAs.

Read alignment. Developer teams were provided with RNA-seq data, human and mouse reference genome sequences, and transcript annotations from the Ensembl database. So that we avoided potential biases, teams were not informed of the final evaluation criteria and were not given the true results for simulated data. Developers providing alignments for evaluation could not access submissions from other teams and were prohibited from participating in the analysis phase as part of the study design. Details of alignment protocols are provided in the **Supplementary Note**.

Evaluation of alignments. Developer teams provided alignments in BAM format. These files were processed to ensure compliance with the SAM specification[22] and eliminate formatting discrepancies that otherwise could have affected the evaluation. Mismatch information (NM and MD tags) was stripped from the files and recomputed using the SAMtools command "calmd" to ensure that mismatches were counted in the same manner for all protocols[22]. The resulting alignment files have been deposited in ArrayExpress under accession E-MTAB-1728.

With inspiration from earlier benchmarking studies[9,18,23], we devised several performance metrics to assess attributes ranging from fundamental (for example, proportion of mapped reads and base-level alignment characteristics) to advanced, including splice junction detection, read placement around indels and suitability of alignments for transcript reconstruction. A detailed description of evaluation metrics is provided in the **Supplementary Note**, and key results are summarized in **Supplementary Table 2**. Unless otherwise noted, evaluation metrics for alignments of K562 RNA-seq data were averaged over the six K562 data sets (**Supplementary Table 1**). A subset of K562 samples were not processed by PALMapper and ReadsMap (**Supplementary Table 3**). Comparisons with gene annotation were performed using the Ensembl annotation that was provided to aligners (**Supplementary Note**).

Treatment of alignment gaps. In the BAM format, alignment gaps in read sequences can be described as either deletions or introns. Small gaps are typically labeled deletions and longer gaps considered introns, but the exact criteria differ among aligners. To prevent the introduction of bias from such differences, we reclassified deletions and introns where appropriate. Specifically,

doi:10.1038/nmeth.2722

for the indel results presented in **Figure 4** and **Supplementary Figure 5** and the evaluation of splice accuracy on simulated data, an alignment gap in the read sequence was considered a deletion if shorter than 19 bp and otherwise counted as an intron. We aimed to select a threshold that would minimize relabeling of gaps in the read sequence, and we observed that only three methods (BAGET, GSNAP and GSTRUCT) reported a substantial frequency of deletions longer than 18 bp from any data set. Up to 2.0% of the deletions in the output from GSNAP and GSTRUCT exceeded 18 bp, compared to 0.16% for BAGET and <0.001% for all other methods. The adjustment noticeably affected the results for GSNAP and GSTRUCT only.

For alignments of simulated RNA-seq data, accuracy metrics were computed by comparison with the alignments produced by the simulator. For computation of basewise and indel accuracy, ambiguity in indel placement was accounted for[18]. For example, in an alignment of the sequences ATTTA and ATTA, there are three equivalent gap placements in the latter sequence (A-TTA, AT-TA and ATT-A), all of which were considered correct. A general strategy was implemented to handle positional ambiguity for indels of any size.

Transcript reconstruction. Transcript assembly was conducted with Cufflinks version 2.0.2. The option library-type was set to fr-firststrand for the K562 data, which are strand specific, and to fr-unstranded for the simulated data, which are not. Default values were used for other parameters.

Cufflinks requires spliced alignments to have a SAM format tag (XS) indicating the genomic strand (plus or minus) on which the transcript represented by the read is likely to be encoded. Alignment programs such as TopHat can set the XS tag by using information about the library construction protocol (for strand-specific libraries) or by inspecting sequence at exon-intron boundaries. Five of the methods evaluated here (BAGET, GEM, ReadsMap, SMALT and STAR) did not provide XS tags; we therefore post-processed the alignment output from these methods to add them. For the strand-specific K562 data, XS tags were set on the basis of alignment orientation and read number (first or second in pair), as done by TopHat. For alignments of simulated reads, we set XS tags according to the initial and terminal dinucleotides of the inferred introns, which are expected to be GT/AG, GC/AG or AT/AC for plus-strand transcripts and CT/AC, CT/GC or GT/AT for minus-strand transcripts[24]. For the XS tag to be added to an alignment, at least one exon junction was required to have these signals, and conflicting signals among junctions were not allowed.

We noted that the annotation-based TopHat2 protocol uses the annotation provided to set the XS tag for unspliced alignments that overlap annotated exons. As this is a unique feature of TopHat2 that might confer an advantage in the evaluation of transcript reconstruction, we investigated the effect of removing the XS tag from unspliced alignments in the TopHat2 output before running Cufflinks. This modification had a negligible effect on the Cufflinks accuracy metrics presented here (data not shown), demonstrating that provision of XS tags for unspliced alignments cannot explain why the annotation-based TopHat2 protocol resulted in better Cufflinks performance than other protocols.

For K562 data, exon precision was defined as the fraction of predicted exons matching GENCODE annotation, and exon recall as the fraction of annotated exons that were predicted. Only exons from protein-coding genes were considered when computing recall, as some noncoding RNA classes are likely to be underrepresented in the RNA-seq libraries. Results on simulated data were benchmarked against simulated gene models, using analogous definitions of precision and recall, such that exon precision measures the proportion of predicted exons matching an exon in the simulated transcriptome, and transcript precision is the fraction of predicted spliced transcripts matching a simulated spliced transcript. To stratify recall by expression, we divided simulated transcripts into three groups of equal size according to expression level (**Fig. 6b**). Internal exons were required to be recovered with exact boundaries, first and terminal exons were required to have correctly predicted internal borders only, and exons constituting unspliced transcripts were scored as correct if covered to at least 60% by a predicted unspliced transcript. For the simulated data, only exons of spliced transcripts were required to be placed on the correct strand, as the orientation of single-exon transcripts cannot be reliably predicted unless RNA-seq libraries are strand specific. Spliced transcripts were considered to be correctly assembled if the strand and all exon junctions matched.

Program availability. Source code for the evaluations performed in this study can be obtained from https://github.com/RGASP-consortium/.

19. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
20. Danecek, P. *et al.* High levels of RNA-editing site conservation amongst 15 laboratory mouse strains. *Genome Biol.* **13**, R26 (2012).
21. Massingham, T. & Goldman, N. All Your Base: a fast and accurate probabilistic approach to base calling. *Genome Biol.* **13**, R13 (2012).
22. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
23. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
24. Iwata, H. & Gotoh, O. Comparative analysis of information contents relevant to recognition of introns in many species. *BMC Genomics* **12**, 45 (2011).

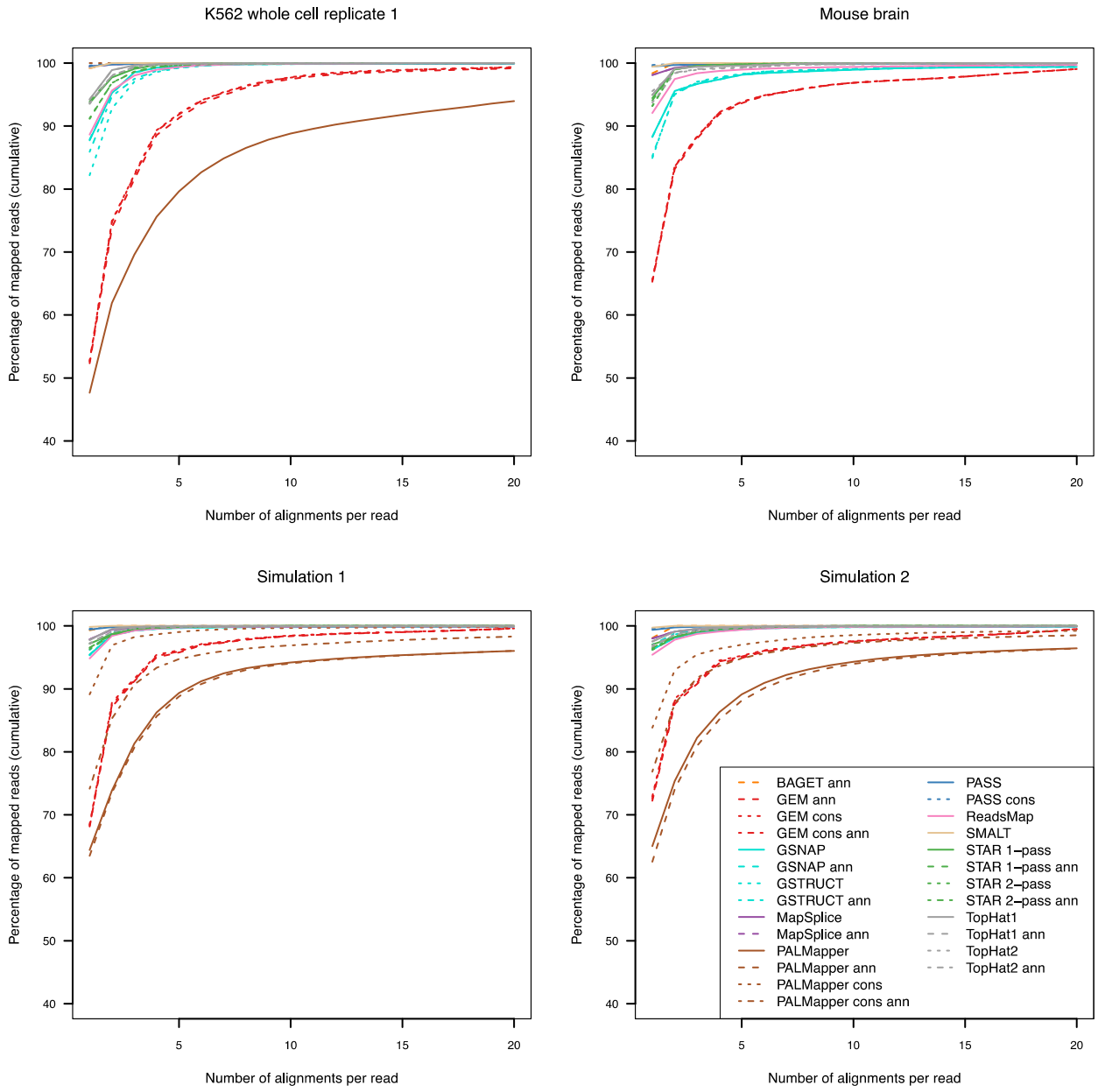
Systematic evaluation of spliced alignment programs for RNA-seq data

Pär G. Engström, Tamara Steijger, Botond Sipos, Gregory R. Grant, André Kahles, RGASP Consortium, Gunnar Rätsch, Nick Goldman, Tim J. Hubbard, Jennifer Harrow, Roderic Guigó and Paul Bertone

Supplement

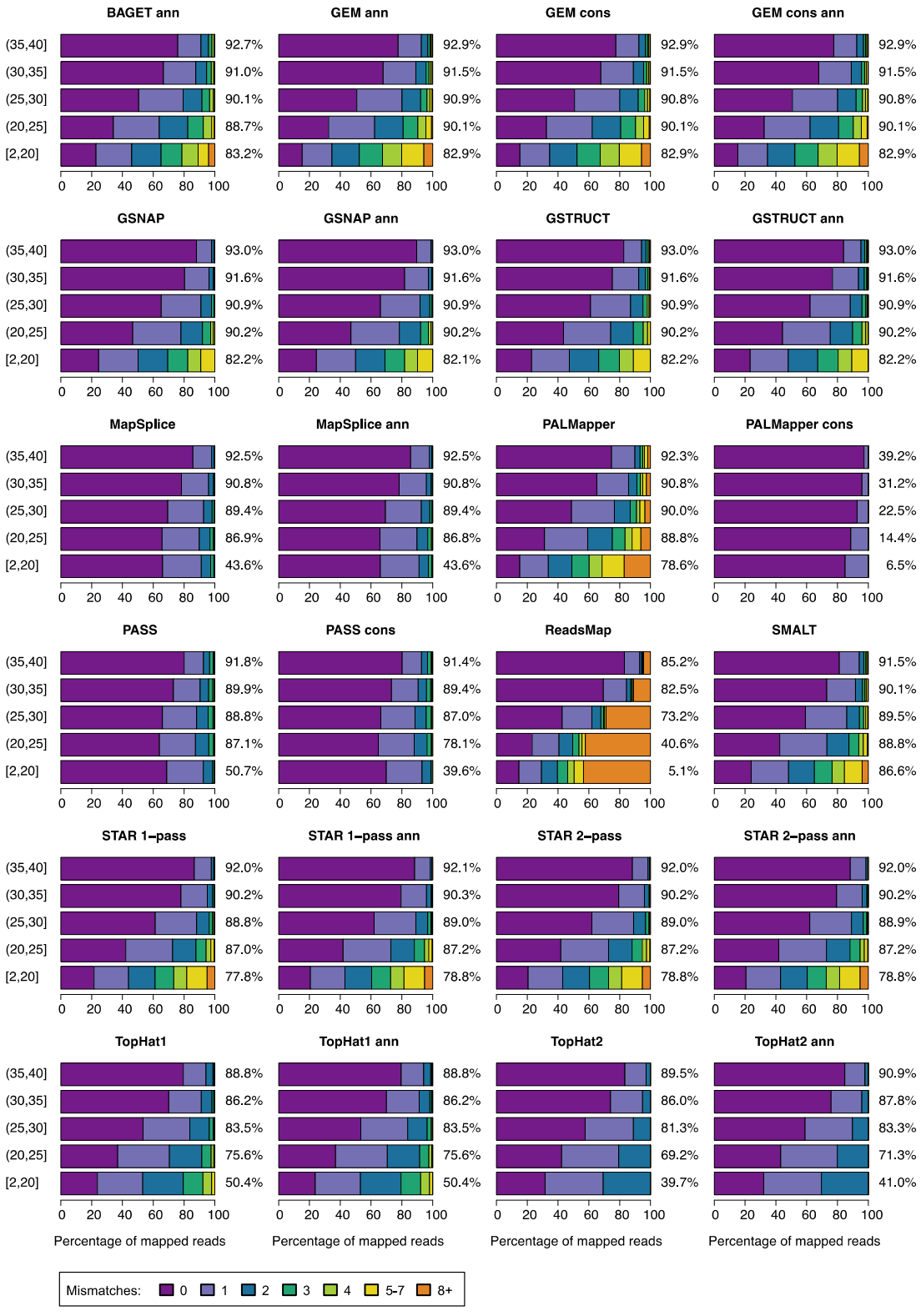
Contents

Nature Methods: doi:10.1038/nmeth.2722



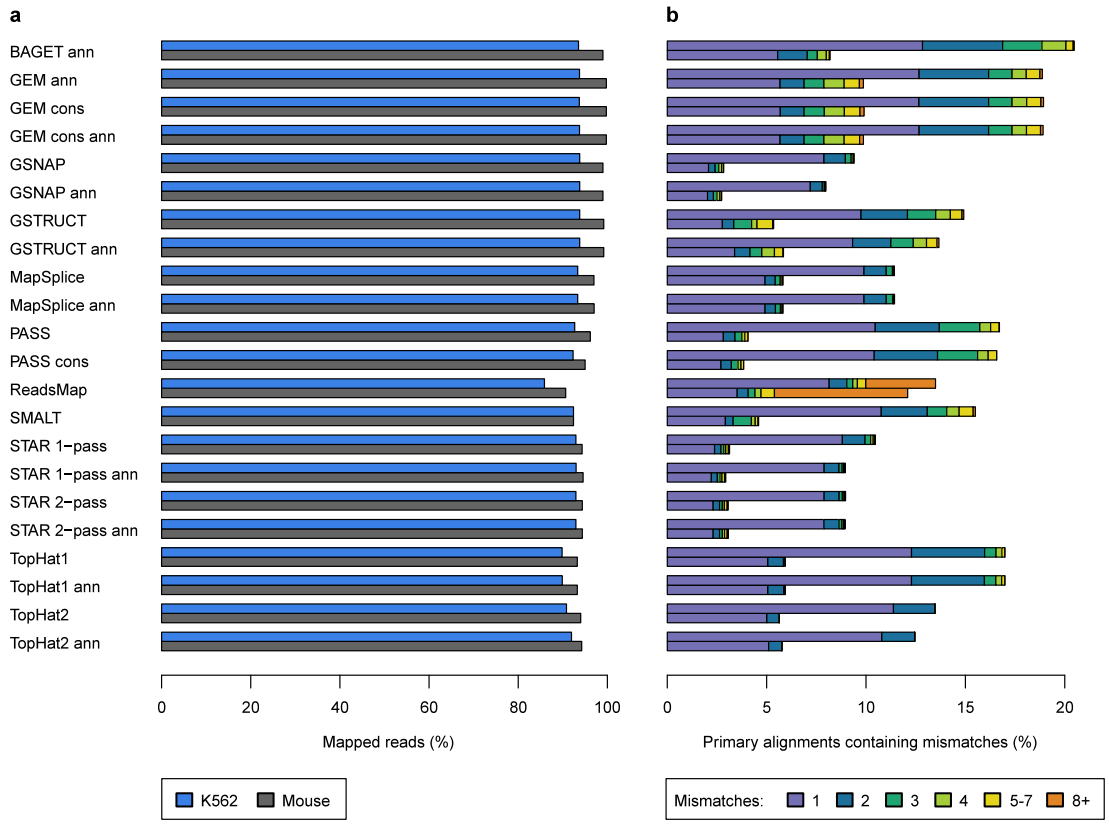
Supplementary Figure 1. Cumulative distribution of number of alignments per read. Distributions are shown for each protocol on four data sets. Note that PALMapper was not run on the mouse data, and only two of the four PALMapper protocols were applied to the K562 data (PALMapper and PALMapper cons).

Nature Methods: doi:10.1038/nmeth.2722



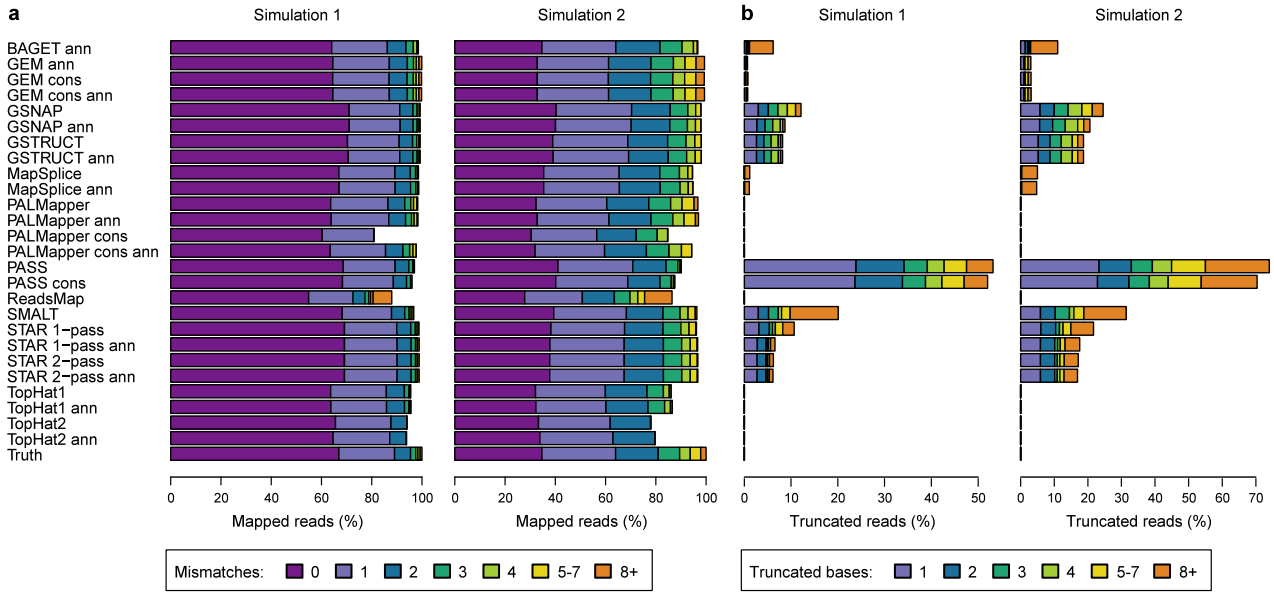
Supplementary Figure 2. Mismatch frequencies stratified by base caller quality scores. Results for K562 whole cell replicate 1 are shown. Reads were divided into five categories by mean quality score. Quality scores range from 2 to 40, with lower scores correponding to less confident base calls. Bars show distribution of mismatches per alignment, demonstrating that most methods tend to align low-quality reads with more mismatches. Percentages of aligned reads are tabulated for each protocol and quality score category, showing that protocols differ in the extent to which mappability depends on quality score.

Nature Methods: doi:10.1038/nmeth.2722



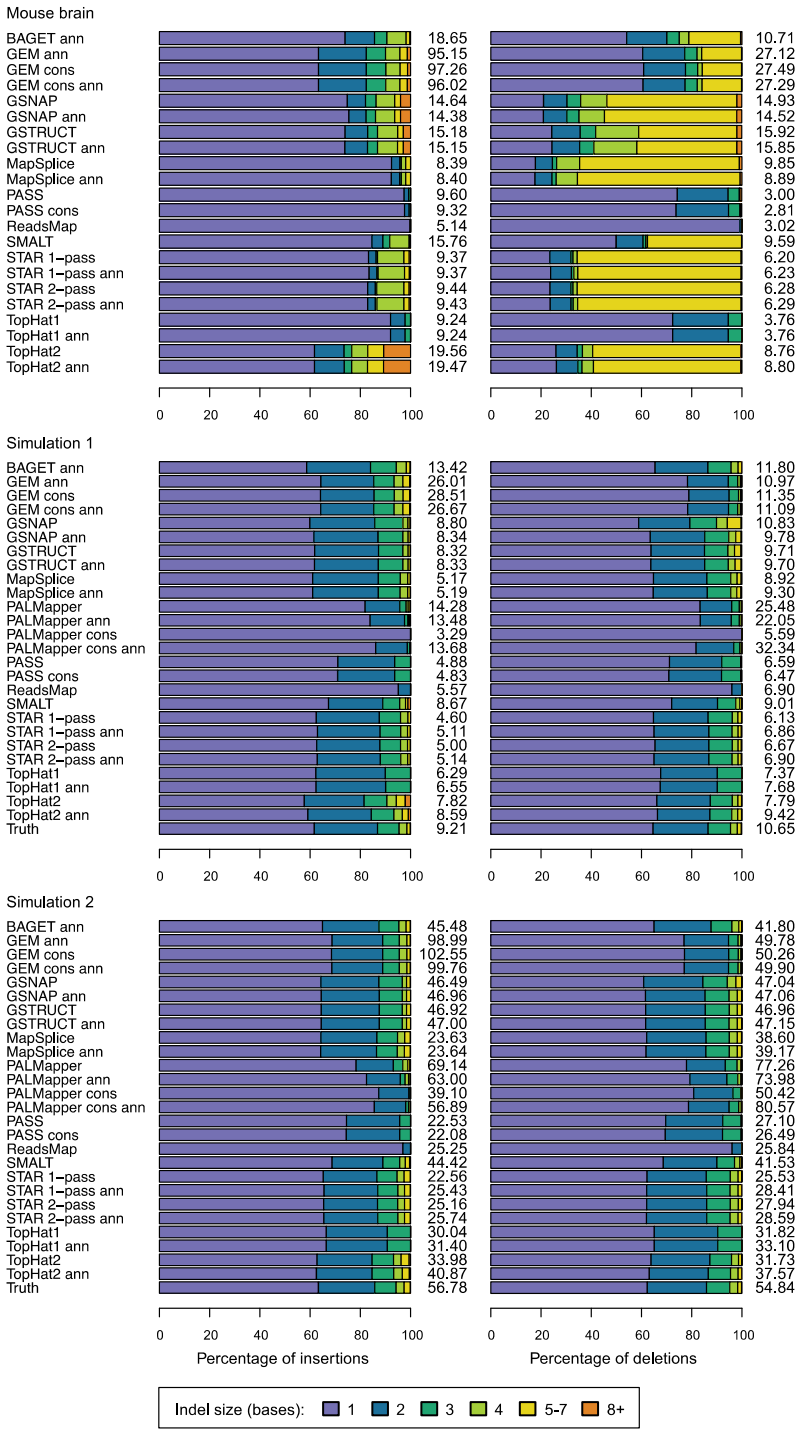
Supplementary Figure 3. Mapping statistics for high-quality reads from K562 and mouse. Mapping yield (a) and mismatch frequencies (b) are shown for reads with a mean base call quality score of at least 38. Results for K562 whole cell RNA replicate 1 (upper bar for each protocol) are compared to those for the mouse data set (lower bar). Mismatch frequencies represent the proportion of mapped reads for which the primary alignment contains the indicated number of mismatches.

Nature Methods: doi:10.1038/nmeth.2722

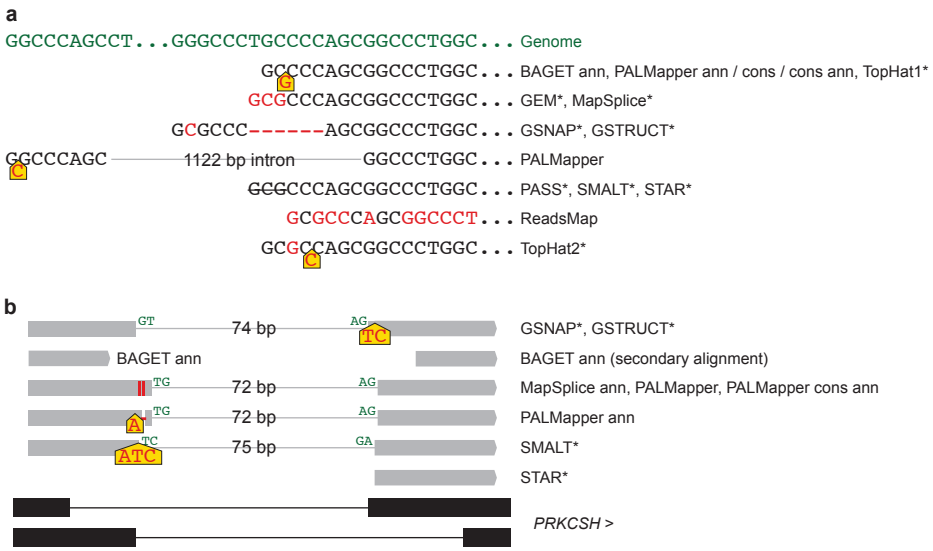


Supplementary Figure 4. Mismatch and truncation frequencies for alignments of simulated data. (a) Percentage of reads aligned with the indicated number of mismatches. (b) Percentage of reads that were truncated at either or both ends (colors indicate the number of bases removed per read). The bars labeled “Truth” show frequencies for the alignments produced by the simulator, corresponding to the results expected from a perfect aligner.

Nature Methods: doi:10.1038/nmeth.2722

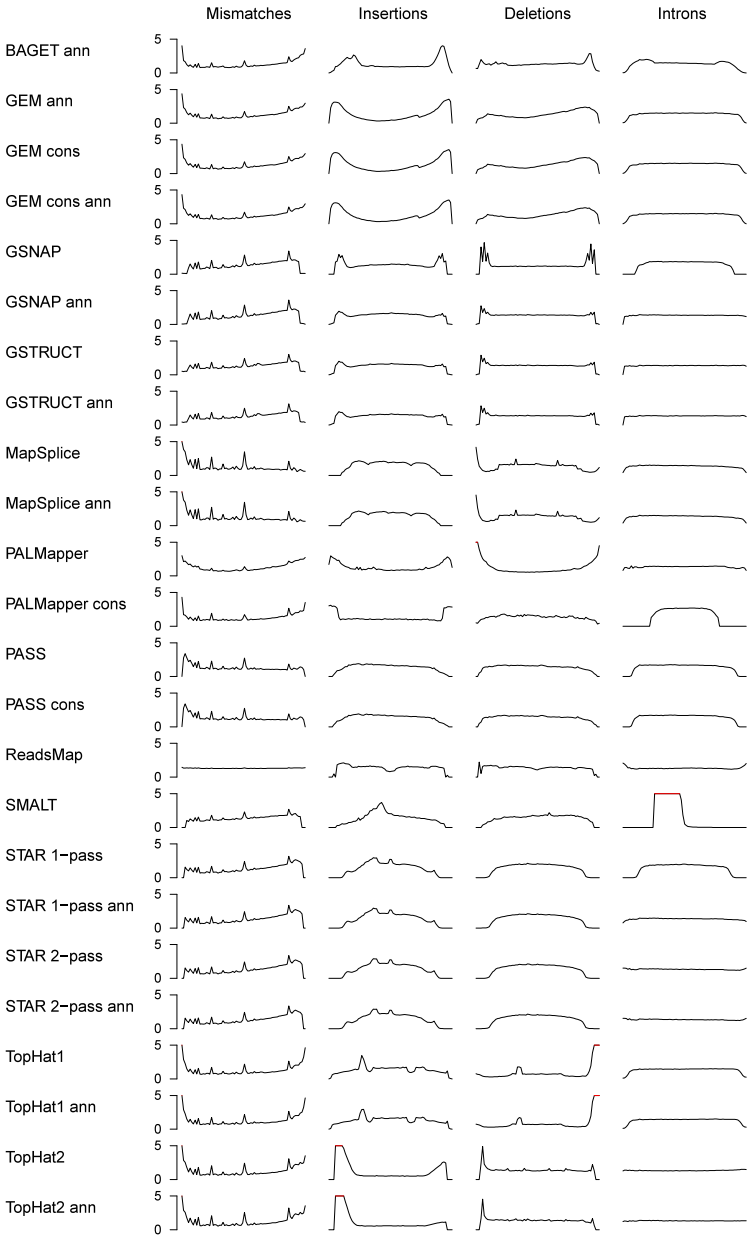


Supplementary Figure 5. Indel frequencies for mouse and simulated data. Bars show size distribution of indels. Indel frequencies are tabulated (number of indels per thousand sequenced reads). The mouse data set contains a significant number of 45S ribosomal RNA reads that align best with a six bp deletion to a locus on chromosome 17. For the two simulated data sets, the last bars show the results expected for a perfect aligner (Truth).



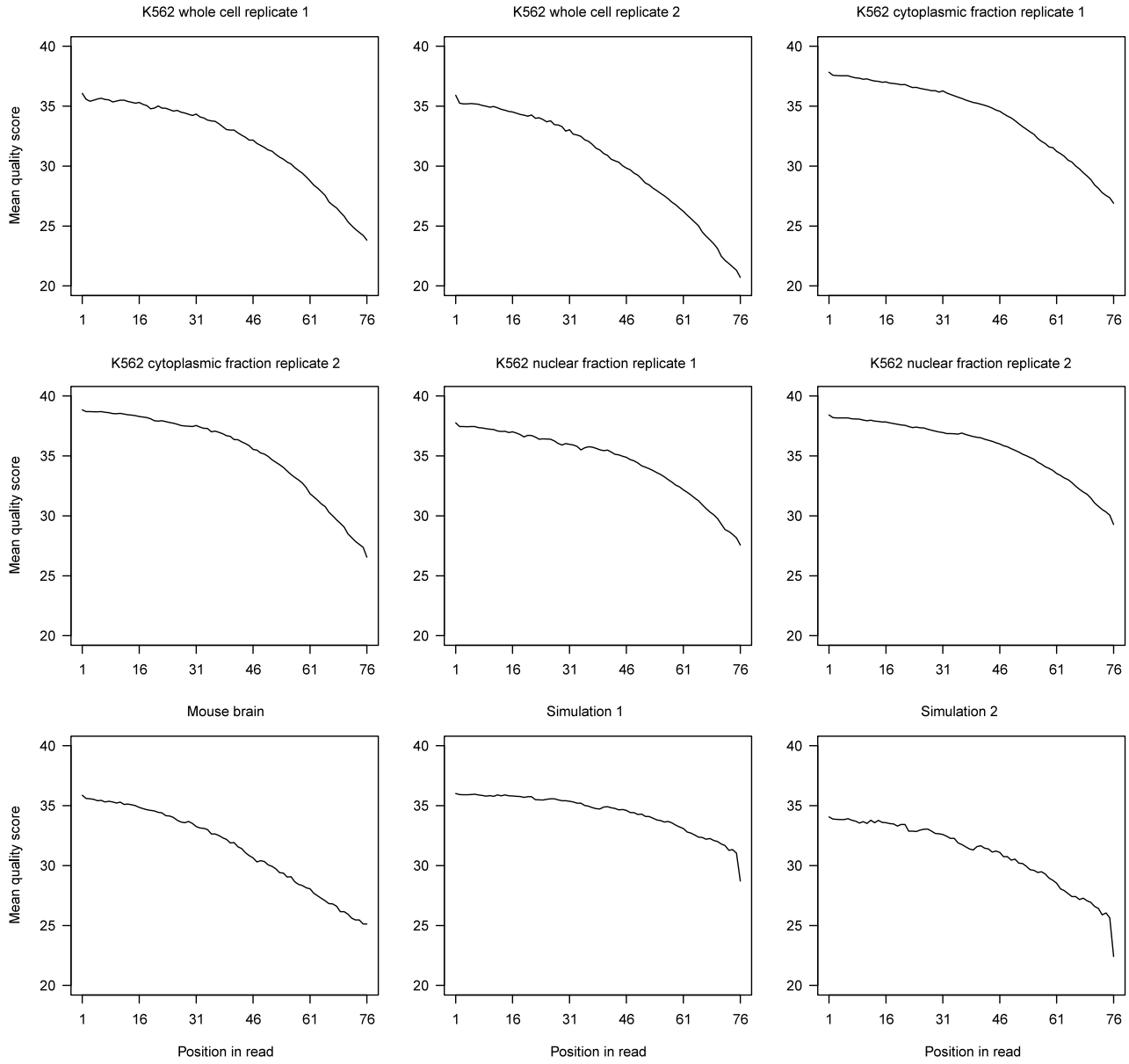
Supplementary Figure 6. Examples of mapping results for reads with small insertions. (a) Alignments of simulated read containing an insertion at the third position. All protocols mapped the read to the correct locus, but the exact simulated alignment was only recovered by BAGET, three PALMapper protocols and TopHat1. The first 18 bases of the read are shown. Mismatches (red), deletions (red dash) and insertions (red on yellow) are indicated. Asterisks indicate aligners for which all protocols produced the same alignment. The PALMapper base protocol errorneously predicted a 1122 bp intron with noncanonical acceptor and donor dinucleotides (CT, GC). PASS, SMALT and STAR truncated the first three positions of the read. ReadsMap placed the read three bases away from its correct location, resulting in 59 mismatches. (b) Alignments of a simulated read containing an insertion near a junction joining two exons of the gene *PRKCSH*. Only GSNAP and GSTRUCT recovered the simulated alignment. Grey bars represent aligned segments in genomic coordinates. Mismatches and gaps are colored as in panel c. Grey lines represent predicted introns. Only the correct aligment has canonical acceptor and donor dinucleotides (GT..AG, green). Annotated *PRKCSH* junctions are shown in black. All reported primary alignments are shown. GEM, PASS, ReadsMap and TopHat did not map the read.

Nature Methods: doi:10.1038/nmeth.2722



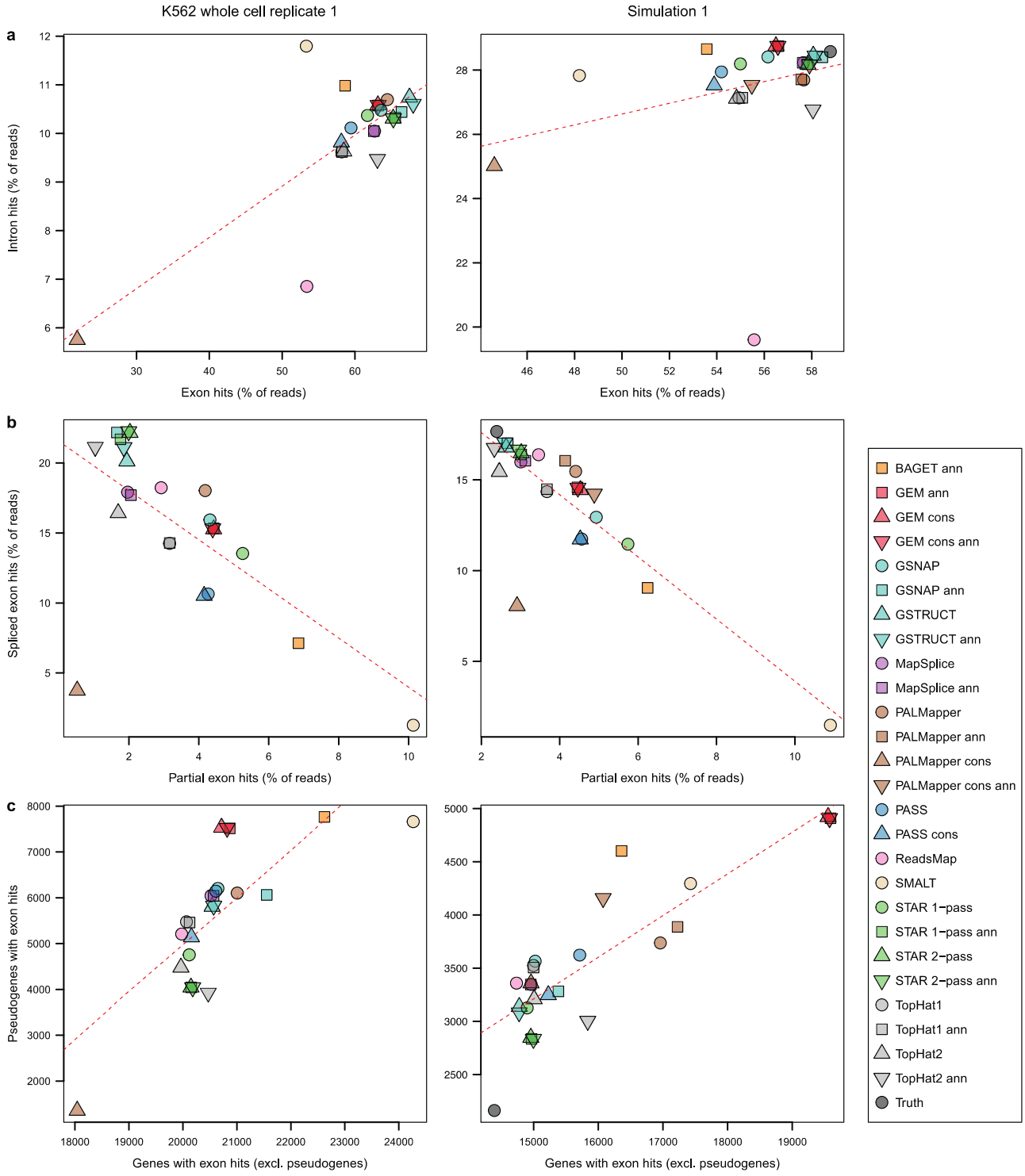
Supplementary Figure 7. Positional distribution of mismatches and gaps over read sequences. Curves show the distribution (percentage) of the indicated operations along the 76 nt read sequences, computed over the primary alignments for K562 whole cell replicate 1. Red lines indicate positions where the frequency exceeds 5%.

Nature Methods: doi:10.1038/nmeth.2722



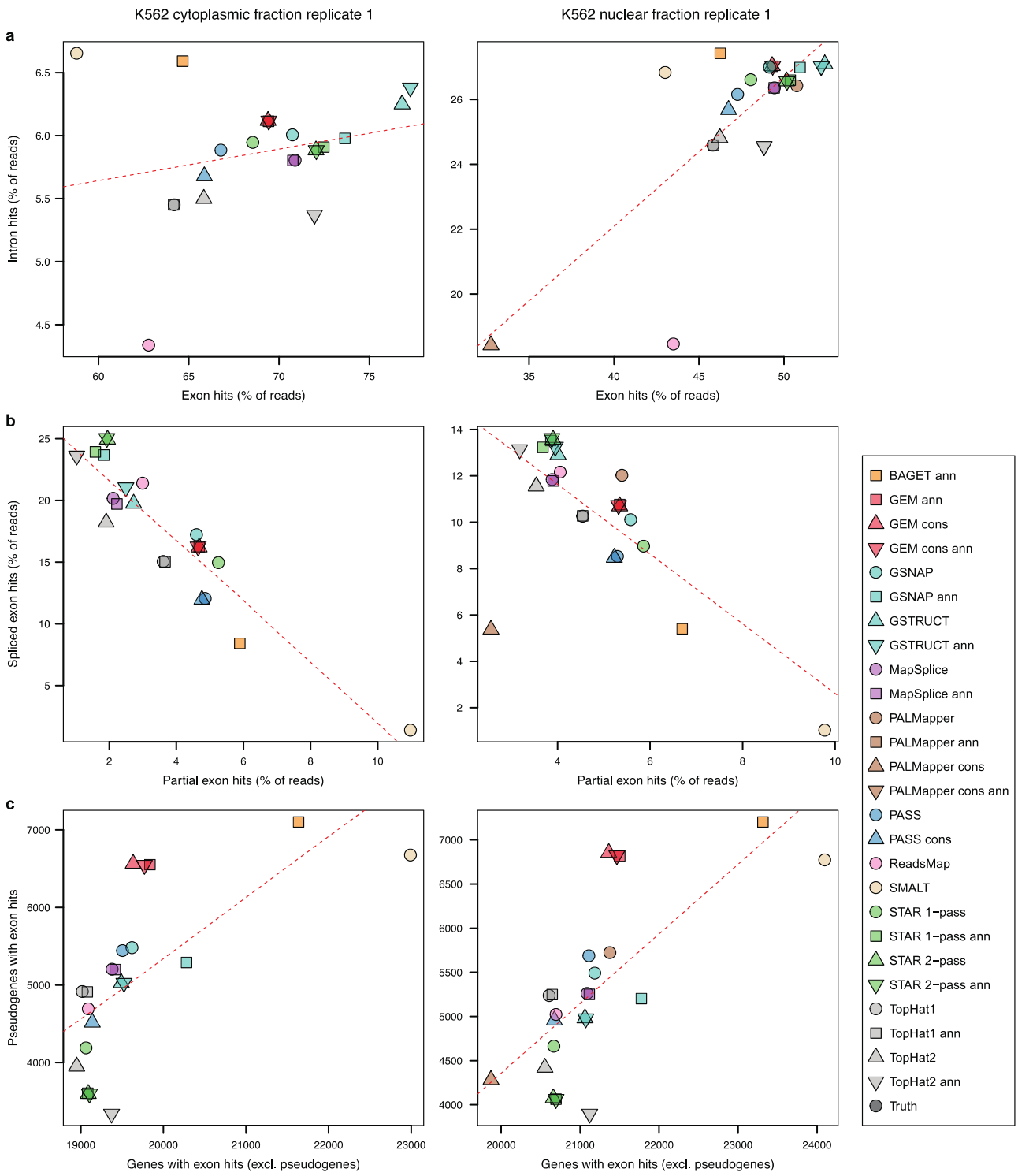
Supplementary Figure 8. Base call quality score distributions for the RNA-seq data sets used in this study.

Nature Methods: doi:10.1038/nmeth.2722



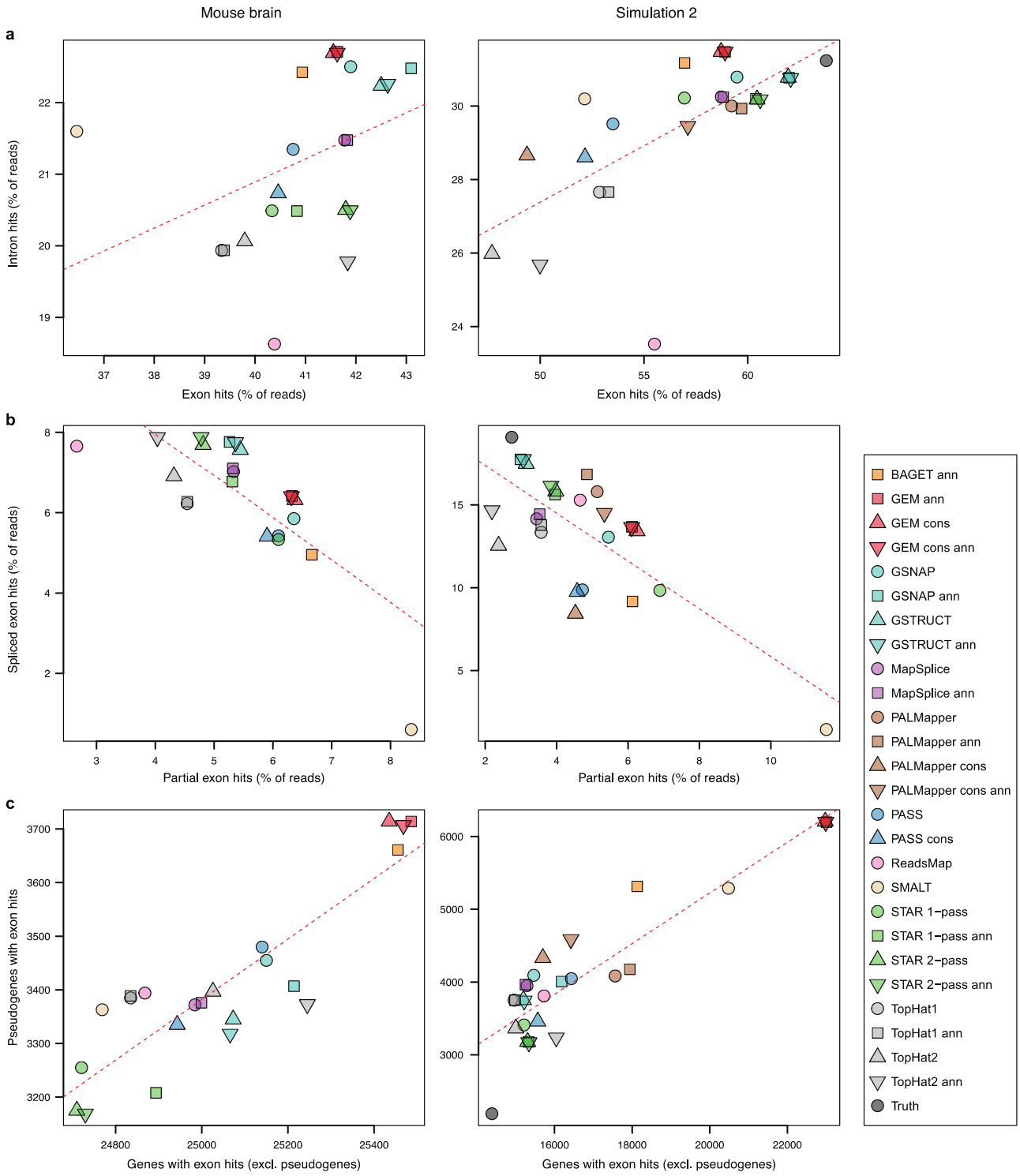
Supplementary Figure 9. Coverage of annotated genes for K562 whole cell and simulation 1. Scatter plots show a range of metrics reflecting coverage of Ensembl genes by RNA-seq read alignments, for K562 whole cell replicate 1 (left) and simulated data set 1 (right). (a) Percentage of sequenced or simulated reads for which all mapped bases fall within exon sequence versus those with all mapped bases confined to introns. (b) Percentage of reads for which mappings partially overlap exons (i.e. alignments where a subset of the genomic positions are annotated as exonic) versus those aligned in a spliced manner with all mapped bases in exon sequence. Note the negative correlation, suggesting that partial exon hits often result from failure to identify splice junctions. (c) Number of genes (including non-coding genes) with fully exonic mappings versus number of pseudogenes with such mappings. For simulated data, "Truth" corresponds to the results expected for a perfect aligner. See also Supplementary Figures 10–12.

Nature Methods: doi:10.1038/nmeth.2722



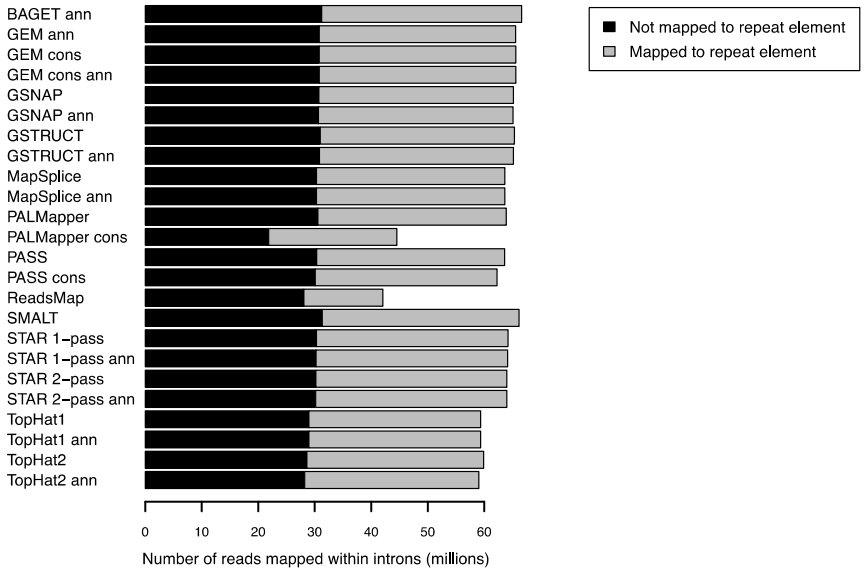
Supplementary Figure 10. Coverage of annotated genes for K562 cytoplasmic and nuclear RNA. Scatter plots show a range of gene coverage metrics as in Supplementary Figure 9.

Nature Methods: doi:10.1038/nmeth.2722



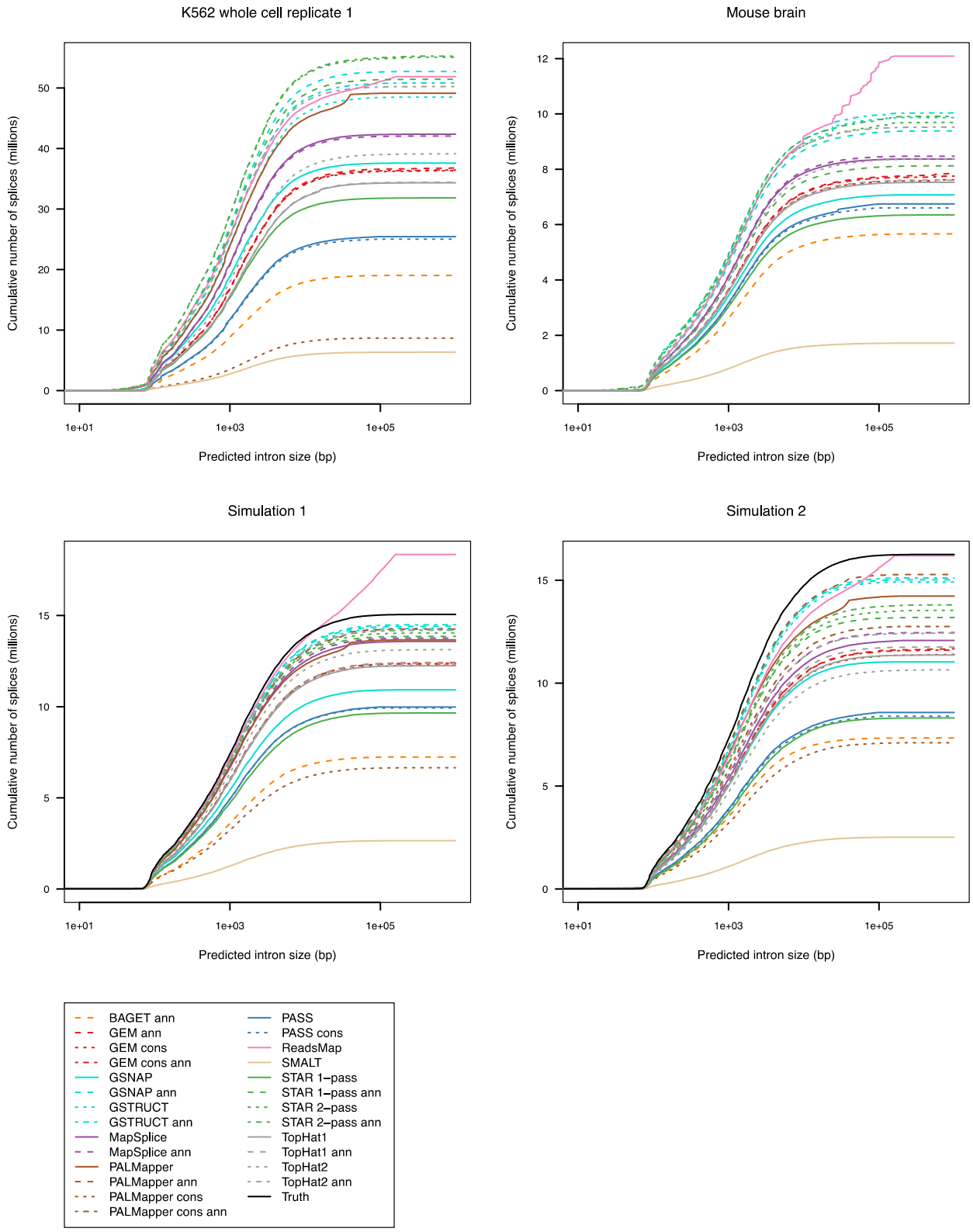
Supplementary Figure 11. Coverage of annotated genes for mouse and simulation 2. Scatter plots show a range of gene coverage metrics as in Supplementary Figure 9.

Nature Methods: doi:10.1038/nmeth.2722



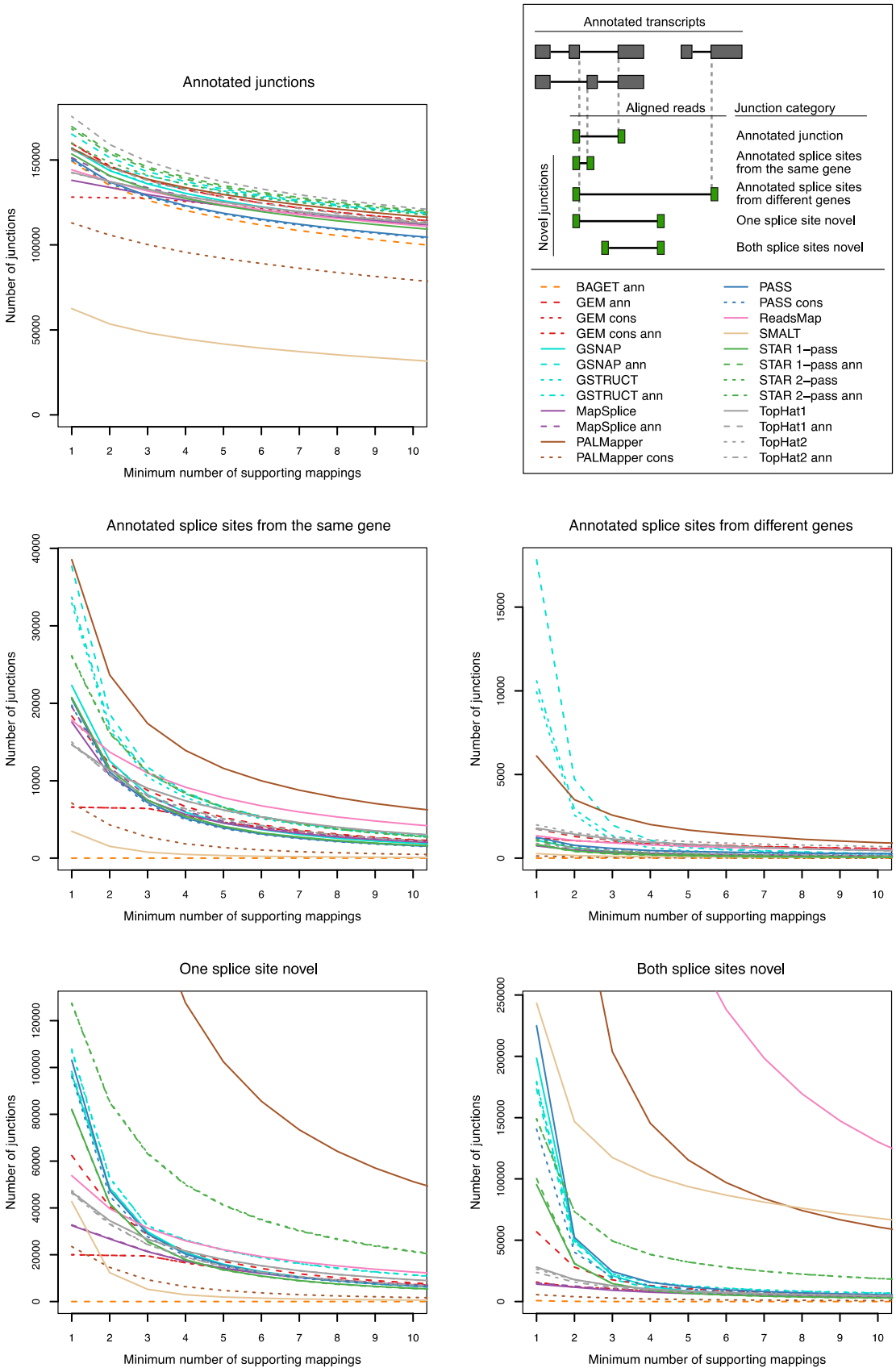
Supplementary Figure 12. Mapping frequency at intronic repeats. Results for K562 nuclear fraction replicate 1 are shown. Grey bar segments indicate the proportion of intronic mappings that overlap with repeat elements. Note the lower proportion of such mappings for ReadsMap.

Nature Methods: doi:10.1038/nmeth.2722



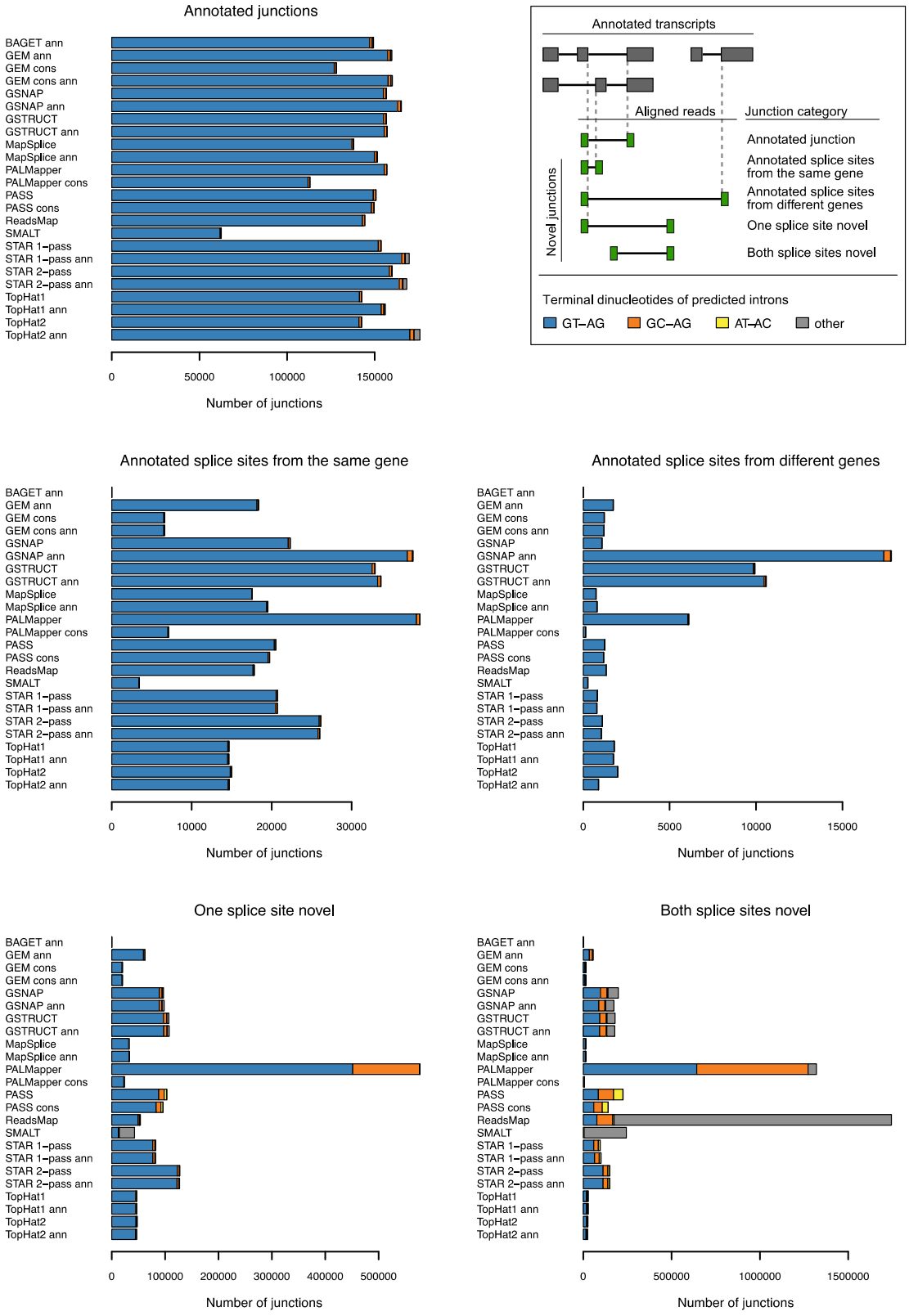
Supplementary Figure 13. Size distribution for splices in primary alignments. Cumulative distributions are shown for each protocol on four data sets. For the two simulated data sets, the true size distribution is also shown (black curves). For PALMapper and ReadsMap, the distributions show an unexpected pattern near the saturation point, suggesting a problem with the scoring of very long splices by these two aligners.

Nature Methods: doi:10.1038/nmeth.2722



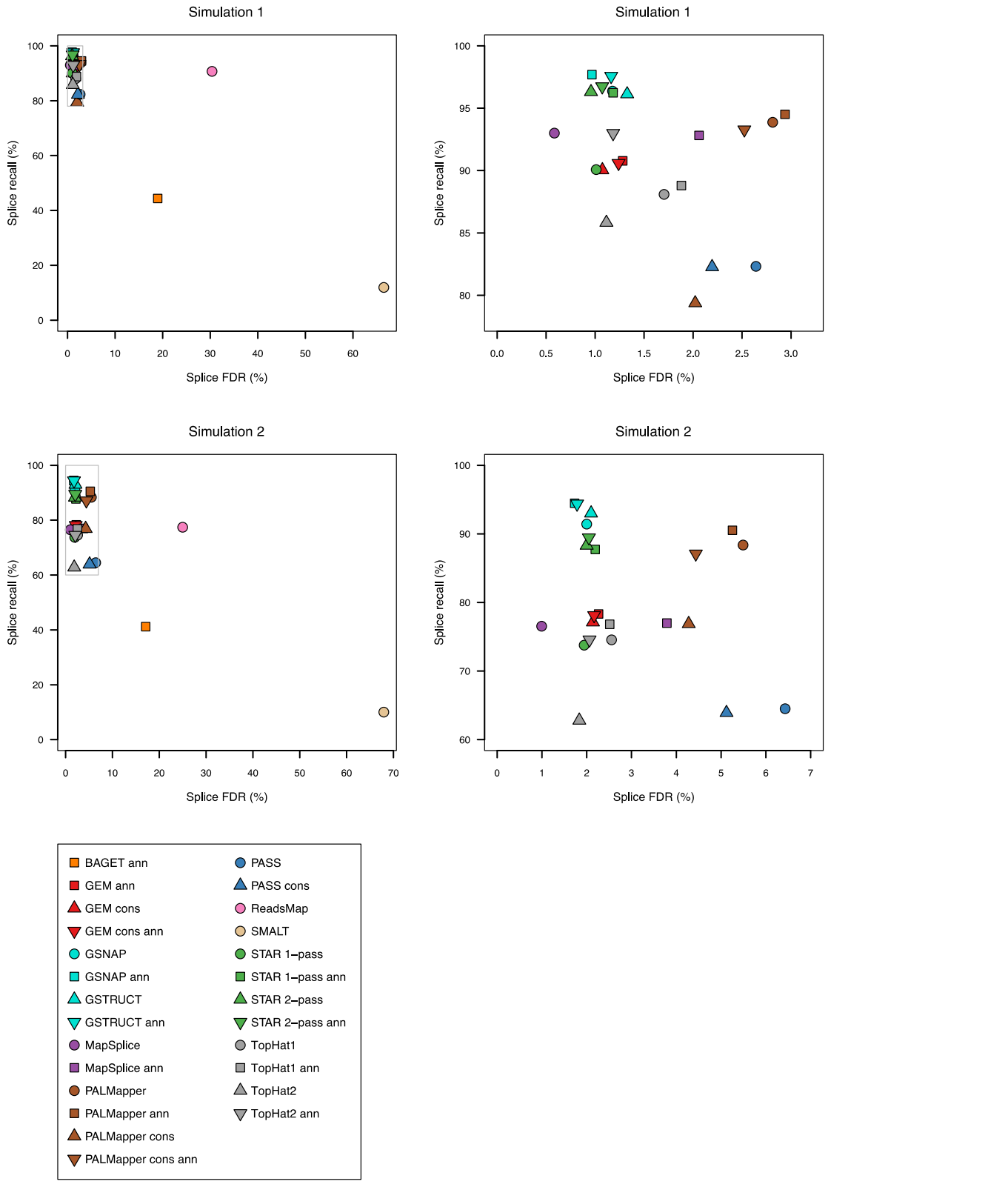
Supplementary Figure 14. Number of supporting alignments for known and novel junctions. Results for K562 whole cell replicate 1 are shown. Curves illustrate the frequency of junctions for different thresholds on the number of supporting primary alignments. Reported junctions were classified into five categories by comparison to junctions annotated in the Ensembl database (see pictogram). Note that known junctions tend to have many supporting alignments (top left plot), while unannotated junctions typically have few (other plots).

Nature Methods: doi:10.1038/nmeth.2722



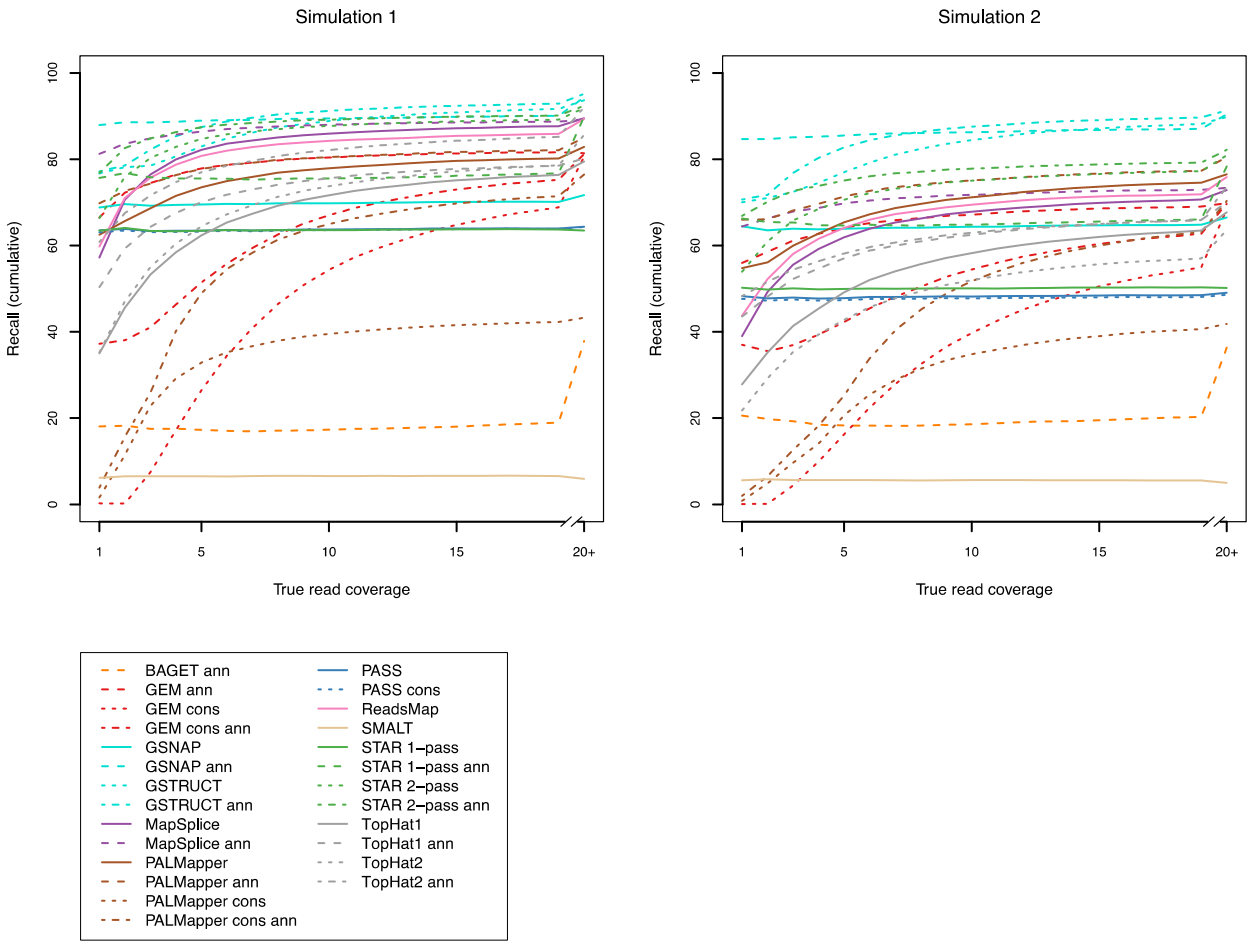
Supplementary Figure 15. Splice signals at known and novel junctions. Results for K562 whole cell replicate 1 are shown. Reported junctions were classified into five categories by comparison to those annotated in the Ensembl database and further stratified according to the first and last dinucleotides of inferred introns (see inset legend). The great majority of known introns begin with GT and end with AG, whereas a small proportion have the sequences GC-AG and AT-AC (see Methods). Directionality was not considered in this analysis (i.e. CT-AC was counted as GT-AG), since RNA-seq data cannot be assumed to be perfectly strand-specific.

Nature Methods: doi:10.1038/nmeth.2722



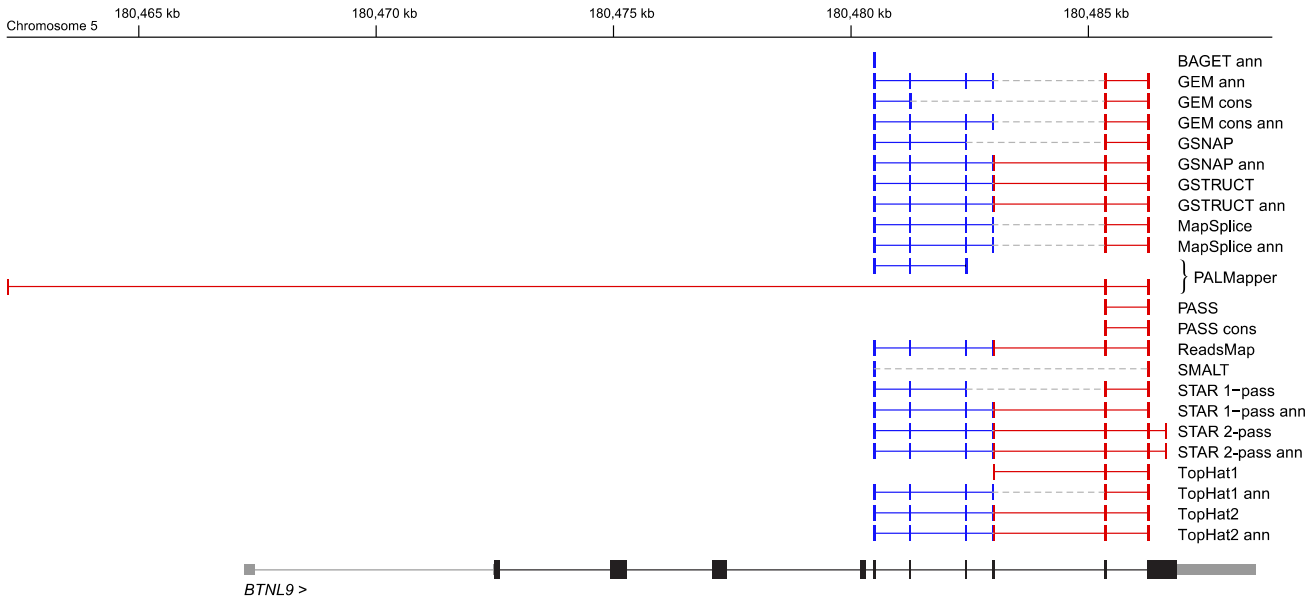
Supplementary Figure 16. Accuracy for anchored splices in primary alignments of simulated reads. Recall and false disovery rate (FDR) is presented for splices located between positions 20 and 57 in the 76 nt reads. Accuracy tends to be higher for this subset of splices compared to those with less flanking sequence (cf. Fig. 5a, where results for all splices are shown). The left plots depict results for all protocols, whereas the right plots show details of the most dense areas (indicated by grey rectangles in the left plots).

Nature Methods: doi:10.1038/nmeth.2722



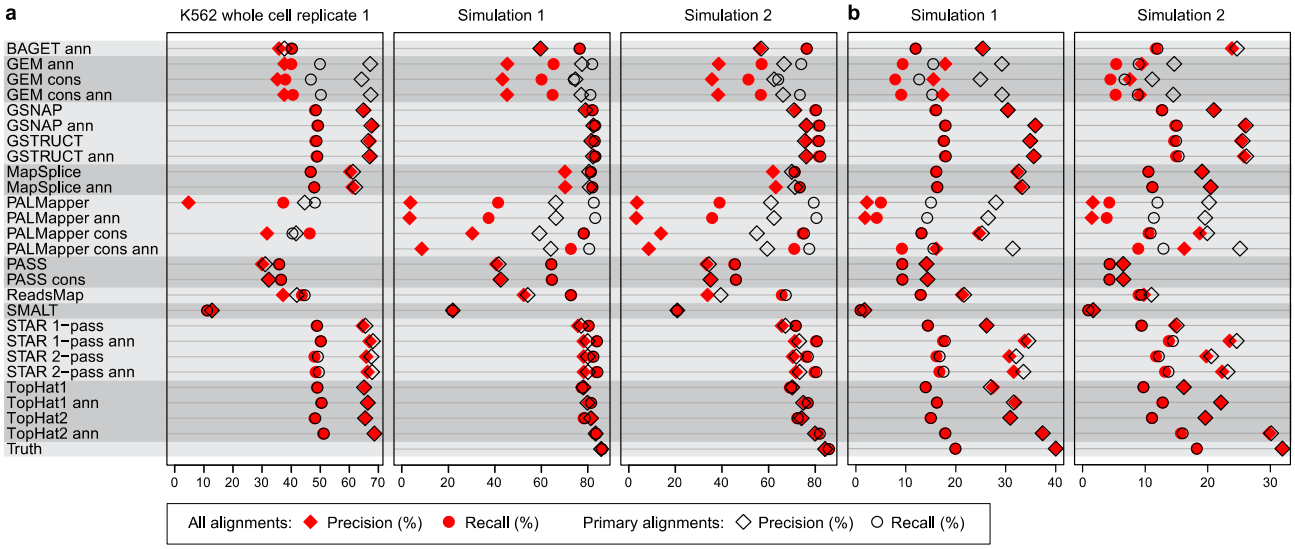
Supplementary Figure 17. Splice recall as a function of true read coverage. Curves depict the cumulative percentage of correctly identified splices as a function of the true number of simulated reads spanning the corresponding exon junctions.

Nature Methods: doi:10.1038/nmeth.2722



Supplementary Figure 18. Examples of alignments with multiple splice junctions. Alignments of a read pair from the K562 data set mapping across six exons of the gene *BTNL9.* The first mate (red) contains two exon junctions, and the second mate (blue) contains three. Paired alignments are connected by dashed lines. All reported primary alignments are shown. *BTLN9* coding sequence is indicated in black and untranslated regions in gray. Nine protocols (GSNAP ann, GSTRUCT, GSTRUCT ann, ReadsMap, STAR 1-pass ann, STAR 2-pass, STAR 2-pass ann, TopHat2 and TopHat2 ann) successfully identified all junctions. However, the STAR 2-pass protocols predicted an additional, most likely errorneous junction separating the first base of mate 1 from the remainder of the read. PASS, PASS cons and TopHat1 only mapped the first mate, whereas BAGET only mapped the second. The PALMapper base protocol produced incompatible alignments of the two mates and the conservative PALMapper protocol did not report alignments for either mate.

Nature Methods: doi:10.1038/nmeth.2722



Supplementary Figure 19. Effect of secondary alignments on transcript assembly by Cufflinks. Performance was assessed by measuring precision and recall for individual exons (a) and spliced transcripts (b), using all alignments from each protocol (red symbols) or the subset of primary alignments (open symbols). For K562 data, precision was defined as the fraction of predicted exons matching Ensembl annotation, and recall as the fraction of annotated exons that were predicted. Only exons from protein-coding genes were considered. Results on simulated data were benchmarked against simulated gene models, using analogous definitions of precision and recall. The last row shows the results obtained when using perfect alignments produced by the simulator (Truth).

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 1. RNA-seq data sets used in this study.

Name	ID	Species	Read pairs	Sequencing lanes
K562 whole cell replicate 1	LID16627	Human (cell line)	113588758	3
K562 whole cell replicate 2	LID16628	Human (cell line)	119053315	3
K562 cytoplasmic fraction replicate 1	LID8465	Human (cell line)	124826068	3
K562 cytoplasmic fraction replicate 2	LID8466	Human (cell line)	88445339	3
K562 nuclear fraction replicate 1	LID8556	Human (cell line)	117113622	3
K562 nuclear fraction replicate 2	LID8557	Human (cell line)	105769104	3
Mouse brain	ERS028664	Mouse strain C57BL/6NJ	57187342	2
Simulation 1	n.a.	Human	40000000	n.a.
Simulation 2	n.a.	Human	40000000	n.a.

n.a., not applicable.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 2. Results on key metrics.

	Mapped reads[a]				Correctly mapped bases[b]		Incorrectly mapped bases[b]		Splice frequency[c]				Junction recall (≥2 mappings)[d]		Junction precision (≥2 mappings)[d]	
	K562	M	S1	S2	S1	S2	S1	S2	K562/1	M	S1	S2	S1	S2	S1	S2
BAGET ann	92.94	95.71	98.58	96.77	90.61	87.49	5.23	4.83	8.38	4.95	9.05	9.17	63.03	61.89	95.56	94.91
GEM ann	**93.87**	**98.33**	**99.90**	**99.40**	96.54	94.33	3.29	4.76	16.23	6.91	15.55	14.62	95.34	90.80	95.60	89.60
GEM cons	93.85	98.31	99.88	99.36	96.49	94.25	3.30	4.80	16.01	6.70	15.35	14.26	84.08	77.39	96.56	91.57
GEM cons ann	93.86	**98.33**	**99.90**	99.39	96.53	94.32	3.29	4.77	16.07	6.81	15.50	14.53	90.14	86.22	96.15	91.51
GSNAP	93.80	96.71	99.24	97.95	96.84	94.55	1.75	2.01	16.55	6.19	13.66	13.79	95.61	95.34	95.58	93.58
GSNAP ann	93.82	96.72	99.25	97.97	97.52	95.27	1.35	1.70	23.21	8.20	18.01	18.78	**98.12**	**97.90**	93.28	91.04
GSTRUCT	**93.87**	97.44	99.26	98.11	96.95	94.85	1.95	2.34	21.35	8.63	17.87	18.65	96.79	96.42	96.95	95.16
GSTRUCT ann	**93.87**	97.43	99.26	98.11	**97.59**	**95.43**	1.31	1.76	22.37	8.77	18.12	18.89	97.24	97.02	97.24	95.51
MapSplice	90.02	93.95	98.61	94.61	96.83	91.46	1.35	1.62	18.65	7.32	16.98	15.09	95.94	90.35	**98.26**	**95.86**
MapSplice ann	90.01	93.98	98.68	94.79	96.95	91.67	1.34	1.64	18.51	7.41	17.20	15.57	97.00	93.54	94.54	90.78
PALMapper	91.15	n.a.	98.35	96.78	95.20	93.03	3.05	3.74	21.62	n.a.	17.09	17.79	94.89	93.14	61.49	58.58
PALMapper ann	n.a.	n.a.	98.42	96.99	94.96	92.99	3.37	4.00	n.a.	n.a.	17.82	19.10	96.27	95.18	58.66	52.07
PALMapper cons	52.14	n.a.	80.81	84.77	78.54	81.91	1.70	2.86	3.82	n.a.	8.31	8.88	87.97	86.59	95.74	91.85
PALMapper cons ann	n.a.	n.a.	97.74	94.32	94.85	90.92	2.78	3.40	n.a.	n.a.	15.44	15.94	92.65	89.47	78.79	71.63
PASS	89.86	92.78	96.97	90.15	90.83	80.52	3.46	3.38	11.20	5.90	12.48	10.72	91.18	85.10	86.33	76.30
PASS cons	87.62	90.29	95.99	87.48	90.47	79.28	3.01	2.80	11.02	5.77	12.42	10.49	91.10	84.94	89.41	80.37
ReadsMap	77.18	72.82	88.00	86.49	77.15	72.65	9.87	13.83	22.84	**10.57**	**22.94**	**20.24**	94.63	89.53	20.68	20.25
SMALT	91.45	92.25	96.73	96.34	91.62	90.13	1.92	2.10	2.80	1.51	3.32	3.15	35.34	34.88	30.69	28.43
STAR 1-pass	91.52	89.23	98.77	96.23	96.20	92.21	1.70	1.96	14.02	5.55	12.07	10.39	93.01	87.24	97.68	95.79
STAR 1-pass ann	91.69	89.26	98.85	96.71	97.19	93.73	1.27	1.60	22.64	7.10	17.32	16.49	96.00	93.23	91.72	89.80
STAR 2-pass	91.68	89.31	98.86	96.77	97.26	93.85	**1.23**	1.58	24.24	8.47	17.55	16.92	96.53	92.38	95.66	92.59
STAR 2-pass ann	91.67	89.34	98.85	96.77	97.26	93.90	1.25	1.59	**24.33**	8.67	17.74	17.25	97.71	95.02	91.66	88.81
TopHat1	84.22	84.92	95.44	86.09	92.79	83.82	2.44	2.27	15.12	6.58	15.31	14.21	91.01	83.85	94.97	92.33
TopHat1 ann	84.25	84.96	95.58	86.53	92.94	84.26	2.45	2.27	15.15	6.65	15.48	14.70	93.59	88.99	94.62	92.15
TopHat2	83.47	85.10	93.96	77.93	91.96	76.18	1.85	1.74	17.23	7.32	16.41	13.31	91.78	86.23	95.04	93.36
TopHat2 ann	84.52	85.41	93.84	79.64	93.16	78.10	1.46	**1.55**	22.11	8.33	17.76	15.54	95.76	92.61	88.40	86.87

Results are based on primary alignments only. Data sets: Mean over K562 samples (K562), K562 whole cell replicate 1 (K562/1), mouse brain (M), simulation 1 (S1) and 2 (S2). Metrics: [a]percentage of sequenced or simulated reads mapped by each protocol; [b]percentage of all simulated bases that were correctly/incorrectly aligned; [c]number of splices in primary alignments divided by the number of sequenced reads; [d]junction discovery accuracy when requiring at least two supporting mappings per junction. All values are given as percentages. Bold indicates the highest or lowest value in each column. The PALMapper protocols were not applied to all data sets, as indicated (n.a.). The lower splice frequencies on mouse data are expected as a result of a more pronounced 3' bias in this data set (not shown).

Supplementary Table 3. Alignment yield.

	Both mates uniquely mapped	Both mates multi-mapped	One mate uniquely and one multi-mapped	One mate uniquely mapped and one unaligned	One mate multi-mapped and one unaligned	Total mapped read pairs	Total mapped reads
A. K562 whole cell replicate 1							
BAGET ann	87.78%	0.13%	0.98%	3.43%	0.24%	92.57%	90.73%
GEM ann	47.13%	42.92%	0.37%	0.77%	0.72%	91.91%	91.17%
GEM cons	47.45%	42.57%	0.37%	0.79%	0.73%	91.91%	91.15%
GEM cons ann	47.38%	42.65%	0.37%	0.78%	0.73%	91.91%	91.16%
GSNAP	79.50%	10.98%	0.04%	0.90%	0.35%	91.77%	91.14%
GSNAP ann	79.61%	10.86%	0.04%	0.92%	0.35%	91.78%	91.15%
GSTRUCT	74.48%	16.01%	0.04%	0.88%	0.39%	91.80%	91.17%
GSTRUCT ann	77.86%	12.63%	0.04%	0.92%	0.35%	91.80%	91.16%
MapSplice	83.31%	0.01%	0.05%	5.81%	0.88%	90.07%	86.72%
MapSplice ann	83.30%	0.01%	0.05%	5.81%	0.89%	90.07%	86.71%
PALMapper	32.84%	36.97%	18.72%	1.50%	1.67%	91.69%	90.11%
PALMapper cons	18.12%	0.00%	0.00%	24.23%	0.00%	42.36%	30.24%
PASS	82.13%	0.33%	0.18%	8.17%	0.05%	90.86%	86.75%
PASS cons	80.95%	0.32%	0.00%	6.65%	0.00%	87.93%	84.60%
ReadsMap	55.49%	4.42%	6.46%	11.17%	1.17%	78.70%	72.54%
SMALT	85.76%	0.03%	1.02%	6.49%	0.38%	93.68%	90.24%
STAR 1-pass	83.76%	5.68%	0.00%	0.00%	0.00%	89.45%	89.45%
STAR 1-pass ann	84.17%	5.45%	0.00%	0.00%	0.00%	89.61%	89.61%
STAR 2-pass	81.75%	7.85%	0.00%	0.00%	0.00%	89.60%	89.60%
STAR 2-pass ann	81.66%	7.93%	0.00%	0.00%	0.00%	89.59%	89.59%
TopHat1	73.35%	4.09%	0.00%	9.65%	1.39%	88.48%	82.96%
TopHat1 ann	73.39%	4.11%	0.00%	9.60%	1.39%	88.49%	82.99%
TopHat2	70.58%	4.49%	0.00%	11.29%	1.58%	87.95%	81.51%
TopHat2 ann	72.57%	4.59%	0.00%	10.63%	1.33%	89.12%	83.14%
B. K562 whole cell replicate 2							
BAGET ann	84.16%	0.14%	1.54%	8.79%	0.51%	95.14%	90.49%
GEM ann	47.12%	41.72%	0.49%	3.09%	2.43%	94.85%	92.09%
GEM cons	47.46%	41.34%	0.49%	3.14%	2.42%	94.85%	92.07%
GEM cons ann	47.40%	41.42%	0.49%	3.12%	2.42%	94.85%	92.08%
GSNAP	78.74%	11.68%	0.05%	2.35%	0.93%	93.75%	92.11%
GSNAP ann	78.86%	11.61%	0.05%	2.32%	0.93%	93.77%	92.14%
GSTRUCT	73.07%	17.45%	0.05%	2.22%	1.03%	93.82%	92.19%
GSTRUCT ann	74.60%	15.91%	0.05%	2.26%	0.99%	93.81%	92.19%
MapSplice	76.45%	0.01%	0.06%	12.23%	2.06%	90.82%	83.68%
MapSplice ann	76.43%	0.01%	0.06%	12.23%	2.07%	90.81%	83.66%
PALMapper	31.12%	35.15%	18.23%	4.72%	4.80%	94.03%	89.27%
PALMapper cons	34.52%	0.00%	0.00%	34.39%	0.00%	68.92%	51.72%
PASS	74.45%	0.32%	0.17%	17.64%	0.13%	92.72%	83.83%
PASS cons	73.19%	0.32%	0.00%	10.36%	0.00%	83.87%	78.69%
SMALT	86.08%	0.02%	0.75%	6.91%	0.21%	93.96%	90.40%
STAR 1-pass	82.90%	5.99%	0.00%	0.00%	0.00%	88.89%	88.89%
STAR 1-pass ann	83.68%	5.68%	0.00%	0.00%	0.00%	89.36%	89.36%
STAR 2-pass	81.25%	8.10%	0.00%	0.00%	0.00%	89.36%	89.36%
STAR 2-pass ann	81.15%	8.20%	0.00%	0.00%	0.00%	89.35%	89.35%
TopHat1	62.54%	3.61%	0.00%	16.54%	2.34%	85.03%	75.59%
TopHat1 ann	62.56%	3.63%	0.00%	16.50%	2.35%	85.04%	75.62%
TopHat2	59.15%	3.97%	0.00%	17.63%	2.37%	83.12%	73.12%
TopHat2 ann	60.74%	3.85%	0.00%	17.39%	2.16%	84.15%	74.37%
C. K562 cytoplasmic fraction replicate 1							
BAGET ann	91.83%	0.11%	1.00%	3.51%	0.29%	96.74%	94.84%
GEM ann	52.24%	42.33%	0.63%	0.72%	0.72%	96.63%	95.91%
GEM cons	52.66%	41.87%	0.63%	0.75%	0.72%	96.63%	95.90%
GEM cons ann	52.57%	41.98%	0.63%	0.74%	0.72%	96.63%	95.90%
GSNAP	82.59%	12.69%	0.12%	0.68%	0.31%	96.39%	95.89%
GSNAP ann	82.53%	12.75%	0.12%	0.69%	0.31%	96.40%	95.90%
GSTRUCT	77.97%	17.34%	0.12%	0.79%	0.31%	96.53%	95.98%
GSTRUCT ann	79.34%	15.97%	0.12%	0.81%	0.30%	96.53%	95.98%
MapSplice	90.31%	0.01%	0.09%	4.29%	0.63%	95.33%	92.87%
MapSplice ann	90.29%	0.01%	0.09%	4.31%	0.63%	95.32%	92.86%
PASS	89.47%	0.19%	0.19%	5.90%	0.03%	95.78%	92.82%
PASS cons	88.33%	0.18%	0.00%	5.51%	0.00%	94.03%	91.27%
ReadsMap	61.37%	5.60%	9.37%	10.00%	1.02%	87.37%	81.86%
SMALT	88.12%	0.00%	0.49%	5.63%	0.14%	94.39%	91.50%
STAR 1-pass	87.75%	5.96%	0.00%	0.00%	0.00%	93.71%	93.71%
STAR 1-pass ann	87.72%	6.12%	0.00%	0.00%	0.00%	93.84%	93.84%
STAR 2-pass	83.73%	10.08%	0.00%	0.00%	0.00%	93.81%	93.81%
STAR 2-pass ann	83.60%	10.20%	0.00%	0.00%	0.00%	93.80%	93.80%
TopHat1	77.44%	4.61%	0.00%	9.24%	1.18%	92.47%	87.26%
TopHat1 ann	77.46%	4.65%	0.00%	9.18%	1.18%	92.48%	87.29%
TopHat2	75.66%	5.96%	0.00%	9.56%	1.25%	92.43%	87.03%
TopHat2 ann	77.35%	6.01%	0.00%	8.96%	1.08%	93.39%	88.37%
D. K562 cytoplasmic fraction replicate 2							
BAGET ann	90.78%	0.12%	1.03%	3.12%	0.25%	95.30%	93.61%
GEM ann	44.72%	49.11%	0.46%	0.51%	0.40%	95.20%	94.74%
GEM cons	45.14%	48.67%	0.46%	0.54%	0.40%	95.20%	94.73%
GEM cons ann	45.05%	48.78%	0.46%	0.52%	0.40%	95.20%	94.74%
GSNAP	83.12%	11.16%	0.11%	0.56%	0.17%	95.12%	94.75%
GSNAP ann	83.11%	11.18%	0.11%	0.57%	0.17%	95.13%	94.76%

Nature Methods: doi:10.1038/nmeth.2722

	Both mates uniquely mapped	Both mates multi-mapped	One mate uniquely and one multi-mapped	One mate uniquely mapped and one unaligned	One mate multi-mapped and one unaligned	Total mapped read pairs	Total mapped reads
GSTRUCT	79.62%	14.68%	0.11%	0.66%	0.17%	95.25%	94.83%
GSTRUCT ann	81.25%	13.06%	0.10%	0.67%	0.16%	95.25%	94.83%
MapSplice	90.66%	0.01%	0.08%	3.31%	0.34%	94.40%	92.58%
MapSplice ann	90.60%	0.01%	0.08%	3.36%	0.34%	94.40%	92.55%
PASS	88.31%	0.20%	0.18%	5.90%	0.03%	94.63%	91.66%
PASS cons	87.05%	0.20%	0.00%	5.63%	0.00%	92.87%	90.05%
SMALT	87.24%	0.00%	0.55%	5.99%	0.17%	93.94%	90.86%
STAR 1-pass	86.92%	5.63%	0.00%	0.00%	0.00%	92.55%	92.55%
STAR 1-pass ann	86.65%	6.02%	0.00%	0.00%	0.00%	92.67%	92.67%
STAR 2-pass	82.97%	9.67%	0.00%	0.00%	0.00%	92.64%	92.64%
STAR 2-pass ann	82.83%	9.79%	0.00%	0.00%	0.00%	92.63%	92.63%
TopHat1	75.71%	4.51%	0.00%	9.96%	1.11%	91.29%	85.75%
TopHat1 ann	75.79%	4.52%	0.00%	9.88%	1.11%	91.30%	85.80%
TopHat2	73.38%	5.91%	0.00%	10.80%	1.20%	91.30%	85.29%
TopHat2 ann	75.16%	6.32%	0.00%	10.00%	1.07%	92.55%	87.02%
E. K562 nuclear fraction replicate 1							
BAGET ann	92.05%	0.25%	1.02%	2.65%	0.40%	96.36%	94.84%
GEM ann	64.76%	29.89%	0.40%	0.72%	0.45%	96.22%	95.63%
GEM cons	65.17%	29.45%	0.40%	0.74%	0.45%	96.22%	95.62%
GEM cons ann	65.11%	29.52%	0.40%	0.73%	0.45%	96.22%	95.62%
GSNAP	87.22%	7.68%	0.06%	0.65%	0.26%	95.87%	95.42%
GSNAP ann	87.25%	7.66%	0.06%	0.65%	0.26%	95.88%	95.43%
GSTRUCT	88.12%	6.84%	0.06%	0.71%	0.21%	95.93%	95.47%
GSTRUCT ann	88.70%	6.25%	0.06%	0.71%	0.21%	95.93%	95.47%
MapSplice	90.43%	0.01%	0.08%	4.10%	0.55%	95.16%	92.84%
MapSplice ann	90.43%	0.01%	0.08%	4.09%	0.55%	95.16%	92.84%
PALMapper	46.25%	24.11%	21.33%	1.87%	2.00%	95.57%	93.64%
PALMapper cons	37.19%	2.26%	3.25%	33.81%	3.40%	79.90%	61.29%
PASS	89.41%	0.39%	0.26%	5.43%	0.04%	95.53%	92.79%
PASS cons	88.22%	0.38%	0.00%	5.23%	0.00%	93.83%	91.22%
ReadsMap	62.17%	2.93%	4.98%	12.99%	1.13%	84.20%	77.14%
SMALT	90.82%	0.01%	0.54%	3.88%	0.17%	95.42%	93.40%
STAR 1-pass	88.94%	4.06%	0.00%	0.00%	0.00%	93.00%	93.00%
STAR 1-pass ann	88.77%	4.31%	0.00%	0.00%	0.00%	93.08%	93.08%
STAR 2-pass	87.00%	6.08%	0.00%	0.00%	0.00%	93.08%	93.08%
STAR 2-pass ann	86.95%	6.13%	0.00%	0.00%	0.00%	93.07%	93.07%
TopHat1	78.26%	3.74%	0.00%	9.99%	1.19%	93.19%	87.59%
TopHat1 ann	78.29%	3.75%	0.00%	9.96%	1.19%	93.19%	87.62%
TopHat2	77.30%	4.10%	0.00%	10.23%	1.26%	92.88%	87.14%
TopHat2 ann	78.12%	3.71%	0.00%	9.97%	1.07%	92.87%	87.35%
F. K562 nuclear fraction replicate 2							
BAGET ann	90.76%	0.19%	0.80%	2.47%	0.34%	94.55%	93.15%
GEM ann	64.61%	28.22%	0.32%	0.66%	0.36%	94.17%	93.66%
GEM cons	64.95%	27.85%	0.32%	0.68%	0.37%	94.17%	93.64%
GEM cons ann	64.89%	27.92%	0.32%	0.67%	0.36%	94.17%	93.65%
GSNAP	86.38%	6.72%	0.05%	0.55%	0.18%	93.88%	93.52%
GSNAP ann	86.39%	6.71%	0.05%	0.55%	0.18%	93.89%	93.52%
GSTRUCT	86.95%	6.19%	0.04%	0.62%	0.15%	93.96%	93.57%
GSTRUCT ann	87.68%	5.46%	0.04%	0.62%	0.15%	93.96%	93.57%
MapSplice	89.59%	0.01%	0.07%	3.22%	0.35%	93.24%	91.46%
MapSplice ann	89.59%	0.01%	0.07%	3.22%	0.35%	93.24%	91.46%
PALMapper	45.82%	22.50%	21.39%	1.80%	1.92%	93.43%	91.57%
PALMapper cons	42.29%	2.43%	3.46%	31.26%	2.97%	82.42%	65.30%
PASS	88.47%	0.43%	0.26%	4.24%	0.04%	93.44%	91.30%
PASS cons	87.31%	0.42%	0.00%	4.37%	0.00%	92.09%	89.91%
SMALT	89.36%	0.01%	0.55%	4.46%	0.23%	94.61%	92.26%
STAR 1-pass	87.62%	3.89%	0.00%	0.00%	0.00%	91.50%	91.50%
STAR 1-pass ann	87.34%	4.23%	0.00%	0.00%	0.00%	91.57%	91.57%
STAR 2-pass	85.67%	5.91%	0.00%	0.00%	0.00%	91.57%	91.57%
STAR 2-pass ann	85.61%	5.96%	0.00%	0.00%	0.00%	91.57%	91.57%
TopHat1	78.17%	3.07%	0.00%	8.95%	0.91%	91.10%	86.17%
TopHat1 ann	78.20%	3.08%	0.00%	8.92%	0.91%	91.10%	86.19%
TopHat2	78.66%	3.68%	0.00%	7.91%	0.87%	91.12%	86.73%
TopHat2 ann	79.32%	3.34%	0.00%	7.64%	0.72%	91.03%	86.85%
G. Mouse brain							
BAGET ann	90.34%	0.28%	1.81%	5.87%	0.67%	98.98%	95.71%
GEM ann	62.53%	31.64%	2.89%	0.42%	2.12%	99.60%	98.33%
GEM cons	62.80%	31.33%	2.89%	0.45%	2.12%	99.60%	98.31%
GEM cons ann	62.72%	31.44%	2.89%	0.43%	2.11%	99.60%	98.33%
GSNAP	83.92%	9.54%	1.51%	1.46%	2.01%	98.45%	96.71%
GSNAP ann	83.88%	9.59%	1.51%	1.46%	2.01%	98.45%	96.72%
GSTRUCT	81.63%	13.29%	1.23%	1.00%	1.56%	98.71%	97.44%
GSTRUCT ann	81.94%	13.00%	1.20%	1.01%	1.56%	98.71%	97.43%
MapSplice	88.42%	0.24%	1.63%	5.89%	1.42%	97.60%	93.95%
MapSplice ann	88.49%	0.24%	1.63%	5.81%	1.43%	97.60%	93.98%
PASS	87.38%	0.31%	0.33%	9.48%	0.04%	97.54%	92.78%
PASS cons	84.99%	0.27%	0.00%	10.07%	0.00%	95.33%	90.29%
ReadsMap	57.26%	3.68%	3.20%	16.36%	0.99%	81.50%	72.82%
SMALT	88.66%	0.01%	0.86%	5.27%	0.18%	94.97%	92.25%
STAR 1-pass	84.28%	4.95%	0.00%	0.00%	0.00%	89.23%	89.23%
STAR 1-pass ann	83.98%	5.28%	0.00%	0.00%	0.00%	89.26%	89.26%
STAR 2-pass	83.23%	6.08%	0.00%	0.00%	0.00%	89.31%	89.31%

Nature Methods: doi:10.1038/nmeth.2722

	Both mates uniquely mapped	Both mates multi-mapped	One mate uniquely and one multi-mapped	One mate uniquely mapped and one unaligned	One mate multi-mapped and one unaligned	Total mapped read pairs	Total mapped reads
STAR 2-pass ann	83.26%	6.07%	0.00%	0.00%	0.00%	89.34%	89.34%
TopHat1	75.09%	2.68%	0.00%	11.08%	3.21%	92.06%	84.92%
TopHat1 ann	75.16%	2.70%	0.00%	11.00%	3.21%	92.07%	84.96%
TopHat2	74.51%	4.14%	0.00%	10.51%	2.38%	91.54%	85.10%
TopHat2 ann	76.35%	2.71%	0.00%	10.52%	2.18%	91.75%	85.41%
H. Simulation 1							
BAGET ann	96.37%	0.12%	0.95%	2.08%	0.18%	99.71%	98.58%
GEM ann	67.92%	31.68%	0.20%	0.11%	0.08%	100.00%	99.90%
GEM cons	68.18%	31.38%	0.20%	0.15%	0.10%	100.00%	99.88%
GEM cons ann	68.04%	31.55%	0.20%	0.12%	0.08%	100.00%	99.90%
GSNAP	94.59%	4.54%	0.00%	0.18%	0.05%	99.35%	99.24%
GSNAP ann	94.65%	4.49%	0.00%	0.19%	0.05%	99.37%	99.25%
GSTRUCT	94.54%	4.60%	0.00%	0.20%	0.04%	99.38%	99.26%
GSTRUCT ann	95.37%	3.77%	0.00%	0.20%	0.04%	99.38%	99.26%
MapSplice	95.80%	2.06%	0.01%	1.38%	0.08%	99.34%	98.61%
MapSplice ann	95.95%	2.06%	0.01%	1.24%	0.08%	99.34%	98.68%
PALMapper	51.06%	22.92%	23.26%	1.30%	0.91%	99.46%	98.35%
PALMapper ann	49.88%	23.48%	24.02%	1.21%	0.88%	99.46%	98.42%
PALMapper cons	57.35%	3.89%	7.10%	22.26%	2.67%	93.27%	80.81%
PALMapper cons ann	62.61%	16.14%	17.49%	2.22%	0.78%	99.25%	97.74%
PASS	94.53%	0.44%	0.23%	3.52%	0.02%	98.73%	96.97%
PASS cons	93.82%	0.44%	0.00%	3.46%	0.00%	97.72%	95.99%
ReadsMap	75.90%	2.17%	4.29%	10.83%	0.45%	93.64%	88.00%
SMALT	95.79%	0.01%	0.25%	1.30%	0.04%	97.39%	96.73%
STAR 1-pass	95.97%	2.80%	0.00%	0.00%	0.00%	98.77%	98.77%
STAR 1-pass ann	95.44%	3.41%	0.00%	0.00%	0.00%	98.85%	98.85%
STAR 2-pass	95.36%	3.50%	0.00%	0.00%	0.00%	98.86%	98.86%
STAR 2-pass ann	95.18%	3.67%	0.00%	0.00%	0.00%	98.85%	98.85%
TopHat1	90.80%	1.98%	0.00%	5.04%	0.27%	98.10%	95.44%
TopHat1 ann	91.05%	2.00%	0.00%	4.78%	0.27%	98.10%	95.58%
TopHat2	88.00%	2.46%	0.00%	6.64%	0.36%	97.46%	93.96%
TopHat2 ann	88.38%	2.45%	0.00%	5.77%	0.26%	96.85%	93.84%
I. Simulation 2							
BAGET ann	91.36%	0.35%	2.47%	4.66%	0.51%	99.36%	96.77%
GEM ann	71.15%	27.08%	0.58%	0.74%	0.44%	99.99%	99.40%
GEM cons	71.76%	26.38%	0.59%	0.81%	0.45%	99.99%	99.36%
GEM cons ann	71.50%	26.72%	0.59%	0.75%	0.44%	99.99%	99.39%
GSNAP	93.95%	3.60%	0.01%	0.65%	0.14%	98.35%	97.95%
GSNAP ann	93.97%	3.58%	0.01%	0.68%	0.14%	98.39%	97.97%
GSTRUCT	94.11%	3.57%	0.01%	0.71%	0.12%	98.52%	98.11%
GSTRUCT ann	94.82%	2.87%	0.01%	0.72%	0.11%	98.52%	98.11%
MapSplice	89.26%	1.75%	0.02%	6.88%	0.26%	98.19%	94.61%
MapSplice ann	89.59%	1.74%	0.03%	6.61%	0.25%	98.21%	94.79%
PALMapper	47.73%	19.09%	27.68%	2.70%	1.87%	99.06%	96.78%
PALMapper ann	44.90%	20.84%	29.17%	2.37%	1.80%	99.08%	96.99%
PALMapper cons	56.73%	5.78%	12.15%	16.52%	3.70%	94.88%	84.77%
PALMapper cons ann	58.91%	10.22%	21.30%	5.85%	1.92%	98.21%	94.32%
PASS	83.60%	0.39%	0.29%	11.70%	0.05%	96.03%	90.15%
PASS cons	82.52%	0.38%	0.00%	9.15%	0.00%	92.06%	87.48%
ReadsMap	73.71%	2.06%	2.99%	14.63%	0.81%	94.21%	86.49%
SMALT	94.92%	0.01%	0.48%	1.82%	0.04%	97.27%	96.34%
STAR 1-pass	93.36%	2.87%	0.00%	0.00%	0.00%	96.23%	96.23%
STAR 1-pass ann	93.33%	3.38%	0.00%	0.00%	0.00%	96.71%	96.71%
STAR 2-pass	93.24%	3.53%	0.00%	0.00%	0.00%	96.77%	96.77%
STAR 2-pass ann	93.09%	3.69%	0.00%	0.00%	0.00%	96.77%	96.77%
TopHat1	75.36%	1.71%	0.00%	17.25%	0.80%	95.11%	86.09%
TopHat1 ann	76.09%	1.74%	0.00%	16.62%	0.79%	95.24%	86.53%
TopHat2	63.27%	1.88%	0.00%	24.52%	1.03%	90.70%	77.93%
TopHat2 ann	65.70%	2.11%	0.00%	22.70%	0.98%	91.48%	79.64%

Percentage of sequenced or simulated read pairs mapped by each protocol, for the data sets used in this study. Read pairs are classified by the number of alignments reported per mate. These results are also shown graphically in Figure 1.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 4. Accuracy among unique and ambiguous mappings of simulated reads.

	Uniquely mapped reads	Multi-mapped reads	Proportion of unique mappings that are perfect	Proportion of multi-mapped reads for which the primary alignment is perfect	Proportion correctly aligned bases for unique mappings	Proportion correctly aligned bases for primary alignments of multi-mapped reads
A. Simulation 1						
BAGET ann	97.75%	0.74%	86.99%	0.00%	94.58%	81.61%
GEM ann	70.79%	29.11%	96.99%	84.92%	99.56%	89.76%
GEM cons	71.08%	28.79%	96.84%	84.91%	99.53%	89.68%
GEM cons ann	70.93%	28.97%	96.95%	84.87%	99.56%	89.71%
GSNAP	95.65%	3.57%	86.76%	53.56%	99.61%	61.17%
GSNAP ann	95.72%	3.52%	90.94%	57.60%	99.92%	63.45%
GSTRUCT	95.70%	3.55%	91.36%	45.81%	99.80%	50.15%
GSTRUCT ann	96.59%	2.66%	91.39%	51.99%	99.81%	57.11%
MapSplice	96.70%	1.85%	97.56%	47.55%	99.59%	48.24%
MapSplice ann	96.79%	1.84%	97.41%	47.35%	99.60%	48.09%
PALMapper	68.50%	29.75%	99.03%	78.85%	99.86%	90.06%
PALMapper ann	67.57%	30.75%	99.42%	78.30%	99.91%	89.23%
PALMapper cons	73.43%	6.81%	98.59%	72.65%	99.80%	77.15%
PALMapper cons ann	77.83%	19.80%	98.01%	69.56%	99.86%	86.54%
PASS	96.24%	0.61%	49.99%	17.00%	96.71%	35.62%
PASS cons	95.35%	0.48%	50.42%	21.36%	97.04%	43.92%
ReadsMap	83.60%	3.42%	89.98%	34.43%	90.85%	35.26%
SMALT	96.72%	0.17%	75.65%	0.00%	97.98%	68.26%
STAR 1-pass	96.14%	2.58%	88.01%	46.40%	99.51%	52.08%
STAR 1-pass ann	95.56%	3.25%	92.96%	55.74%	99.83%	65.83%
STAR 2-pass	95.48%	3.34%	93.32%	58.30%	99.83%	67.75%
STAR 2-pass ann	95.29%	3.53%	93.41%	58.83%	99.84%	68.81%
TopHat1	93.37%	1.86%	96.52%	40.24%	98.53%	42.38%
TopHat1 ann	93.51%	1.88%	96.47%	40.64%	98.53%	43.01%
TopHat2	91.38%	2.43%	98.93%	43.04%	99.41%	46.14%
TopHat2 ann	92.00%	2.62%	99.29%	49.48%	99.67%	55.92%
B. Simulation 2						
BAGET ann	94.93%	1.84%	84.65%	0.00%	94.84%	83.16%
GEM ann	71.85%	27.55%	92.36%	75.75%	98.83%	85.69%
GEM cons	72.50%	26.85%	92.09%	75.59%	98.79%	85.33%
GEM cons ann	72.21%	27.18%	92.29%	75.64%	98.82%	85.51%
GSNAP	94.28%	3.67%	75.25%	43.56%	99.49%	57.32%
GSNAP ann	94.31%	3.66%	79.16%	46.14%	99.80%	58.16%
GSTRUCT	94.47%	3.64%	80.51%	37.28%	99.55%	46.34%
GSTRUCT ann	95.18%	2.93%	80.64%	41.23%	99.62%	51.29%
MapSplice	92.72%	1.90%	92.82%	46.45%	99.26%	49.13%
MapSplice ann	92.90%	1.88%	92.70%	46.31%	99.24%	48.97%
PALMapper	62.96%	33.82%	97.57%	70.30%	99.59%	89.69%
PALMapper ann	60.71%	36.28%	98.61%	71.06%	99.74%	89.41%
PALMapper cons	71.11%	13.66%	96.29%	59.52%	99.52%	81.57%
PALMapper cons ann	72.53%	21.79%	97.15%	55.24%	99.74%	85.28%
PASS	89.59%	0.56%	28.98%	8.66%	96.35%	33.04%
PASS cons	87.10%	0.38%	29.79%	12.13%	96.81%	43.88%
ReadsMap	82.51%	3.97%	84.58%	27.13%	86.71%	27.99%
SMALT	96.07%	0.26%	67.19%	0.00%	97.76%	73.00%
STAR 1-pass	93.36%	2.87%	77.71%	40.46%	99.32%	52.22%
STAR 1-pass ann	93.33%	3.38%	82.00%	46.36%	99.61%	62.46%
STAR 2-pass	93.24%	3.53%	82.37%	49.17%	99.57%	65.70%
STAR 2-pass ann	93.08%	3.69%	82.56%	49.36%	99.61%	65.88%
TopHat1	83.98%	2.11%	96.38%	39.08%	98.72%	43.37%
TopHat1 ann	84.40%	2.13%	96.36%	39.47%	98.73%	44.02%
TopHat2	75.53%	2.39%	98.39%	40.69%	99.47%	43.87%
TopHat2 ann	77.05%	2.60%	98.56%	45.98%	99.58%	52.60%

Results are shown for simulated reads from the nuclear genome. The percentages in the first two columns are relative to the total number of such reads, whereas the values in subsequent columns are relative to the number of unique or ambiguous mappings (or mapped bases) from each protocol.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 5. Mapping accuracy for simulated data (all reads).

	Uniquely mapped reads						All reads (primary alignment counted)					
	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases
A. Simulation 1												
BAGET ann	97.75%	85.04%	7.63%	0.02%	90.41%	5.18%	98.49%	85.04%	8.24%	0.02%	90.61%	5.23%
GEM ann	70.79%	68.66%	1.98%	0.00%	70.45%	0.31%	99.90%	93.38%	3.53%	0.01%	96.54%	3.29%
GEM cons	71.08%	68.84%	2.09%	0.00%	70.71%	0.33%	99.87%	93.28%	3.61%	0.01%	96.49%	3.30%
GEM cons ann	70.93%	68.76%	2.01%	0.00%	70.57%	0.31%	99.89%	93.35%	3.56%	0.01%	96.53%	3.29%
GSNAP	95.65%	82.99%	12.39%	0.01%	94.68%	0.37%	99.23%	84.90%	12.66%	0.01%	96.84%	1.75%
GSNAP ann	95.72%	87.04%	8.62%	0.01%	95.30%	0.07%	99.24%	89.07%	8.82%	0.01%	97.52%	1.35%
GSTRUCT	95.70%	87.43%	8.08%	0.01%	95.18%	0.19%	99.24%	89.05%	8.24%	0.01%	96.95%	1.95%
GSTRUCT ann	96.59%	88.27%	8.14%	0.01%	96.08%	0.18%	99.24%	89.65%	8.28%	0.01%	97.59%	1.31%
MapSplice	96.70%	94.34%	1.98%	0.01%	95.94%	0.40%	98.55%	95.22%	1.99%	0.02%	96.83%	1.35%
MapSplice ann	96.79%	94.28%	2.15%	0.01%	96.07%	0.39%	98.63%	95.16%	2.16%	0.02%	96.95%	1.34%
PALMapper	68.50%	67.84%	0.63%	0.00%	68.41%	0.09%	98.25%	91.30%	4.20%	0.02%	95.20%	3.05%
PALMapper ann	67.57%	67.18%	0.37%	0.00%	67.51%	0.06%	98.33%	91.26%	3.94%	0.02%	94.96%	3.37%
PALMapper cons	73.43%	72.39%	0.92%	0.00%	73.28%	0.15%	80.24%	77.34%	1.24%	0.01%	78.54%	1.70%
PALMapper cons ann	77.83%	76.28%	1.51%	0.00%	77.72%	0.11%	97.62%	90.05%	5.10%	0.01%	94.85%	2.78%
PASS	96.24%	48.11%	45.10%	0.02%	90.62%	3.08%	96.85%	48.21%	45.21%	0.02%	90.83%	3.46%
PASS cons	95.35%	48.08%	44.60%	0.02%	90.27%	2.75%	95.83%	48.18%	44.71%	0.02%	90.47%	3.01%
ReadsMap	83.60%	75.22%	0.82%	3.90%	75.95%	7.65%	87.02%	76.40%	0.86%	3.95%	77.15%	9.87%
SMALT	96.72%	73.17%	21.80%	0.00%	91.56%	1.89%	96.89%	73.17%	21.90%	0.00%	91.62%	1.92%
STAR 1-pass	96.14%	84.61%	11.20%	0.00%	94.87%	0.47%	98.72%	85.81%	11.35%	0.01%	96.20%	1.70%
STAR 1-pass ann	95.56%	88.83%	6.60%	0.00%	95.06%	0.16%	98.81%	90.64%	6.94%	0.01%	97.19%	1.27%
STAR 2-pass	95.48%	89.11%	6.24%	0.00%	95.01%	0.16%	98.82%	91.05%	6.57%	0.01%	97.26%	1.23%
STAR 2-pass ann	95.29%	89.00%	6.16%	0.00%	94.84%	0.15%	98.81%	91.08%	6.52%	0.01%	97.26%	1.25%
TopHat1	93.37%	90.12%	1.96%	0.02%	92.00%	1.37%	95.23%	90.87%	2.00%	0.02%	92.79%	2.44%
TopHat1 ann	93.51%	90.21%	2.00%	0.02%	92.13%	1.37%	95.39%	90.97%	2.05%	0.02%	92.94%	2.45%
TopHat2	91.38%	90.41%	0.46%	0.01%	90.84%	0.54%	93.81%	91.45%	0.56%	0.02%	91.96%	1.85%
TopHat2 ann	92.00%	91.35%	0.36%	0.01%	91.69%	0.31%	94.62%	92.64%	0.54%	0.02%	93.16%	1.46%
B. Simulation 2												
BAGET ann	94.93%	80.35%	9.77%	0.01%	86.98%	4.73%	96.77%	80.35%	11.31%	0.01%	87.49%	4.83%
GEM ann	71.85%	66.36%	5.21%	0.01%	70.84%	0.84%	99.40%	87.23%	8.29%	0.02%	94.33%	4.76%
GEM cons	72.50%	66.77%	5.45%	0.01%	71.44%	0.88%	99.36%	87.06%	8.40%	0.02%	94.25%	4.80%
GEM cons ann	72.21%	66.64%	5.28%	0.01%	71.19%	0.85%	99.39%	87.21%	8.30%	0.02%	94.32%	4.77%
GSNAP	94.28%	70.94%	22.95%	0.01%	92.48%	0.47%	97.95%	72.54%	23.45%	0.01%	94.55%	2.01%
GSNAP ann	94.31%	74.66%	19.48%	0.01%	93.18%	0.19%	97.97%	76.35%	19.91%	0.01%	95.27%	1.70%
GSTRUCT	94.47%	76.06%	18.01%	0.01%	93.18%	0.42%	98.11%	77.42%	18.33%	0.01%	94.85%	2.34%
GSTRUCT ann	95.18%	76.75%	18.08%	0.01%	93.95%	0.36%	98.11%	77.96%	18.37%	0.01%	95.43%	1.76%
MapSplice	92.72%	86.05%	6.00%	0.01%	90.55%	0.68%	94.61%	86.94%	6.05%	0.02%	91.46%	1.62%
MapSplice ann	92.90%	86.13%	6.11%	0.01%	90.77%	0.70%	94.78%	87.00%	6.16%	0.01%	91.67%	1.64%
PALMapper	62.96%	61.43%	1.43%	0.00%	62.70%	0.26%	96.78%	85.21%	8.47%	0.02%	93.03%	3.74%
PALMapper ann	60.71%	59.87%	0.77%	0.00%	60.55%	0.16%	96.99%	85.65%	7.83%	0.02%	92.99%	4.00%
PALMapper cons	71.11%	68.47%	2.41%	0.00%	70.77%	0.34%	84.76%	76.60%	5.55%	0.02%	81.91%	2.86%
PALMapper cons ann	72.53%	70.46%	1.96%	0.00%	72.34%	0.19%	94.32%	82.50%	8.88%	0.02%	90.92%	3.40%
PASS	89.59%	25.96%	60.51%	0.02%	80.35%	3.04%	90.15%	26.01%	60.64%	0.02%	80.52%	3.38%
PASS cons	87.10%	25.94%	58.58%	0.02%	79.12%	2.60%	87.48%	25.99%	58.71%	0.02%	79.28%	2.80%
ReadsMap	82.51%	69.79%	2.02%	7.49%	71.54%	10.97%	86.48%	70.87%	2.06%	7.59%	72.65%	13.83%
SMALT	96.07%	64.55%	29.38%	0.00%	90.04%	2.07%	96.34%	64.55%	29.55%	0.00%	90.13%	2.10%
STAR 1-pass	93.36%	72.55%	20.39%	0.00%	90.75%	0.62%	96.23%	73.72%	20.74%	0.01%	92.21%	1.96%
STAR 1-pass ann	93.33%	76.53%	16.55%	0.00%	91.66%	0.36%	96.71%	78.10%	17.11%	0.01%	93.73%	1.60%
STAR 2-pass	93.24%	76.80%	16.14%	0.00%	91.58%	0.39%	96.77%	78.54%	16.74%	0.01%	93.85%	1.58%
STAR 2-pass ann	93.08%	76.85%	15.98%	0.00%	91.51%	0.35%	96.77%	78.67%	16.61%	0.01%	93.90%	1.59%
TopHat1	83.98%	80.94%	2.04%	0.01%	82.90%	1.08%	86.09%	81.76%	2.14%	0.01%	83.82%	2.27%
TopHat1 ann	84.40%	81.32%	2.08%	0.01%	83.32%	1.07%	86.53%	82.16%	2.19%	0.01%	84.26%	2.27%
TopHat2	75.53%	74.31%	0.87%	0.01%	75.13%	0.40%	77.92%	75.29%	0.97%	0.01%	76.18%	1.74%
TopHat2 ann	77.05%	75.94%	0.83%	0.01%	76.73%	0.32%	79.65%	77.14%	1.02%	0.01%	78.10%	1.55%

Results are shown for simulated reads from the nuclear genome, and percentages are relative to the total number of such reads. Perfectly mapped reads have all 76 bases correctly placed (accounting for ambiguity in indel placement as described in Methods). Part correctly mapped reads have at least one base correctly placed, but not all 76. Reads mapped near the correct location are those for which no base is correctly placed, but the mapping overlaps with the correct mapping (this may occur in repetitive regions or indicate a bug in the aligner, as for ReadsMap).

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 6. Mapping accuracy for simulated data (spliced reads).

	Uniquely mapped reads						All reads (primary alignment counted)					
	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases
A. Simulation 1												
BAGET ann	91.73%	39.23%	35.43%	0.01%	64.14%	17.83%	94.59%	39.23%	37.95%	0.01%	64.96%	17.96%
GEM ann	21.58%	13.33%	7.71%	0.01%	20.19%	1.25%	99.52%	80.46%	14.78%	0.01%	93.57%	5.64%
GEM cons	22.39%	13.58%	8.23%	0.01%	20.84%	1.37%	99.38%	79.94%	15.16%	0.01%	93.31%	5.72%
GEM cons ann	21.93%	13.54%	7.84%	0.01%	20.50%	1.27%	99.49%	80.30%	14.91%	0.01%	93.50%	5.67%
GSNAP	96.44%	64.51%	31.00%	0.00%	93.21%	1.44%	99.31%	65.51%	31.61%	0.00%	94.78%	2.68%
GSNAP ann	96.82%	85.14%	11.58%	0.00%	96.20%	0.18%	99.36%	86.57%	11.81%	0.00%	97.84%	1.07%
GSTRUCT	95.09%	84.77%	10.13%	0.00%	94.45%	0.26%	99.38%	86.57%	10.35%	0.00%	96.46%	2.52%
GSTRUCT ann	97.33%	86.94%	10.23%	0.00%	96.73%	0.23%	99.37%	87.93%	10.39%	0.00%	97.88%	1.11%
MapSplice	97.09%	89.22%	7.01%	0.00%	95.09%	0.96%	97.46%	89.27%	7.04%	0.00%	95.17%	1.24%
MapSplice ann	97.51%	88.89%	7.82%	0.00%	95.69%	0.93%	97.86%	88.91%	7.86%	0.00%	95.75%	1.21%
PALMapper	35.11%	32.21%	2.85%	0.00%	34.77%	0.33%	98.58%	81.36%	14.58%	0.00%	94.77%	3.81%
PALMapper ann	33.21%	31.60%	1.59%	0.00%	33.05%	0.16%	98.95%	83.57%	12.49%	0.00%	95.22%	3.74%
PALMapper cons	41.10%	35.99%	4.56%	0.00%	40.42%	0.68%	52.14%	44.62%	5.54%	0.00%	49.96%	2.18%
PALMapper cons ann	63.26%	55.90%	7.30%	0.00%	62.88%	0.39%	97.29%	75.16%	19.30%	0.00%	93.27%	4.02%
PASS	92.26%	56.31%	29.43%	0.01%	82.78%	7.01%	92.44%	56.35%	29.50%	0.01%	82.86%	7.09%
PASS cons	91.61%	56.31%	29.01%	0.01%	82.52%	6.81%	91.75%	56.34%	29.07%	0.01%	82.59%	6.87%
ReadsMap	94.52%	87.94%	2.26%	4.19%	89.99%	4.53%	97.44%	89.05%	2.32%	4.24%	91.14%	6.29%
SMALT	96.10%	5.52%	83.88%	0.00%	72.96%	8.06%	96.65%	5.52%	84.39%	0.00%	73.27%	8.09%
STAR 1-pass	96.68%	59.73%	35.86%	0.00%	91.57%	1.79%	98.81%	60.31%	36.32%	0.00%	92.53%	2.88%
STAR 1-pass ann	94.73%	82.28%	12.26%	0.00%	93.43%	0.35%	99.14%	84.77%	13.53%	0.00%	97.11%	1.03%
STAR 2-pass	94.46%	83.70%	10.47%	0.00%	93.22%	0.41%	99.18%	86.72%	11.66%	0.00%	97.34%	0.95%
STAR 2-pass ann	93.82%	83.73%	9.90%	0.00%	92.80%	0.29%	99.16%	87.13%	11.23%	0.00%	97.46%	0.93%
TopHat1	91.77%	78.88%	9.43%	0.00%	87.93%	3.84%	93.03%	79.29%	9.64%	0.00%	88.53%	4.50%
TopHat1 ann	92.48%	79.36%	9.66%	0.00%	88.63%	3.85%	93.81%	79.82%	9.89%	0.00%	89.30%	4.51%
TopHat2	88.01%	84.78%	1.78%	0.00%	86.42%	1.59%	90.02%	85.57%	2.23%	0.00%	87.56%	2.46%
TopHat2 ann	91.24%	90.04%	1.08%	0.00%	91.06%	0.18%	94.51%	91.82%	1.75%	0.00%	93.47%	1.04%
B. Simulation 2												
BAGET ann	85.25%	36.99%	33.78%	0.01%	59.31%	14.73%	90.25%	36.99%	38.15%	0.01%	60.74%	14.96%
GEM ann	27.43%	12.78%	13.67%	0.03%	23.77%	2.95%	97.25%	66.53%	25.21%	0.05%	87.09%	8.79%
GEM cons	28.68%	13.04%	14.63%	0.04%	24.78%	3.14%	97.02%	65.67%	25.79%	0.05%	86.66%	8.98%
GEM cons ann	27.85%	13.00%	13.87%	0.03%	24.14%	2.99%	97.20%	66.39%	25.29%	0.05%	87.02%	8.82%
GSNAP	94.22%	51.61%	41.43%	0.00%	89.48%	1.66%	97.36%	52.53%	42.26%	0.00%	91.15%	3.02%
GSNAP ann	94.43%	70.60%	23.51%	0.00%	92.82%	0.44%	97.45%	71.93%	23.97%	0.00%	94.59%	1.65%
GSTRUCT	93.53%	72.18%	20.61%	0.00%	91.67%	0.85%	97.73%	73.62%	21.03%	0.00%	93.50%	3.16%
GSTRUCT ann	95.04%	73.83%	20.68%	0.00%	93.40%	0.63%	97.72%	74.86%	21.01%	0.00%	94.74%	1.93%
MapSplice	88.03%	71.09%	15.41%	0.00%	82.84%	1.63%	88.44%	71.12%	15.51%	0.00%	82.95%	1.88%
MapSplice ann	88.97%	71.60%	15.83%	0.00%	84.03%	1.69%	89.35%	71.61%	15.93%	0.00%	84.12%	1.93%
PALMapper	30.30%	24.79%	5.36%	0.00%	29.51%	0.79%	95.42%	69.75%	22.73%	0.00%	90.24%	5.18%
PALMapper ann	26.37%	23.94%	2.37%	0.00%	26.05%	0.32%	96.47%	75.00%	18.43%	0.00%	92.08%	4.39%
PALMapper cons	41.40%	30.85%	9.70%	0.00%	40.07%	1.34%	59.63%	40.60%	15.50%	0.00%	55.29%	4.34%
PALMapper cons ann	58.04%	50.08%	7.82%	0.00%	57.55%	0.49%	91.45%	65.20%	23.19%	0.00%	86.99%	4.47%
PASS	78.31%	31.92%	40.48%	0.02%	66.53%	6.17%	78.50%	31.94%	40.57%	0.02%	66.61%	6.25%
PASS cons	75.77%	31.92%	38.45%	0.02%	65.27%	5.76%	75.90%	31.94%	38.52%	0.02%	65.34%	5.81%
ReadsMap	87.63%	72.81%	4.81%	9.70%	77.02%	10.61%	90.82%	73.88%	4.93%	9.82%	78.17%	12.65%
SMALT	94.88%	4.13%	83.91%	0.00%	70.66%	7.84%	95.85%	4.13%	84.75%	0.00%	71.16%	7.91%
STAR 1-pass	91.80%	42.50%	48.05%	0.00%	82.96%	2.14%	94.50%	43.11%	48.93%	0.00%	84.26%	3.30%
STAR 1-pass ann	91.98%	63.42%	28.13%	0.00%	87.96%	0.80%	96.47%	65.30%	29.87%	0.00%	91.37%	1.69%
STAR 2-pass	91.89%	65.27%	25.99%	0.00%	87.96%	0.95%	96.70%	67.61%	27.80%	0.00%	91.90%	1.62%
STAR 2-pass ann	91.34%	65.81%	25.10%	0.00%	87.89%	0.73%	96.71%	68.46%	26.98%	0.00%	92.22%	1.59%
TopHat1	77.46%	66.62%	8.04%	0.00%	74.35%	3.11%	79.17%	67.23%	8.38%	0.00%	75.26%	3.91%
TopHat1 ann	79.59%	68.59%	8.23%	0.00%	76.49%	3.10%	81.39%	69.26%	8.60%	0.00%	77.50%	3.90%
TopHat2	65.76%	62.38%	2.35%	0.00%	64.57%	1.19%	67.56%	63.07%	2.73%	0.00%	65.54%	2.02%
TopHat2 ann	73.10%	70.93%	1.99%	0.00%	72.81%	0.29%	76.50%	72.68%	2.67%	0.00%	75.18%	1.32%

Results are shown for simulated spliced reads from the nuclear genome, and percentages are relative to the total number of such reads. Perfectly mapped reads have all 76 bases correctly placed (accounting for ambiguity in indel placement as described in Methods). Part correctly mapped reads have at least one base correctly placed, but not all 76. Reads mapped near the correct location are those for which no base is correctly placed, but the mapping overlaps with the correct mapping (this may occur in repetitive regions or indicate a bug in the aligner, as for ReadsMap).

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 7. Mapping accuracy for simulated data (unspliced reads).

	Uniquely mapped reads						All reads (primary alignment counted)					
	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases	Mapped reads	Perfectly mapped reads	Part correctly mapped reads	Reads mapped near correct location	Correctly mapped bases	Incorrectly mapped bases
A. Simulation 1												
BAGET ann	99.22%	96.23%	0.84%	0.02%	96.83%	2.09%	99.45%	96.23%	0.98%	0.02%	96.88%	2.11%
GEM ann	82.81%	82.18%	0.58%	0.00%	82.73%	0.08%	99.99%	96.53%	0.79%	0.02%	97.27%	2.71%
GEM cons	82.98%	82.34%	0.59%	0.00%	82.89%	0.08%	99.99%	96.53%	0.78%	0.02%	97.27%	2.71%
GEM cons ann	82.90%	82.26%	0.59%	0.00%	82.81%	0.08%	99.99%	96.53%	0.78%	0.02%	97.27%	2.71%
GSNAP	95.46%	87.50%	7.85%	0.01%	95.03%	0.11%	99.21%	89.64%	8.03%	0.02%	97.35%	1.52%
GSNAP ann	95.45%	87.51%	7.89%	0.01%	95.08%	0.05%	99.21%	89.68%	8.09%	0.02%	97.44%	1.42%
GSTRUCT	95.84%	88.08%	7.59%	0.01%	95.36%	0.17%	99.21%	89.66%	7.72%	0.02%	97.07%	1.81%
GSTRUCT ann	96.40%	88.60%	7.63%	0.01%	95.92%	0.17%	99.21%	90.07%	7.76%	0.02%	97.52%	1.36%
MapSplice	96.61%	95.60%	0.75%	0.01%	96.15%	0.26%	98.82%	96.68%	0.75%	0.02%	97.24%	1.38%
MapSplice ann	96.61%	95.60%	0.76%	0.01%	96.16%	0.25%	98.82%	96.68%	0.77%	0.02%	97.25%	1.37%
PALMapper	76.66%	76.55%	0.09%	0.00%	76.62%	0.04%	98.17%	93.73%	1.66%	0.03%	95.31%	2.87%
PALMapper ann	75.97%	75.87%	0.07%	0.00%	75.93%	0.04%	98.17%	93.14%	1.85%	0.03%	94.89%	3.28%
PALMapper cons	81.33%	81.28%	0.03%	0.00%	81.31%	0.02%	87.11%	85.34%	0.19%	0.01%	85.52%	1.59%
PALMapper cons ann	81.39%	81.26%	0.09%	0.00%	81.35%	0.04%	97.71%	93.69%	1.62%	0.02%	95.23%	2.47%
PASS	97.21%	46.10%	48.92%	0.02%	92.53%	2.12%	97.93%	46.23%	49.05%	0.02%	92.78%	2.57%
PASS cons	96.26%	46.06%	48.41%	0.02%	92.16%	1.76%	96.82%	46.18%	48.54%	0.02%	92.40%	2.07%
ReadsMap	80.93%	72.11%	0.47%	3.83%	72.52%	8.41%	84.47%	73.31%	0.50%	3.88%	73.73%	10.74%
SMALT	96.87%	89.70%	6.63%	0.00%	96.11%	0.38%	96.95%	89.70%	6.63%	0.00%	96.11%	0.41%
STAR 1-pass	96.01%	90.69%	5.17%	0.00%	95.68%	0.15%	98.70%	92.04%	5.25%	0.01%	97.10%	1.41%
STAR 1-pass ann	95.76%	90.43%	5.22%	0.00%	95.46%	0.12%	98.73%	92.08%	5.33%	0.01%	97.21%	1.32%
STAR 2-pass	95.73%	90.42%	5.21%	0.00%	95.45%	0.10%	98.73%	92.11%	5.33%	0.01%	97.24%	1.30%
STAR 2-pass ann	95.65%	90.29%	5.24%	0.00%	95.34%	0.12%	98.73%	92.04%	5.37%	0.01%	97.21%	1.33%
TopHat1	93.76%	92.87%	0.13%	0.02%	93.00%	0.77%	95.77%	93.70%	0.14%	0.02%	93.83%	1.94%
TopHat1 ann	93.76%	92.86%	0.13%	0.02%	92.99%	0.77%	95.77%	93.70%	0.14%	0.02%	93.83%	1.94%
TopHat2	92.20%	91.78%	0.14%	0.01%	91.92%	0.29%	94.73%	92.89%	0.15%	0.02%	93.03%	1.70%
TopHat2 ann	92.19%	91.67%	0.19%	0.01%	91.85%	0.34%	94.65%	92.85%	0.25%	0.02%	93.08%	1.57%
B. Simulation 2												
BAGET ann	97.27%	90.86%	3.94%	0.01%	93.68%	2.31%	98.35%	90.86%	4.81%	0.01%	93.97%	2.38%
GEM ann	82.62%	79.35%	3.16%	0.00%	82.25%	0.33%	99.92%	92.25%	4.19%	0.02%	96.09%	3.78%
GEM cons	83.12%	79.79%	3.22%	0.00%	82.75%	0.33%	99.92%	92.25%	4.18%	0.02%	96.08%	3.79%
GEM cons ann	82.96%	79.65%	3.20%	0.00%	82.59%	0.33%	99.92%	92.25%	4.18%	0.02%	96.09%	3.78%
GSNAP	94.29%	75.62%	18.47%	0.01%	93.21%	0.18%	98.10%	77.39%	18.89%	0.02%	95.38%	1.77%
GSNAP ann	94.29%	75.64%	18.51%	0.01%	93.26%	0.13%	98.10%	77.42%	18.93%	0.02%	95.44%	1.71%
GSTRUCT	94.70%	77.00%	17.38%	0.01%	93.55%	0.31%	98.20%	78.34%	17.68%	0.02%	95.17%	2.14%
GSTRUCT ann	95.21%	77.46%	17.46%	0.01%	94.08%	0.29%	98.20%	78.71%	17.74%	0.02%	95.60%	1.72%
MapSplice	93.85%	89.68%	3.72%	0.01%	92.41%	0.44%	96.11%	90.77%	3.76%	0.02%	93.53%	1.56%
MapSplice ann	93.86%	89.65%	3.76%	0.01%	92.40%	0.46%	96.10%	90.73%	3.79%	0.02%	93.51%	1.57%
PALMapper	70.87%	70.31%	0.48%	0.00%	70.75%	0.13%	97.11%	88.95%	5.02%	0.02%	93.71%	3.40%
PALMapper ann	69.03%	68.58%	0.38%	0.00%	68.91%	0.12%	97.12%	88.23%	5.27%	0.03%	93.21%	3.91%
PALMapper cons	78.31%	77.59%	0.64%	0.00%	78.21%	0.10%	90.86%	85.32%	3.13%	0.02%	88.36%	2.50%
PALMapper cons ann	76.04%	75.40%	0.55%	0.00%	75.92%	0.12%	95.01%	86.69%	5.41%	0.02%	91.87%	3.14%
PASS	92.32%	24.52%	65.36%	0.02%	83.71%	2.29%	92.97%	24.58%	65.51%	0.02%	83.89%	2.69%
PASS cons	89.84%	24.50%	63.46%	0.02%	82.47%	1.84%	90.29%	24.55%	63.60%	0.02%	82.65%	2.08%
ReadsMap	81.27%	69.06%	1.35%	6.95%	70.22%	11.06%	85.43%	70.14%	1.37%	7.05%	71.32%	14.11%
SMALT	96.36%	79.19%	16.17%	0.00%	94.73%	0.67%	96.45%	79.19%	16.17%	0.00%	94.73%	0.69%
STAR 1-pass	93.74%	79.84%	13.68%	0.00%	92.64%	0.26%	96.65%	81.13%	13.90%	0.01%	94.14%	1.64%
STAR 1-pass ann	93.66%	79.71%	13.74%	0.00%	92.55%	0.25%	96.77%	81.20%	14.02%	0.01%	94.30%	1.58%
STAR 2-pass	93.56%	79.59%	13.76%	0.00%	92.45%	0.26%	96.79%	81.18%	14.07%	0.01%	94.33%	1.57%
STAR 2-pass ann	93.51%	79.52%	13.77%	0.00%	92.39%	0.26%	96.79%	81.14%	14.09%	0.01%	94.31%	1.59%
TopHat1	85.55%	84.41%	0.59%	0.02%	84.97%	0.58%	87.77%	85.29%	0.63%	0.02%	85.89%	1.88%
TopHat1 ann	85.56%	84.41%	0.59%	0.02%	84.98%	0.58%	87.78%	85.29%	0.63%	0.02%	85.90%	1.87%
TopHat2	77.90%	77.20%	0.51%	0.01%	77.69%	0.21%	80.44%	78.25%	0.54%	0.02%	78.76%	1.68%
TopHat2 ann	78.01%	77.15%	0.55%	0.01%	77.68%	0.33%	80.41%	78.22%	0.62%	0.02%	78.80%	1.61%

Results are shown for simulated unspliced reads from the nuclear genome, and percentages are relative to the total number of such reads. Perfectly mapped reads have all 76 bases correctly placed (accounting for ambiguity in indel placement as described in Methods). Part correctly mapped reads have at least one base correctly placed, but not all 76. Reads mapped near the correct location are those for which no base is correctly placed, but the mapping overlaps with the correct mapping (this may occur in repetitive regions or indicate a bug in the aligner, as for ReadsMap).

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 8. Consistency of novel junction calls among protocols.

	BAGET ann	GEM ann	GEM cons	GEM cons ann	GSNAP	GSNAP ann	GSTRUCT	GSTRUCT ann	MapSplice	MapSplice ann	PALMapper	PALMapper cons	PASS	PASS cons	ReadsMap	SMALT	STAR 1-pass	STAR 1-pass ann	STAR 2-pass	STAR 2-pass ann	TopHat1	TopHat1 ann	TopHat2	TopHat2 ann	Union excluding protocol	Union excluding team
A. All novel junctions																										
BAGET ann	**0.7**	0.1	0.0	0.0	0.1	0.0	0.1	0.1	0.0	0.3	0.8	0.0	0.0	0.0	0.0	11.6	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.0	12.8	12.8
GEM ann	0.0	**139.3**	30.7	30.8	78.8	78.1	77.8	77.7	39.0	39.	69.5	24.8	69.9	68.4	55.6	13.7	79.2	78.6	79.8	79.7	53.6	53.5	54.3	53.2	90.7	86.6
GEM cons	0.0	98.1	**43.6**	99.8	78.1	77.3	77.4	77.1	63.0	63.	74.4	50.1	74.7	73.9	61.7	22.1	78.2	77.2	77.9	77.8	69.7	69.6	70.5	67.8	99.9	85.6
GEM cons	0.0	98.3	99.9	**43.6**	78.1	77.3	77.4	77.1	63.0	63.	74.4	50.2	74.7	73.9	61.8	22.1	78.2	77.2	77.9	77.8	69.7	69.6	70.5	67.9	100.0	85.7
GSNAP	0.0	34.4	10.7	10.7	**319.1**	86.9	82.0	81.8	19.3	19.	38.5	11.4	39.1	38.1	26.9	8.3	45.2	45.2	47.3	47.2	24.6	24.6	24.4	24.1	90.0	54.4
GSNAP	0.0	33.3	10.3	10.3	84.8	**327.1**	88.1	88.8	18.8	19.	38.1	11.1	37.8	36.8	25.9	7.6	43.5	43.7	46.2	46.2	23.8	23.7	23.7	23.6	94.9	52.6
GSTRUCT	0.0	32.9	10.2	10.2	79.4	87.4	**329.5**	97.0	18.6	19.	38.1	11.0	37.5	36.6	25.7	7.5	42.9	43.1	45.7	45.7	23.7	23.6	23.6	23.3	98.3	52.5
GSTRUCT	0.0	32.8	10.2	10.2	78.9	87.9	96.7	**330.7**	18.5	19.	38.0	11.0	37.4	36.4	25.6	7.5	42.8	42.9	45.6	45.6	23.5	23.5	23.4	23.3	98.3	52.3
MapSplice	0.0	83.5	42.2	42.2	94.5	94.4	94.2	94.2	**65.1**	98.	86.9	42.7	88.0	86.7	67.7	20.4	93.5	93.1	94.8	94.7	66.1	66.0	67.1	66.4	99.6	98.5
MapSplice	0.0	81.5	40.6	40.6	92.6	92.9	92.6	92.7	93.6	**68.**	85.2	41.0	86.1	84.8	66.5	20.1	91.6	91.2	93.1	93.0	64.2	64.2	65.2	64.6	98.7	97.6
PALMappe	0.0	5.0	1.7	1.7	6.3	6.4	6.5	6.5	2.9	3.0	**1942.**	1.9	5.7	5.6	4.0	1.0	6.0	6.0	6.8	6.8	3.8	3.8	3.8	3.7	7.8	7.8
PALMappe	0.0	94.8	60.1	60.1	99.5	99.6	99.6	99.6	76.3	76.	100.0	**36.4**	97.7	96.9	73.6	28.0	99.3	99.3	99.6	99.6	83.2	83.2	83.1	83.3	100.0	100.0
PASS	0.0	27.8	9.3	9.3	35.6	35.3	35.3	35.3	16.4	16.	31.8	10.2	**349.8**	72.3	22.8	5.6	34.2	34.1	35.7	35.7	20.3	20.3	20.5	20.2	74.0	40.0
PASS cons	0.0	37.0	12.5	12.5	47.1	46.7	46.8	46.7	21.9	22.	41.9	13.7	98.2	**257.7**	30.2	7.5	45.3	45.1	47.2	47.2	27.1	27.1	27.4	27.0	98.5	52.4
ReadsMap	0.0	4.3	1.5	1.5	4.7	4.7	4.7	4.6	2.4	2.5	4.2	1.5	4.4	4.3	**1817.**	3.4	4.7	4.7	4.7	4.7	3.2	3.2	3.3	3.3	7.9	7.9
SMALT	0.0	6.6	3.3	3.3	9.1	8.6	8.5	8.5	4.6	4.7	7.0	3.5	6.7	6.7	21.5	**289.8**	7.1	7.0	7.1	7.1	5.8	5.8	5.7	5.7	26.6	26.6
STAR 1-	0.0	55.4	17.1	17.1	72.3	71.5	71.0	71.0	30.5	31.	58.5	18.1	60.0	58.6	42.8	10.3	**199.2**	95.1	95.1	94.9	38.6	38.5	38.6	38.3	97.3	78.9
STAR 1-	0.0	53.7	16.5	16.5	70.8	70.2	69.7	69.7	29.7	30.	57.3	17.7	58.5	57.1	41.7	10.0	93.0	**203.7**	97.8	98.0	37.4	37.4	37.5	37.5	99.6	77.2
STAR 2-	0.0	36.6	11.2	11.2	49.6	49.7	49.6	49.6	20.3	20.	43.5	11.9	41.1	40.1	28.4	6.8	62.3	65.6	**303.8**	99.6	25.4	25.4	25.5	25.3	99.8	59.1
STAR 2-	0.0	36.6	11.2	11.2	49.7	49.8	49.7	49.7	20.3	21.	43.6	12.0	41.2	40.1	28.4	6.8	62.4	65.9	99.8	**303.2**	25.5	25.4	25.6	25.4	100.0	59.2
TopHat1	0.0	81.8	33.3	33.3	86.0	85.2	85.4	85.2	47.1	48.	80.9	33.2	77.8	76.5	64.2	18.6	84.3	83.5	84.7	84.6	**91.2**	97.6	81.5	78.7	99.1	93.6
TopHat1	0.0	81.9	33.4	33.3	86.1	85.3	85.6	85.3	47.2	48.	81.0	33.2	77.8	76.6	64.4	18.6	84.3	83.6	84.8	84.7	97.8	**91.0**	81.5	78.9	99.2	93.7
TopHat2	0.0	83.2	33.9	33.8	85.9	85.4	85.5	85.3	48.1	49.	81.2	33.3	79.0	77.8	66.7	18.3	84.7	84.2	85.4	85.3	81.8	81.7	**90.8**	90.6	97.4	93.8
TopHat2	0.0	86.6	34.6	34.6	90.0	90.0	89.9	90.0	50.5	51.	84.8	35.4	82.7	81.3	69.3	19.3	89.2	89.3	90.0	90.0	83.9	83.9	96.2	**85.6**	99.3	96.0
B. Novel junctions with at least two mappings																										
BAGET ann	**0.1**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.5	7.5
GEM ann	0.0	**82.1**	46.8	46.8	84.1	83.5	83.3	83.1	54.6	55.	77.1	37.7	78.1	76.9	62.5	18.0	84.8	84.1	85.0	84.9	63.7	63.6	64.9	63.6	94.0	91.2
GEM cons	0.0	99.0	**39.0**	99.9	86.3	85.5	85.6	85.4	69.9	70.	82.0	56.0	82.6	81.9	67.8	24.6	86.6	85.5	86.1	86.0	76.7	76.6	77.6	75.1	100.0	93.1
GEM cons	0.0	99.1	99.9	**39.0**	86.3	85.5	85.7	85.4	69.9	70.	82.0	56.1	82.7	81.9	67.8	24.7	86.6	85.6	86.2	86.1	76.7	76.6	77.6	75.1	100.0	93.1
GSNAP	0.0	56.8	29.8	29.8	**112.3**	92.2	88.6	87.9	41.8	42.	61.3	28.1	62.8	61.8	45.2	14.9	67.9	67.7	69.3	69.3	45.0	45.0	45.4	44.8	95.1	76.1
GSNAP	0.0	55.3	28.5	28.5	89.6	**116.3**	90.9	91.6	41.2	42.	60.7	27.4	61.2	60.2	44.0	14.0	66.1	66.2	68.3	68.3	43.7	43.7	44.2	43.9	96.4	74.2
GSTRUCT	0.0	53.3	27.5	27.5	83.3	87.7	**121.0**	97.1	39.8	40.	59.2	26.3	59.3	58.3	42.5	13.4	63.6	63.6	66.0	66.0	42.3	42.3	42.8	42.3	98.5	72.5
GSTRUCT	0.0	53.1	27.3	27.2	82.5	88.2	96.8	**121.5**	39.6	40.	59.1	26.2	59.0	58.0	42.2	13.3	63.3	63.4	65.8	65.8	42.0	42.0	42.5	42.2	98.4	72.2
MapSplice	0.0	85.8	54.7	54.7	94.4	94.1	94.0	93.9	**49.7**	98.	87.9	51.2	89.4	88.3	68.2	23.7	93.7	93.3	94.5	94.4	71.0	71.0	71.7	71.0	99.6	98.3
MapSplice	0.0	83.4	52.8	52.8	92.2	92.2	92.0	92.0	94.3	**51.**	85.9	49.4	87.1	86.0	66.8	23.4	91.4	91.1	92.4	92.4	68.8	68.8	69.5	68.9	98.4	97.1
PALMappe	0.0	10.0	5.0	5.0	11.9	12.0	12.1	12.1	7.2	7.3	**641.6**	5.3	11.3	11.0	8.2	2.4	11.4	11.4	12.9	12.9	8.1	8.1	8.2	8.1	14.5	14.5
PALMappe	0.0	95.9	80.0	80.0	99.7	99.7	99.7	99.7	83.8	84.	100.0	**22.9**	98.8	98.2	73.5	33.3	99.6	99.6	99.7	99.6	86.6	86.5	85.9	86.1	100.0	100.0
PASS	0.0	49.4	28.2	28.2	57.6	57.0	57.1	57.0	37.5	38.	54.0	27.0	**111.7**	89.8	40.2	12.3	56.1	55.8	57.3	57.3	40.0	40.0	40.6	39.9	90.9	62.2
PASS cons	0.0	55.7	32.2	32.2	64.7	64.0	64.1	64.0	42.5	43.	60.6	30.7	99.7	**96.7**	45.2	14.0	63.1	62.7	64.3	64.2	45.2	45.2	45.9	45.1	99.8	69.3
ReadsMap	0.0	6.7	2.9	2.9	7.2	7.1	7.1	7.1	4.2	4.3	6.5	2.7	6.8	6.6	**902.4**	5.1	7.2	7.2	7.3	7.3	5.2	5.2	5.4	5.3	11.9	11.9
SMALT	0.0	5.3	4.0	4.0	6.5	6.2	6.2	6.2	4.3	4.5	5.7	3.7	5.3	5.3	28.9	**160.8**	5.5	5.5	5.5	5.5	4.9	4.8	4.8	4.8	32.4	32.4
STAR 1-	0.0	72.7	39.6	39.6	84.9	84.0	83.6	83.4	53.0	53.	74.6	36.5	77.0	75.8	56.9	17.2	**84.8**	97.0	97.0	96.9	57.1	57.0	57.6	57.1	98.7	90.0
STAR 1-	0.0	72.0	39.0	39.0	84.7	84.2	83.6	83.6	52.7	53.	74.5	36.5	76.7	75.5	56.6	17.2	96.9	**84.6**	98.6	98.8	56.5	56.4	57.1	57.2	99.9	89.8
STAR 2-	0.0	44.5	19.1	19.1	56.1	56.1	56.0	56.0	30.0	30.	52.5	19.1	49.4	48.3	35.4	9.3	63.5	64.0	**175.8**	99.7	33.2	33.1	33.6	33.3	99.8	65.2
STAR 2-	0.0	44.5	19.2	19.2	56.2	56.2	56.1	56.1	30.1	30.	52.6	19.2	49.5	48.4	35.5	9.4	63.6	64.3	99.9	**175.1**	33.2	33.2	33.6	33.4	100.0	65.3
TopHat1	0.0	84.6	45.5	45.5	87.6	86.8	87.2	86.9	57.8	58.	82.9	42.5	81.5	80.5	69.1	21.3	86.5	85.7	86.7	86.6	**65.3**	98.9	86.1	83.1	99.6	94.5
TopHat1	0.0	84.7	45.5	45.5	87.7	86.9	87.3	87.0	57.9	58.	83.1	42.6	81.6	80.6	69.2	21.3	86.5	85.7	86.8	86.7	99.1	**65.1**	86.1	83.3	99.7	94.5
TopHat2	0.0	85.5	47.8	47.8	87.0	86.4	86.7	86.5	59.1	59.	83.0	43.4	82.0	81.1	70.6	21.0	86.2	85.6	86.7	86.6	85.0	84.9	**63.1**	92.5	98.3	94.6
TopHat2	0.0	88.6	48.3	48.3	90.9	90.9	90.8	90.8	61.8	62.	86.2	45.7	85.3	84.1	73.1	22.0	90.5	90.6	91.2	91.2	86.4	86.3	97.4	**60.3**	99.6	96.4

Results are shown for K562 whole cell replicate 1. Values on the diagonal (bold) indicate the number of unannotated junctions reported by each protocol (thousands). Off-diagonal values measure pairwise agreement between protocols. Specifically, for row *R* and column *C*, we define N(*k*;*R*,*C*) as the number of novel junctions that are supported by at least *k* mappings from protocol *R* and at least one mapping from protocol *C*, and display the proportion N(*k*;*R*,*C*) / N(*k*;*R*,*R*) for *k*=1 (table A) and *k*=2 (table B) in units of percent. Note that the denominator N(*k*;*R*,*R*) is effectively the number of junctions supported by at least *k* mappings from to protocol *R*, i.e. the values tabulated on the diagonal. In the two rightmost columns, *R* is compared to the combined output from all other protocols, or all protocols from other developer teams.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 9. Accuracy of junction discovery on simulated data.

	True junctions							False junctions						
	1	2	3	4	5	6	7	1	2	3	4	5	6	7
A. Simulation 1														
BAGET ann	77524	73129	69942	67549	65546	63737	62130	5144	3397	2752	2372	2150	1972	1810
GEM ann	116613	110636	105531	101265	97581	94269	91300	9193	5090	3887	3238	2774	2488	2266
GEM cons	97700	97566	97408	96169	94150	91740	89266	4689	3479	2986	2638	2330	2125	1956
GEM cons ann	108106	104597	101922	99273	96423	93548	90804	6251	4192	3456	2999	2593	2345	2142
GSNAP	118830	110947	104976	100267	96217	92799	89723	13287	5128	3814	3103	2613	2321	2097
GSNAP ann	120817	113861	108589	104300	100616	97362	94503	18539	8208	5822	4546	3784	3276	2863
GSTRUCT	119584	112315	106793	103804	100703	97788	95044	8890	3531	2832	2434	2184	2017	1869
GSTRUCT ann	119779	112836	108494	104821	101315	98144	95322	8447	3207	2522	2172	1925	1743	1599
MapSplice	115689	111331	106663	102584	99019	95922	93075	4071	1970	1595	1348	1190	1072	991
MapSplice ann	119040	112564	107469	103222	99589	96460	93553	22445	6504	3917	2911	2369	2066	1862
PALMapper	117210	110112	105936	102172	98686	95554	92752	283036	68956	41034	29074	22520	18426	15638
PALMapper ann	118654	111714	107418	103686	100219	97046	94199	325933	78723	49658	37117	29991	25270	21978
PALMapper cons	106353	102086	95731	90320	85874	81982	78553	7272	4538	3554	3032	2691	2421	2240
PALMapper cons ann	108253	107507	105178	101959	98239	94967	91997	43234	28946	23061	19391	16956	15108	13703
PASS	114014	105797	99743	94885	90900	87485	84486	62605	16760	10401	7826	6305	5292	4683
PASS cons	113828	105707	99696	94840	90868	87453	84450	37293	12528	8437	6607	5486	4722	4221
ReadsMap	114148	109812	105452	101661	98320	95360	92693	898713	421115	272289	199817	156865	128492	108555
SMALT	50497	41008	35546	31504	28431	25900	23692	140685	92591	77404	67930	60578	54614	49584
STAR 1-pass	116236	107929	101799	96986	92896	89402	86334	6528	2563	2082	1796	1604	1457	1357
STAR 1-pass ann	119007	111394	105572	100953	97035	93664	90790	20226	10056	7323	5832	4871	4182	3678
STAR 2-pass	117081	112014	107278	103202	99619	96518	93727	11579	5088	3789	3105	2640	2327	2092
STAR 2-pass ann	119222	113383	108425	104253	100619	97497	94668	21203	10324	7305	5776	4765	4066	3570
TopHat1	108779	105599	101942	98446	95189	92246	89518	7709	5594	4515	3859	3425	3090	2828
TopHat1 ann	113180	108599	104170	100270	96754	93677	90817	8373	6179	5044	4306	3822	3431	3181
TopHat2	109673	106504	102741	99093	95857	93022	90460	7891	5565	4471	3847	3405	3042	2789
TopHat2 ann	115945	111117	106709	102838	99397	96425	93768	24336	14583	10354	8036	6571	5573	4817
Truth	122745	116040	110976	106744	103132	99965	97158	0	0	0	0	0	0	0
B. Simulation 2														
BAGET ann	76953	72955	70245	68042	66147	64426	62806	6331	3900	3077	2651	2417	2238	2059
GEM ann	112359	107048	102924	99412	96309	93360	90819	22293	12427	9024	7213	6154	5368	4820
GEM cons	91373	91235	91121	90396	89195	87536	85811	12622	8403	6631	5539	4854	4320	3908
GEM cons ann	105415	101642	99134	96816	94557	92150	89914	14782	9433	7314	6059	5271	4663	4214
GSNAP	119276	112394	107561	103605	100140	97025	94184	30694	7716	5394	4324	3651	3238	2904
GSNAP ann	121420	115406	111206	107760	104783	102064	99575	36639	11359	7861	6122	5071	4369	3834
GSTRUCT	119916	113667	109264	106977	104638	102230	99927	23071	5778	4306	3672	3243	2984	2747
GSTRUCT ann	120215	114371	111153	108142	105364	102725	100330	22613	5374	3990	3378	2943	2675	2468
MapSplice	109651	106509	102957	99736	96811	94145	91687	9306	4601	3771	3314	2914	2682	2473
MapSplice ann	116470	110275	105742	101987	98721	95808	93193	33963	11203	7265	5771	5000	4495	4121
PALMapper	115685	109797	106564	103574	100882	98237	95802	383917	77636	47553	34831	27834	23504	20514
PALMapper ann	118141	112197	108627	105585	102755	100080	97660	528217	103303	63781	48099	39411	33881	29788
PALMapper cons	103937	102080	98619	94685	90861	87281	84023	12261	9066	7610	6713	6090	5580	5172
PALMapper cons ann	105888	105474	104422	102703	100485	97936	95344	59119	41790	34471	29873	26574	24057	21988
PASS	107833	100322	95225	91084	87380	84195	81312	125292	31174	19369	14724	12254	10596	9388
PASS cons	107558	100128	95030	90853	87142	83988	81057	77363	24454	16074	12487	10536	9205	8258
ReadsMap	109047	105544	102270	99308	96640	94122	91777	942684	415590	259006	184541	141596	113825	94466
SMALT	50726	41116	35239	30947	27516	24694	22277	181841	103528	82418	70436	61738	54887	49302
STAR 1-pass	110301	102851	97694	93452	89795	86566	83641	14893	4526	3471	2883	2563	2322	2130
STAR 1-pass ann	116771	109902	105140	101334	97941	94936	92329	31696	12485	9022	7239	6063	5289	4655
STAR 2-pass	113032	108903	105346	102261	99503	96811	94492	22749	8721	6366	5282	4581	4082	3690
STAR 2-pass ann	117144	112022	108147	104872	101968	99188	96774	32855	14117	9979	8051	6782	5913	5236
TopHat1	101390	98839	96117	93382	90861	88449	86201	11384	8223	6557	5540	4875	4351	3951
TopHat1 ann	108919	104901	101269	98035	95122	92459	89994	12275	8946	7196	6076	5320	4735	4303
TopHat2	104273	101654	98660	95567	92706	90090	87728	9568	7242	6027	5280	4722	4266	3921
TopHat2 ann	113564	109175	105494	102149	99129	96421	93937	26389	16514	12196	9786	8179	7059	6227
Truth	123581	117890	113826	110530	107667	105088	102713	0	0	0	0	0	0	0

Number of unique junctions reported for the two simulated data sets, at a range of thresholds (1-7) for the number of primary alignments supporting a junction. Higher thresholds correspond to a more conservative interpretation of alignment results. Junctions were classified as true and false by comparison to the true simulated alignments. The row labeled "Truth" shows the result expected for a perfect aligner.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 10. Number of introns reported per alignment.

	Primary alignments						All alignments					
	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns
A. K562 whole cell replicate 1												
BAGET ann	187077254	19044438	0	0	0	0	190657737	19044438	0	0	0	0
GEM ann	171310851	34730775	1064776	1509	0	0	499923501	76562905	13247099	429798	3108	0
GEM cons	171601899	34567687	898840	209	0	0	494153538	69729534	8070043	378798	2443	0
GEM cons ann	171519716	34628760	938481	1012	0	0	495520648	72100675	8478289	385549	2445	0
GSNAP	169869232	36780143	403327	2387	14	0	222746918	39531697	411451	2541	14	0
GSNAP ann	155959666	49495787	1595726	11463	35	0	207664707	53698921	1707068	12197	36	0
GSTRUCT	160220367	45290527	1586733	9674	38	0	222315512	54330580	1848701	10801	39	0
GSTRUCT ann	157984101	47426882	1683304	10286	37	0	209374754	54242555	1855777	11163	39	0
MapSplice	155715733	40229115	1062125	1353	0	0	158023362	40229125	1062125	1353	0	0
MapSplice ann	156017042	39908235	1068363	1869	0	0	158330301	39908238	1068363	1869	0	0
PALMapper	158313084	43659317	2733642	0	0	0	1615733787	98720318	9598992	0	0	0
PALMapper cons	60036227	8633289	25053	0	0	0	145387747	19789755	82347	0	0	0
PASS	171743999	25222924	115456	6	0	0	173860480	26029419	115739	6	0	0
PASS cons	167272372	24802361	114869	6	0	0	168590407	25496149	115053	6	0	0
ReadsMap	115308436	47132457	2282147	61828	135	1	142468047	54354636	2811288	104747	158	1
SMALT	198636832	6370637	0	0	0	0	200297027	6370637	0	0	0	0
STAR 1-pass	171619910	31320692	260829	357	0	0	189704784	33803026	272426	374	0	0
STAR 1-pass ann	153632518	48484160	1426139	35255	16	0	168619810	52590167	1583953	37982	20	0
STAR 2-pass	150202827	51651315	1653897	39505	68	0	168220716	62174408	2662407	52111	164	0
STAR 2-pass ann	150004846	51809144	1671503	40203	70	0	168011972	62573092	2707290	53504	168	0
TopHat1	155458487	31685988	1321699	4249	2	0	169288810	32281213	1508011	7658	6	0
TopHat1 ann	155458440	31765329	1311500	7347	9	0	169300835	32399210	1499216	12984	14	0
TopHat2	147253729	36710072	1204310	6029	3	0	163857967	38962028	1249600	6498	4	0
TopHat2 ann	140335591	46878285	1637296	27952	23	0	154155848	51536690	1841780	55839	46	0
B. K562 whole cell replicate 2												
BAGET ann	197805101	17665445	0	0	0	0	204566639	17665445	0	0	0	0
GEM ann	185275025	33132187	851569	15660	0	0	539610202	69130436	8433960	388289	215	0
GEM cons	185615881	32821350	795067	209	0	0	533876729	61377573	4873485	204158	2170	0
GEM cons ann	185536089	32882879	833009	861	0	0	535319029	63471618	5318079	228545	2170	0
GSNAP	182769289	36179016	364673	2077	8	0	246051100	38691591	373129	2257	10	0
GSNAP ann	167694005	50170089	1526399	11917	25	0	229895554	53925349	1628445	13450	27	0
GSTRUCT	173861658	44234621	1412799	9059	22	0	262536766	54202009	1640658	10504	22	0
GSTRUCT ann	172181493	45847374	1466966	10111	28	0	254420149	54043823	1657727	11527	28	0
MapSplice	161505723	36819919	910651	1197	0	0	167168530	36819931	910651	1197	0	0
MapSplice ann	161906407	36398227	886835	1465	0	0	167579746	36398233	886835	1465	0	0
PALMapper	164821984	44270272	3460350	0	0	0	1692507126	102137645	11501769	0	0	0
PALMapper cons	107664915	15415151	67570	0	0	0	107664915	15415151	67570	0	0	0
PASS	176544216	22968986	102988	3	0	0	178806957	23480856	103265	6	0	0
PASS cons	165095387	22162161	102235	4	0	0	166432987	22527135	102409	6	0	0
SMALT	209130532	6127506	0	0	0	0	210307069	6127506	0	0	0	0
STAR 1-pass	181284498	30131002	236178	252	0	0	201282547	32495226	246466	265	0	0
STAR 1-pass ann	163348660	48075196	1316387	37029	6	0	180032268	52109009	1454382	40190	13	0
STAR 2-pass	159706127	51458084	1556531	42043	41	0	179508185	62397070	2526198	55231	116	0
STAR 2-pass ann	159514745	51624217	1571320	42549	43	0	179324495	62854431	2567410	56279	121	0
TopHat1	149782758	29146862	1049421	4511	28	0	164950835	29832399	1242963	19291	56	1
TopHat1 ann	149783322	29183515	1072472	6439	24	0	164962013	29912688	1282587	24737	42	0
TopHat2	139741709	33321108	1028619	5325	0	0	156890641	35541225	1062079	5479	0	0
TopHat2 ann	132677412	42947041	1435715	29132	6	0	146628377	46898096	1615166	56043	7	0
C. K562 cytoplasmic fraction replicate 1												
BAGET ann	208521101	28241977	0	0	0	0	212070658	28241977	0	0	0	0
GEM ann	197702900	40606845	1134440	1230	0	0	493631577	77253508	9906030	216131	1004	0
GEM cons	197967587	40352093	1087093	190	0	0	488624837	69430179	7101212	101765	426	0
GEM cons ann	197887670	40413389	1125713	1137	0	0	489553614	71590585	7396418	94242	162	0
GSNAP	195321629	43621344	451596	2673	15	0	250621888	46533423	460979	2806	17	0
GSNAP ann	179534877	58044452	1822973	9814	24	0	233740631	62778133	1962993	10388	24	0
GSTRUCT	188613775	49280508	1717291	8100	26	0	257506322	64135677	2357715	9061	26	0
GSTRUCT ann	185601418	52112656	1894798	9080	28	0	249941538	64060885	2380522	9518	28	0
MapSplice	180858050	49673933	1318983	1094	0	0	182639049	49673956	1318983	1094	0	0
MapSplice ann	181889965	48648848	1276719	2309	0	0	183671968	48648855	1276719	2309	0	0
PASS	200240345	31338049	142379	7	0	0	201821496	31749762	142891	7	0	0
PASS cons	196768072	30949027	141723	6	0	0	197592326	31222304	142046	6	0	0
ReadsMap	140851097	60522677	2872481	107427	28	0	178414125	69875721	3461753	234391	39	0
SMALT	220887722	7544213	0	0	0	0	221684991	7544213	0	0	0	0
STAR 1-pass	195727486	37897252	324449	373	0	0	215758379	40394998	335875	386	0	0
STAR 1-pass ann	173885738	58529123	1736876	127424	11	0	191298592	63538256	1915457	129617	12	0
STAR 2-pass	166816809	65075051	2176309	136258	54	1	190222297	81106244	3686507	156631	159	2
STAR 2-pass ann	166571761	65289207	2186183	137206	58	1	190002411	81790333	3749852	158565	163	2
TopHat1	179661623	36831248	1344808	6148	3	0	195310467	37339628	1491492	13720	9	0
TopHat1 ann	179660581	36913991	1351375	6896	4	0	195359319	37471945	1492452	12005	9	0
TopHat2	170911303	44935030	1410915	5034	0	0	195605668	49344971	1519378	6374	0	0
TopHat2 ann	161001926	57456987	2047066	116223	34	0	181974554	64861853	2429409	212221	51	0
D. K562 cytoplasmic fraction replicate 2												
BAGET ann	146679343	18913128	0	0	0	0	149493992	18913128	0	0	0	0
GEM ann	133664041	33050924	878073	1183	0	0	364341755	69827925	9762875	283774	227	0
GEM cons	133880600	32852609	832767	227	0	0	360205170	63087139	7306424	116213	60	0
GEM cons ann	133794150	32919330	871043	968	0	0	361093746	65076885	7579619	117473	75	0
GSNAP	132905278	34363515	337927	1960	3	0	165065576	36381786	343986	2063	3	0
GSNAP ann	121020777	45309882	1287171	6066	16	0	152256757	48498217	1357665	6492	16	0
GSTRUCT	126275712	40116144	1346547	5748	18	0	163482441	49159012	1687419	6252	18	0

Nature Methods: doi:10.1038/nmeth.2722

	Primary alignments						All alignments					
	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns
GSTRUCT ann	124224907	42031798	1480673	5994	22	0	157225878	49052558	1718690	6442	22	0
MapSplice	122865216	39932973	961603	642	0	0	123668821	39932980	961603	642	0	0
MapSplice ann	123621200	39143336	939761	1171	0	0	124425977	39143339	939761	1171	0	0
PASS	136440767	25599042	103327	2	0	0	137655948	25954539	103689	2	0	0
PASS cons	133861132	25334401	102917	2	0	0	134571977	25602161	103163	2	0	0
SMALT	154486033	6245646	0	0	0	0	155123972	6245646	0	0	0	0
STAR 1-pass	133548716	29927431	240562	341	0	0	146845730	32174114	249305	345	0	0
STAR 1-pass ann	117478919	45195504	1209188	49421	8	0	129335370	49198277	1335727	50591	9	0
STAR 2-pass	113072761	49193805	1555212	54778	50	0	128635253	60047230	2680991	67904	86	0
STAR 2-pass ann	112880538	49349458	1563581	55596	53	0	128413460	60494713	2731292	69343	90	0
TopHat1	120392290	30257696	1035799	4436	184	0	130952871	30719274	1157794	10051	2135	0
TopHat1 ann	120391265	30338687	1041674	5541	142	0	130960638	30820610	1166744	11262	1586	0
TopHat2	113472066	36323162	1078664	3395	1	0	130552639	39703856	1187823	3773	6	0
TopHat2 ann	108225194	44237331	1446875	15471	18	0	124134286	49790480	1737874	22532	18	0
E. K562 nuclear fraction replicate 1												
BAGET ann	206705594	15431582	0	0	0	0	210513852	15431582	0	0	0	0
GEM ann	197660121	25731659	603640	457	1	0	538778273	55458548	6666902	142307	564	0
GEM cons	198011964	25376373	571676	143	0	0	534612925	48611961	3671708	58066	18	0
GEM cons ann	197935331	25446958	591336	397	0	0	535168811	49677122	3944727	62174	18	0
GSNAP	198639580	24590060	262861	1334	16	0	256546541	26582431	274731	1408	16	0
GSNAP ann	190841927	31745811	918286	6212	26	0	248218509	34499476	1018447	6398	26	0
GSTRUCT	191751353	30929967	925084	4376	24	0	234821855	34917124	1039863	5648	25	0
GSTRUCT ann	191009154	31640474	955681	4636	23	0	232797536	34747221	1045346	5741	23	0
MapSplice	188915608	27849966	693371	1118	2	0	190811556	27849975	693371	1118	2	0
MapSplice ann	188993965	27766670	696752	975	8	0	190883732	27766675	696752	975	8	0
PALMapper	186629755	30952119	1738721	0	0	0	2554195311	69199335	4910535	0	0	0
PALMapper cons	130671178	12853961	42233	0	0	0	173449918	15341457	51658	0	0	0
PASS	195909426	21349613	88759	11	0	0	198364701	21972613	89174	11	0	0
PASS cons	192644938	20923625	88109	7	0	0	194283800	21322859	88327	7	0	0
ReadsMap	141145652	37445251	2033113	54916	516	0	167139782	43692129	2500076	74332	699	0
SMALT	213355159	5407107	0	0	0	0	214209294	5407107	0	0	0	0
STAR 1-pass	195898559	21759750	167482	119	0	0	212012843	23263830	184064	133	0	0
STAR 1-pass ann	186225918	30943289	827329	13566	2	0	201653603	33228051	945262	16152	2	0
STAR 2-pass	183252391	33771914	974031	17624	25	1	200881258	40044714	1515066	30013	62	1
STAR 2-pass ann	183132331	33871096	981924	17919	25	1	200789609	40384352	1543146	30642	64	1
TopHat1	180403610	23968166	795467	4044	18	0	194534527	24543719	907738	7399	51	0
TopHat1 ann	180403717	24002859	809383	4819	18	0	194547983	24605666	924258	8167	57	0
TopHat2	176258365	27079714	765656	3311	1	0	195355523	29138021	812726	4028	1	0
TopHat2 ann	173129155	30494610	966483	12655	15	0	188174689	33348476	1105695	21125	22	0
F. K562 nuclear fraction replicate 2												
BAGET ann	183216474	13830210	0	0	0	0	186375573	13830210	0	0	0	0
GEM ann	174331871	23282063	510180	487	0	0	495481257	49632919	5445987	144699	723	0
GEM cons	174618785	22998748	475199	115	0	0	492567840	44678365	2919699	68430	7	0
GEM cons ann	174535471	23073017	497136	453	0	0	493060432	45622893	3173625	71260	30	0
GSNAP	175703134	21897666	218107	1175	18	0	228881337	23886945	228125	1219	21	0
GSNAP ann	169165577	27931661	732314	4283	47	0	221805293	30627969	813255	4383	50	0
GSTRUCT	169585880	27600513	745440	3448	45	0	209168236	30808839	819232	4025	46	0
GSTRUCT ann	169136046	28030982	763899	3681	46	0	207369719	30707614	815426	4146	46	0
MapSplice	167648664	25235898	579628	696	0	0	169001827	25235908	579628	696	0	0
MapSplice ann	167680834	25198102	583636	875	0	0	169027659	25198107	583636	875	0	0
PALMapper	164784505	27454461	1464692	0	0	0	2648436542	59977471	3825122	0	0	0
PALMapper cons	125464945	12642822	36162	0	0	0	169630360	15090565	44023	0	0	0
PASS	172713017	20332388	82128	7	0	0	175158312	20940060	82626	9	0	0
PASS cons	170121237	19986490	81579	6	0	0	171768050	20398941	81940	7	0	0
SMALT	190114114	5057768	0	0	0	0	190958188	5057768	0	0	0	0
STAR 1-pass	173679550	19741734	144171	139	0	0	188503852	21223846	159341	143	0	0
STAR 1-pass ann	165699327	27329037	664747	6623	16	0	180111098	29433467	761984	8159	28	0
STAR 2-pass	163151433	29763532	782416	10236	39	0	179388811	35235658	1224137	17880	70	0
STAR 2-pass ann	163046387	29852674	788567	10554	40	0	179313087	35576486	1246049	18437	77	0
TopHat1	159567534	22051237	652260	3266	8	0	170842949	22579102	740729	6198	39	0
TopHat1 ann	159567334	22085062	668529	4089	11	0	170850087	22632500	760666	6961	44	0
TopHat2	157907882	24910089	648982	2180	8	0	173381001	26705557	690659	2610	8	0
TopHat2 ann	155678376	27241854	788145	6765	30	0	167766127	29760695	899188	9923	32	0
G. Mouse brain												
BAGET ann	103801100	5664502	0	0	0	0	106626404	5664502	0	0	0	0
GEM ann	104783501	7457537	221379	1310	0	0	1680600227	13138441	651413	4024	2	0
GEM cons	104968184	7288614	186078	2	0	0	1680148684	11536810	433152	47	0	0
GEM cons ann	104887829	7350261	219901	1306	0	0	1680274464	11796821	525613	3358	2	0
GSNAP	103644953	6862434	104655	1095	12	0	184900539	7565613	105529	1102	12	0
GSNAP ann	101540035	8778287	296948	3678	72	0	182680583	9705444	311914	3688	72	0
GSTRUCT	101873894	9269701	294801	3204	54	0	172915091	10623869	323401	3443	54	0
GSTRUCT ann	101706267	9423372	300023	3347	58	0	173240656	10703268	325579	3589	58	0
MapSplice	99314039	7907439	230820	261	0	0	103798423	7907441	230820	261	0	0
MapSplice ann	99269151	7976595	245448	1465	1	0	103777468	7976598	245448	1465	1	0
PASS	99405349	6668953	38963	8	0	0	102761515	6822469	39213	8	0	0
PASS cons	96704058	6527373	38602	8	0	0	97577969	6625437	38765	8	0	0
ReadsMap	71638776	11207219	435295	3756	59	0	105324775	12711034	558163	6808	59	0
SMALT	103786857	1723772	0	0	0	0	104394448	1723772	0	0	0	0
STAR 1-pass	95776345	6210490	69044	209	0	0	103446176	6367339	69751	210	0	0
STAR 1-pass ann	94157652	7746192	185813	1914	47	0	101968585	8188551	192088	1961	53	0
STAR 2-pass	92751597	9109186	283715	4850	72	0	100940213	10100321	309283	6020	75	0
STAR 2-pass ann	92573028	9304825	295568	5523	82	0	100633165	10421178	325112	6749	94	0
TopHat1	89822037	7075370	224883	1230	0	0	97517105	7113106	233819	1359	0	0

Nature Methods: doi:10.1038/nmeth.2722

	Primary alignments						All alignments					
	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns	0 introns	1 intron	2 introns	3 introns	4 introns	5 introns
TopHat1 ann	89822004	7109187	244436	2428	62	0	97520713	7159649	253887	2588	64	0
TopHat2	89216210	7853672	258284	1668	0	0	101874048	8062729	262267	1711	0	0
TopHat2 ann	88490497	8865297	321994	4119	88	0	98587259	9357618	341839	4300	88	0
H. Simulation 1												
BAGET ann	71619289	7243380	0	0	0	0	72438310	7243380	0	0	0	0
GEM ann	67895614	11620668	404930	2038	0	0	165877216	21162135	1610192	11556	22	0
GEM cons	67993410	11540114	369445	285	0	0	165099634	19616501	1116746	2305	0	0
GEM cons ann	67923682	11592880	400762	1902	0	0	165500190	20423970	1355142	7442	19	0
GSNAP	68655070	10552570	183396	2029	19	0	80840006	10999735	187525	2031	19	0
GSNAP ann	65523836	13357541	511492	8496	193	0	77578022	13982407	536732	8498	193	0
GSTRUCT	65645586	13235700	517317	8182	173	0	72687872	13849148	531732	8207	173	0
GSTRUCT ann	65455071	13417975	524927	8385	173	0	71751504	13774977	531934	8388	173	0
MapSplice	65759487	12673956	450754	2255	7	0	71522223	12682474	450844	2255	7	0
MapSplice ann	65654189	12822124	466648	2542	8	0	71412112	12829010	466776	2542	8	0
PALMapper	65223152	13245519	211792	74	8	0	650778760	34807250	296256	276	21	0
PALMapper ann	64762109	13687402	285014	60	8	0	649459058	37952667	447335	276	21	0
PALMapper cons	58013224	6615783	17596	0	0	0	86278432	8411815	19808	0	0	0
PALMapper cons ann	66026622	11986230	182774	10	0	0	308967867	19314514	225714	32	2	0
PASS	67661818	9840645	71539	27	0	0	68505088	9924704	71705	27	0	0
PASS cons	66930406	9789895	71387	27	0	0	67528822	9830463	71431	27	0	0
ReadsMap	53769449	15043852	1461515	121373	4855	30	60544097	15956506	1583285	132148	5132	30
SMALT	74725141	2655012	0	0	0	0	74854021	2655012	0	0	0	0
STAR 1-pass	69477331	9422613	116704	346	0	0	73248740	9807635	121345	346	0	0
STAR 1-pass ann	65688823	12927586	452392	8511	228	0	69505024	13655950	495790	9510	233	1
STAR 2-pass	65533283	13070836	472720	8758	209	0	69397907	13798019	530437	10221	226	0
STAR 2-pass ann	65396794	13190256	486595	9129	220	0	69287811	14134611	559707	10894	238	1
TopHat1	64563528	11337574	443740	6547	41	0	66733540	11483082	465724	6915	41	0
TopHat1 ann	64560841	11429576	465337	7337	142	0	66734561	11583780	488154	7702	153	0
TopHat2	62535958	12139147	483319	7601	45	0	66201271	12550205	500929	7705	45	0
TopHat2 ann	61436183	13067748	555565	10513	292	0	64401859	13755211	606975	11567	295	0
I. Simulation 2												
BAGET ann	70075321	7339683	0	0	0	0	72248546	7339683	0	0	0	0
GEM ann	68175403	11000317	343047	2378	5	0	168057919	21101019	1893735	14365	32	0
GEM cons	68381701	10800863	302917	436	0	0	167084140	19138153	1371492	5152	0	0
GEM cons ann	68230123	10945100	337369	2119	5	0	167467722	19973685	1596525	9704	16	0
GSNAP	67502298	10688746	170716	1275	13	0	80649602	11221314	176598	1283	13	0
GSNAP ann	63884518	13975304	512648	7390	102	2	76939310	14761371	554446	7409	102	2
GSTRUCT	64091183	13880674	509423	6581	62	1	71586151	14577724	535411	6624	62	1
GSTRUCT ann	63903567	14056596	518470	6965	95	1	70899214	14590806	538133	6975	95	1
MapSplice	63972131	11366972	350883	1924	6	0	69298637	11373139	351091	1924	6	0
MapSplice ann	63736012	11730852	358910	2314	14	0	69046142	11734861	359129	2318	14	0
PALMapper	63372043	13871065	180109	48	2	0	615694813	43911626	306150	335	14	0
PALMapper ann	62590372	14732160	272967	45	2	0	615958105	51083521	562036	335	14	0
PALMapper cons	60732770	7062992	19571	0	0	0	168635025	9846617	22824	0	0	0
PALMapper cons ann	62858130	12438785	157020	3	0	0	276166832	21541293	230330	5	0	0
PASS	63588965	8487179	45659	11	0	0	64555002	8590501	45820	11	0	0
PASS cons	61639107	8302610	45296	11	0	0	62249886	8347417	45362	11	0	0
ReadsMap	53673659	14854263	649065	12091	143	0	61876899	16658795	776727	13173	159	0
SMALT	74549890	2520940	0	0	0	0	74761393	2520940	0	0	0	0
STAR 1-pass	68751704	8161114	74265	83	0	0	73570492	8673633	80486	85	0	0
STAR 1-pass ann	64558922	12434337	369404	5882	83	0	69221948	13312964	417012	6571	85	0
STAR 2-pass	64293445	12714005	401107	6073	70	0	69034173	13549823	460189	7024	79	0
STAR 2-pass ann	64051280	12942089	419355	6527	79	0	68779769	14066697	497348	7696	88	0
TopHat1	57899907	10586748	382874	4434	1	0	60653984	10893767	425260	5195	2	0
TopHat1 ann	57893669	10914044	412454	5855	44	0	60648703	11232934	458045	6815	47	0
TopHat2	52071715	9893020	370824	4710	8	0	56467249	10487685	397883	4826	8	0
TopHat2 ann	51775131	11460488	470507	8668	155	0	55456172	12423060	552738	9462	180	0

There were no alignments with more than five introns.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 11. Accuracy of multi-intron alignments.

	Recall					Precision				
	≥ 1 introns	≥ 2 introns	≥ 3 introns	≥ 4 introns	≥ 5 introns	≥ 1 introns	≥ 2 introns	≥ 3 introns	≥ 4 introns	≥ 5 introns
A. Simulation 1										
BAGET ann	39.5%	0.0%	0.0%	0.0%	0.0%	78.7%	n.a.	n.a.	n.a.	n.a.
GEM ann	82.2%	65.2%	15.8%	0.0%	0.0%	98.6%	98.3%	97.6%	n.a.	n.a.
GEM cons	81.6%	59.7%	2.3%	0.0%	0.0%	98.9%	98.7%	97.5%	n.a.	n.a.
GEM cons ann	82.0%	64.6%	14.8%	0.0%	0.0%	98.7%	98.5%	98.8%	n.a.	n.a.
GSNAP	73.6%	29.1%	11.4%	3.1%	0.0%	98.8%	95.7%	68.5%	89.5%	n.a.
GSNAP ann	94.9%	82.2%	67.6%	34.2%	0.0%	98.7%	96.5%	95.9%	96.9%	n.a.
GSTRUCT	94.1%	83.9%	65.8%	30.7%	0.0%	98.7%	97.5%	97.1%	97.1%	n.a.
GSTRUCT ann	95.6%	85.2%	66.2%	30.7%	0.0%	98.9%	97.5%	95.3%	97.1%	n.a.
MapSplice	90.3%	72.5%	17.9%	1.3%	0.0%	99.3%	97.8%	97.7%	100.0%	n.a.
MapSplice ann	90.0%	74.1%	19.2%	1.3%	0.0%	97.7%	96.4%	93.0%	87.5%	n.a.
PALMapper	85.4%	24.4%	0.1%	0.0%	0.0%	91.8%	70.4%	19.5%	0.0%	n.a.
PALMapper ann	87.2%	36.5%	0.1%	0.0%	0.0%	90.3%	78.3%	20.6%	0.0%	n.a.
PALMapper cons	45.0%	2.8%	0.0%	0.0%	0.0%	98.0%	98.2%	n.a.	n.a.	n.a.
PALMapper cons ann	78.6%	28.7%	0.1%	0.0%	0.0%	93.4%	95.7%	80.0%	n.a.	n.a.
PASS	66.7%	11.3%	0.2%	0.0%	0.0%	97.0%	96.7%	100.0%	n.a.	n.a.
PASS cons	66.6%	11.2%	0.2%	0.0%	0.0%	97.5%	96.8%	100.0%	n.a.	n.a.
ReadsMap	89.7%	82.7%	72.1%	40.8%	0.0%	77.8%	31.8%	7.0%	4.6%	0.0%
SMALT	6.2%	0.0%	0.0%	0.0%	0.0%	33.5%	n.a.	n.a.	n.a.	n.a.
STAR 1-pass	65.4%	19.0%	2.8%	0.0%	0.0%	99.0%	99.0%	99.4%	n.a.	n.a.
STAR 1-pass ann	90.7%	71.9%	64.7%	40.2%	0.0%	97.8%	95.4%	91.3%	96.5%	n.a.
STAR 2-pass	92.6%	76.8%	70.7%	38.2%	0.0%	98.6%	97.3%	97.2%	100.0%	n.a.
STAR 2-pass ann	93.0%	78.4%	72.4%	40.0%	0.0%	98.1%	96.7%	95.6%	99.5%	n.a.
TopHat1	80.0%	67.4%	47.6%	7.5%	0.0%	98.0%	91.4%	89.2%	100.0%	n.a.
TopHat1 ann	80.6%	70.2%	54.3%	21.6%	0.0%	97.8%	90.7%	89.4%	83.1%	n.a.
TopHat2	86.0%	78.3%	60.9%	8.2%	0.0%	98.2%	97.4%	98.2%	100.0%	n.a.
TopHat2 ann	92.3%	87.4%	79.3%	52.7%	0.0%	97.7%	94.5%	92.4%	98.6%	n.a.
Number of simulated reads	13808336	598297	11781	493	54					
B. Simulation 2										
BAGET ann	37.9%	0.0%	0.0%	0.0%	0.0%	80.5%	n.a.	n.a.	n.a.	n.a.
GEM ann	70.7%	50.8%	13.8%	0.0%	0.0%	97.4%	94.5%	86.4%	0.0%	n.a.
GEM cons	69.4%	45.2%	2.9%	0.0%	0.0%	97.8%	95.4%	82.2%	n.a.	n.a.
GEM cons ann	70.5%	50.2%	13.2%	0.0%	0.0%	97.7%	95.1%	94.3%	0.0%	n.a.
GSNAP	68.3%	24.2%	9.1%	4.0%	0.0%	98.0%	89.1%	84.9%	84.6%	n.a.
GSNAP ann	91.0%	76.4%	56.1%	33.0%	0.0%	98.0%	93.2%	89.5%	85.7%	0.0%
GSTRUCT	90.3%	78.4%	53.1%	18.3%	0.0%	97.8%	96.5%	95.7%	78.1%	0.0%
GSTRUCT ann	91.8%	80.0%	56.6%	29.7%	0.0%	98.2%	96.6%	95.9%	83.5%	0.0%
MapSplice	73.8%	51.1%	15.2%	2.2%	0.0%	98.2%	92.0%	94.1%	100.0%	n.a.
MapSplice ann	74.3%	51.2%	17.6%	0.7%	0.0%	95.9%	90.0%	90.5%	14.3%	n.a.
PALMapper	79.0%	17.1%	0.0%	0.0%	0.0%	87.7%	60.2%	10.0%	0.0%	n.a.
PALMapper ann	82.7%	28.3%	0.0%	0.0%	0.0%	85.9%	65.9%	10.6%	0.0%	n.a.
PALMapper cons	43.5%	3.0%	0.0%	0.0%	0.0%	95.7%	97.1%	n.a.	n.a.	n.a.
PALMapper cons ann	72.4%	22.0%	0.0%	0.0%	0.0%	89.6%	88.8%	66.7%	n.a.	n.a.
PASS	50.8%	6.6%	0.1%	0.0%	0.0%	93.3%	93.9%	100.0%	n.a.	n.a.
PASS cons	50.4%	6.6%	0.1%	0.0%	0.0%	94.5%	94.2%	100.0%	n.a.	n.a.
ReadsMap	76.2%	67.8%	56.3%	30.8%	0.0%	76.6%	65.1%	55.1%	58.7%	n.a.
SMALT	5.2%	0.0%	0.0%	0.0%	0.0%	32.1%	n.a.	n.a.	n.a.	n.a.
STAR 1-pass	51.8%	11.3%	0.7%	0.0%	0.0%	98.1%	96.3%	94.0%	n.a.	n.a.
STAR 1-pass ann	79.3%	54.1%	44.3%	28.2%	0.0%	96.6%	91.8%	91.3%	92.8%	n.a.
STAR 2-pass	81.8%	59.9%	47.7%	25.6%	0.0%	97.3%	93.5%	93.0%	100.0%	n.a.
STAR 2-pass ann	83.1%	61.8%	48.5%	28.9%	0.0%	96.9%	92.4%	90.1%	100.0%	n.a.
TopHat1	68.3%	54.9%	32.8%	0.0%	0.0%	97.0%	89.9%	88.6%	0.0%	n.a.
TopHat1 ann	70.5%	58.8%	44.0%	14.7%	0.0%	97.0%	89.3%	89.3%	90.9%	n.a.
TopHat2	64.0%	56.6%	37.7%	0.0%	0.0%	97.2%	95.7%	95.7%	0.0%	n.a.
TopHat2 ann	73.8%	68.7%	64.0%	40.3%	0.0%	96.5%	91.4%	88.8%	71.0%	n.a.
Number of simulated reads	14962090	622980	11701	270	3					

For each intron count *n*, the tabulated percentages were computed as follows:
recall = number of primary alignments with at least *n* correctly identified introns / number of simulated reads with at least *n* introns;
precision = number of primary alignments with at least *n* correctly identified introns / number of primary alignments with at least *n* reported introns.
The number of simulated reads with *n* introns is given on the last row of each table. Precision is n.a. (not applicable) where no aligments were reported.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 12. Transcript reconstruction accuracy.

	Exon recall			Exon precision			Spliced transcript recall			Spliced transcript precision		
	All	Known	Novel	All	Known	Novel	All	Known	Novel	All	Known	Novel
A. Simulation 1												
BAGET ann	76.5%	81.6%	6.5%	59.6%	59.5%	0.8%	12.0%	38.6%	3.8%	25.4%	20.5%	7.7%
GEM ann	81.8%	84.1%	50.5%	77.4%	76.7%	12.3%	15.5%	39.6%	8.1%	29.2%	19.8%	14.2%
GEM cons	74.0%	76.2%	44.1%	74.6%	73.9%	10.4%	12.7%	31.8%	6.8%	24.9%	16.4%	12.0%
GEM cons ann	81.1%	83.8%	44.5%	77.1%	76.5%	10.9%	15.3%	40.3%	7.6%	29.3%	20.4%	13.6%
GSNAP	82.0%	84.1%	54.2%	78.9%	78.2%	14.1%	16.1%	40.0%	8.7%	30.4%	20.3%	15.4%
GSNAP ann	83.1%	85.3%	53.3%	82.3%	81.7%	16.6%	18.0%	45.6%	9.5%	35.9%	25.0%	18.5%
GSTRUCT	83.0%	85.0%	55.7%	81.5%	80.8%	16.4%	17.7%	43.7%	9.7%	34.9%	23.7%	18.4%
GSTRUCT ann	83.2%	85.3%	55.2%	82.2%	81.5%	16.9%	18.0%	44.8%	9.8%	35.6%	24.4%	18.7%
MapSplice	81.3%	83.3%	54.0%	80.5%	79.8%	15.4%	16.1%	40.0%	8.8%	32.6%	22.0%	16.8%
MapSplice ann	82.0%	84.5%	47.6%	80.7%	80.1%	13.9%	16.3%	42.5%	8.3%	33.3%	23.4%	16.3%
PALMapper	82.5%	84.6%	54.0%	66.2%	65.2%	7.8%	15.0%	37.1%	8.3%	28.0%	18.4%	14.1%
PALMapper ann	83.1%	85.3%	53.7%	66.3%	65.3%	7.7%	14.3%	35.5%	7.8%	26.5%	17.4%	13.1%
PALMapper cons	78.2%	80.4%	47.2%	59.2%	58.2%	5.5%	13.2%	32.2%	7.3%	25.2%	16.2%	12.6%
PALMapper cons ann	80.6%	82.7%	50.7%	64.0%	63.0%	6.9%	15.5%	38.3%	8.5%	31.4%	20.9%	16.2%
PASS	64.3%	66.3%	36.5%	41.6%	40.7%	2.6%	9.3%	23.5%	4.9%	14.1%	8.9%	6.3%
PASS cons	64.6%	66.6%	37.0%	42.5%	41.6%	2.8%	9.3%	23.4%	5.0%	14.3%	9.0%	6.4%
ReadsMap	72.6%	74.5%	46.7%	54.1%	53.0%	4.8%	13.0%	31.4%	7.3%	21.7%	13.6%	10.7%
SMALT	21.6%	22.2%	14.3%	21.9%	21.2%	1.1%	1.0%	2.0%	0.6%	1.7%	0.9%	0.9%
STAR 1-pass	80.3%	82.3%	53.3%	77.1%	76.3%	12.9%	14.4%	36.2%	7.8%	26.2%	17.3%	12.7%
STAR 1-pass ann	83.9%	86.2%	52.8%	79.9%	79.2%	14.3%	17.8%	46.4%	9.0%	34.6%	24.4%	17.0%
STAR 2-pass	82.4%	84.3%	55.7%	80.0%	79.2%	15.3%	16.8%	41.5%	9.2%	32.1%	21.5%	16.6%
STAR 2-pass ann	84.1%	86.2%	55.1%	80.0%	79.3%	14.9%	17.6%	44.8%	9.2%	33.5%	23.2%	16.9%
TopHat1	77.6%	79.8%	46.9%	78.0%	77.3%	12.4%	14.0%	35.4%	7.4%	27.0%	18.0%	13.0%
TopHat1 ann	81.4%	83.9%	47.0%	79.8%	79.2%	13.2%	16.2%	42.4%	8.2%	31.6%	22.1%	15.2%
TopHat2	78.7%	80.9%	47.7%	81.3%	80.7%	14.9%	15.0%	38.0%	8.0%	30.9%	20.9%	15.4%
TopHat2 ann	83.6%	86.3%	46.5%	83.4%	82.8%	15.7%	17.9%	48.1%	8.7%	37.4%	27.3%	18.1%
Truth	86.0%	87.6%	65.2%	85.7%	85.1%	23.1%	19.9%	48.4%	11.2%	40.0%	27.5%	22.3%
B. Simulation 2												
BAGET ann	76.5%	81.7%	7.3%	56.8%	56.7%	0.7%	12.0%	38.2%	3.7%	24.7%	20.0%	7.2%
GEM ann	74.0%	76.4%	42.4%	66.6%	65.8%	7.0%	9.0%	21.0%	5.2%	14.6%	8.8%	7.0%
GEM cons	64.1%	66.2%	36.4%	62.4%	61.5%	5.8%	6.7%	14.3%	4.3%	11.1%	6.0%	5.8%
GEM cons ann	73.5%	76.2%	36.7%	66.2%	65.5%	6.1%	8.8%	21.7%	4.8%	14.5%	9.0%	6.6%
GSNAP	80.4%	82.6%	50.4%	70.9%	70.0%	9.2%	12.7%	30.1%	7.3%	21.0%	13.0%	10.4%
GSNAP ann	81.9%	84.3%	49.7%	76.2%	75.5%	11.5%	15.0%	36.5%	8.3%	26.1%	16.9%	12.9%
GSTRUCT	81.6%	83.8%	52.2%	75.6%	74.8%	11.7%	14.9%	35.1%	8.6%	25.6%	16.1%	13.1%
GSTRUCT ann	82.2%	84.5%	52.1%	76.1%	75.3%	11.9%	15.3%	35.9%	8.9%	26.2%	16.5%	13.5%
MapSplice	71.2%	73.3%	43.5%	70.0%	69.1%	8.7%	10.5%	24.1%	6.3%	19.1%	11.4%	9.7%
MapSplice ann	73.5%	76.1%	37.8%	71.4%	70.7%	7.8%	11.2%	27.1%	6.2%	20.5%	13.0%	9.8%
PALMapper	79.4%	81.7%	49.2%	61.0%	60.0%	6.0%	12.0%	28.3%	6.9%	20.2%	12.5%	10.0%
PALMapper ann	80.6%	82.9%	49.1%	62.3%	61.4%	6.2%	11.4%	27.8%	6.2%	19.6%	12.4%	9.2%
PALMapper cons	75.2%	77.7%	43.0%	55.0%	54.0%	4.4%	10.9%	25.0%	6.5%	19.9%	12.0%	10.1%
PALMapper cons ann	77.4%	79.9%	45.1%	59.4%	58.5%	5.4%	12.9%	30.9%	7.3%	25.1%	16.1%	12.6%
PASS	45.6%	47.1%	25.8%	34.4%	33.5%	2.0%	4.3%	9.7%	2.6%	6.5%	3.6%	3.1%
PASS cons	46.0%	47.5%	26.3%	35.1%	34.2%	2.1%	4.3%	9.7%	2.6%	6.5%	3.6%	3.1%
ReadsMap	67.4%	69.3%	42.7%	39.5%	38.4%	2.7%	9.3%	21.1%	5.7%	11.0%	6.2%	5.4%
SMALT	20.7%	21.2%	14.0%	20.8%	20.2%	1.0%	0.9%	1.6%	0.7%	1.7%	0.7%	1.0%
STAR 1-pass	71.7%	73.8%	44.5%	67.2%	66.3%	7.7%	9.5%	21.1%	5.8%	15.0%	8.6%	7.7%
STAR 1-pass ann	80.8%	83.4%	45.3%	73.1%	72.4%	9.1%	14.4%	37.1%	7.3%	24.6%	16.7%	11.2%
STAR 2-pass	76.9%	79.1%	48.1%	71.9%	71.0%	9.6%	12.2%	28.3%	7.1%	20.5%	12.5%	10.3%
STAR 2-pass ann	80.5%	82.9%	48.4%	73.3%	72.5%	9.9%	13.7%	33.4%	7.6%	23.2%	14.9%	11.2%
TopHat1	69.6%	71.7%	41.4%	70.1%	69.3%	8.5%	9.7%	21.5%	6.1%	16.2%	9.2%	8.4%
TopHat1 ann	76.9%	79.5%	42.1%	74.8%	74.1%	9.8%	12.8%	31.5%	7.0%	22.1%	14.2%	10.5%
TopHat2	72.6%	74.9%	42.3%	74.1%	73.4%	10.0%	11.1%	25.2%	6.7%	19.6%	11.6%	10.1%
TopHat2 ann	82.1%	85.1%	41.9%	79.9%	79.4%	12.0%	16.0%	41.7%	8.0%	30.1%	21.1%	14.1%
Truth	85.9%	87.5%	63.6%	84.2%	83.6%	20.8%	18.2%	41.4%	11.0%	31.9%	20.2%	17.7%

The exons and transcripts constituting the simulated transcriptomes were classified as known or novel, depending whether they were included in the annotation provided to aligners. Note that lower accuracy for novel transcripts is expected even for protocols not using annotation, as the expression levels are lower for novel transcripts on average.

The precision estimates for known and novel features serve to assess the effect on precision when excluding a defined subset of matches. Precision for known features was computed as TP_{known} / (TP_{known} + FP), i.e. by excluding predictions matching novel transcripts. Similarly, precision for novel features was computed as TP_{novel} / (TP_{novel} + FP). These values should not be interpreted as absolute precision estimates, but in a relative manner, for comparison among methods.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Table 13. Cufflinks incorporation rates for exon junctions in alignments of simulated RNA-seq data.

	Junction type	Incorporated	Discarded	Percent incorporated	Percent incorporated, stratified by number of mappings supporting junction									
					1	2	3	4	5	6	7	8	9	10+
A. Simulation 1														
BAGET ann	True	71457	6067	92.2%	60.2%	79.1%	87.0%	89.2%	90.5%	90.4%	91.4%	92.0%	92.7%	95.8%
	False	2648	2496	51.5%	28.6%	40.8%	48.7%	58.1%	56.7%	59.9%	63.2%	58.7%	67.0%	78.7%
GEM ann	True	81780	34833	70.1%	22.2%	34.5%	41.9%	47.5%	47.2%	49.8%	52.3%	53.7%	52.6%	81.2%
	False	1681	7512	18.3%	7.8%	10.2%	12.9%	19.2%	22.4%	24.8%	25.7%	25.4%	39.4%	44.5%
GEM cons	True	73935	23765	75.7%	23.9%	32.9%	29.0%	36.4%	39.9%	45.9%	50.9%	50.5%	52.2%	81.4%
	False	1398	3291	29.8%	19.2%	18.7%	13.5%	20.8%	25.4%	26.6%	22.4%	25.7%	42.4%	47.4%
GEM cons ann	True	80343	27763	74.3%	29.8%	52.2%	55.9%	54.9%	50.3%	51.7%	53.6%	52.9%	53.4%	81.4%
	False	1635	4616	26.2%	14.7%	15.9%	15.8%	20.4%	23.8%	25.1%	24.8%	25.6%	44.1%	47.4%
GSNAP	True	81905	36925	68.9%	19.6%	31.9%	39.8%	44.3%	45.8%	49.3%	49.9%	54.2%	55.4%	82.4%
	False	879	12408	6.6%	1.7%	6.3%	7.2%	9.0%	14.0%	14.7%	15.4%	12.7%	17.0%	25.8%
GSNAP ann	True	82283	38534	68.1%	16.3%	28.5%	36.7%	42.0%	42.7%	45.1%	46.6%	49.6%	49.6%	80.5%
	False	697	17842	3.8%	1.5%	3.5%	5.3%	4.9%	3.9%	6.5%	4.3%	5.6%	7.7%	12.7%
GSTRUCT	True	82639	36945	69.1%	20.6%	33.4%	26.4%	38.9%	41.6%	46.0%	47.0%	52.9%	50.3%	81.0%
	False	624	8266	7.0%	1.5%	5.4%	9.0%	9.2%	11.4%	13.5%	11.1%	15.2%	15.7%	23.5%
GSTRUCT ann	True	82815	36964	69.1%	18.0%	24.2%	36.8%	42.7%	43.5%	47.0%	47.9%	52.5%	50.9%	80.9%
	False	667	7780	7.9%	1.5%	5.4%	9.1%	10.5%	11.5%	16.7%	14.5%	17.5%	20.3%	30.9%
MapSplice	True	80694	34995	69.8%	17.0%	30.1%	38.7%	43.1%	44.7%	47.8%	49.4%	49.9%	52.4%	80.3%
	False	613	3458	15.1%	3.6%	8.8%	18.6%	15.2%	16.9%	19.8%	19.0%	12.2%	26.5%	43.9%
MapSplice ann	True	81525	37515	68.5%	17.9%	30.8%	39.9%	43.0%	44.2%	47.0%	49.0%	49.8%	51.4%	80.1%
	False	943	21502	4.2%	0.9%	1.2%	3.7%	4.6%	4.6%	5.9%	5.0%	3.5%	8.5%	43.6%
PALMapper	True	81806	35404	69.8%	23.6%	35.5%	41.6%	45.2%	45.3%	48.1%	49.3%	51.2%	50.5%	80.7%
	False	6235	276801	2.2%	1.4%	1.8%	2.3%	3.3%	3.8%	5.0%	4.7%	5.4%	6.9%	16.2%
PALMapper ann	True	82171	36483	69.3%	23.0%	42.3%	47.4%	49.2%	45.1%	47.4%	49.7%	49.6%	50.4%	79.1%
	False	8092	317841	2.5%	1.6%	2.2%	2.9%	3.6%	3.8%	4.2%	4.3%	6.1%	6.0%	13.8%
PALMapper cons	True	79278	27075	74.5%	34.9%	42.8%	49.9%	51.1%	55.5%	59.0%	61.5%	62.5%	64.2%	86.4%
	False	2075	5197	28.5%	7.9%	15.4%	27.8%	35.8%	33.3%	43.6%	47.2%	49.3%	46.1%	58.8%
PALMapper cons ann	True	79822	28431	73.7%	24.9%	28.5%	40.3%	48.6%	48.5%	50.8%	52.9%	54.0%	53.6%	81.3%
	False	3475	39759	8.0%	3.0%	4.4%	6.3%	7.1%	8.4%	10.1%	9.0%	10.8%	11.6%	16.6%
PASS	True	70211	43803	61.6%	9.5%	19.6%	28.3%	34.5%	37.5%	40.2%	41.3%	45.8%	45.5%	77.6%
	False	1600	61005	2.6%	0.5%	1.7%	3.6%	4.0%	6.5%	4.4%	8.2%	8.5%	14.1%	25.4%
PASS cons	True	70269	43559	61.7%	9.3%	19.9%	28.1%	35.3%	37.5%	40.7%	41.9%	46.9%	46.7%	77.6%
	False	1425	35868	3.8%	0.6%	2.2%	4.0%	5.7%	7.7%	5.4%	9.4%	11.7%	14.8%	26.5%
ReadsMap	True	74211	39937	65.0%	19.1%	33.5%	39.6%	41.6%	42.7%	44.3%	47.1%	47.6%	46.7%	73.8%
	False	10531	888182	1.2%	0.1%	0.2%	0.3%	0.5%	0.7%	1.1%	1.4%	1.7%	2.0%	11.6%
SMALT	True	26213	24284	51.9%	28.5%	39.6%	46.5%	51.9%	54.6%	54.2%	57.3%	57.9%	59.9%	66.2%
	False	55687	84998	39.6%	14.6%	32.2%	42.7%	48.6%	51.0%	54.6%	57.1%	57.2%	57.9%	62.3%
STAR 1-pass	True	80600	35636	69.3%	22.0%	34.6%	42.5%	45.7%	48.5%	52.0%	52.0%	55.3%	57.4%	83.0%
	False	1219	5309	18.7%	3.4%	13.3%	18.9%	32.3%	31.3%	36.0%	35.3%	41.6%	37.7%	65.4%
STAR 1-pass ann	True	81623	37384	68.6%	15.7%	27.8%	36.0%	40.3%	42.5%	44.9%	45.2%	48.2%	51.1%	83.0%
	False	2010	18216	9.9%	3.3%	7.0%	7.9%	9.1%	12.2%	14.7%	13.5%	15.0%	17.3%	35.9%
STAR 2-pass	True	82229	34852	70.2%	16.8%	29.5%	38.5%	43.6%	45.5%	48.0%	48.3%	51.6%	52.6%	81.3%
	False	1105	10474	9.5%	2.5%	6.1%	7.9%	9.5%	15.3%	15.7%	16.5%	17.6%	17.7%	36.8%
STAR 2-pass ann	True	83680	35542	70.2%	18.1%	31.3%	39.4%	44.9%	45.1%	47.9%	48.7%	51.5%	53.1%	81.5%
	False	1820	19383	8.6%	2.7%	5.8%	7.2%	8.0%	8.7%	11.5%	14.4%	12.0%	15.3%	34.2%
TopHat1	True	77950	30829	71.7%	18.3%	31.0%	39.7%	45.2%	46.7%	49.1%	53.2%	52.6%	54.5%	81.0%
	False	1471	6238	19.1%	3.8%	7.8%	10.7%	12.7%	15.8%	18.3%	25.0%	21.9%	24.3%	41.3%
TopHat1 ann	True	81345	31835	71.9%	24.6%	38.9%	45.7%	50.2%	50.1%	51.2%	54.9%	54.0%	55.4%	81.3%
	False	1570	6803	18.8%	4.5%	9.0%	10.6%	14.3%	14.8%	15.2%	20.5%	22.4%	26.1%	38.3%
TopHat2	True	78218	31455	71.3%	16.3%	28.8%	37.2%	43.1%	44.8%	46.5%	49.2%	51.5%	52.9%	81.2%
	False	588	7303	7.5%	2.3%	3.8%	5.4%	8.1%	5.5%	9.9%	11.7%	11.8%	11.6%	14.0%
TopHat2 ann	True	82301	33644	71.0%	21.4%	35.9%	42.4%	46.1%	46.9%	48.4%	49.2%	51.2%	50.5%	80.9%
	False	1276	23060	5.2%	3.1%	4.5%	4.3%	4.6%	4.2%	5.3%	5.3%	7.3%	6.6%	13.0%
Truth	True	85827	36918	69.9%	17.8%	32.7%	40.6%	46.3%	46.9%	48.0%	48.4%	52.8%	53.1%	81.1%
	False	0	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B. Simulation 2														
BAGET ann	True	70765	6188	92.0%	58.1%	78.5%	86.2%	87.7%	89.3%	90.6%	90.0%	91.8%	91.9%	95.4%
	False	2751	3580	43.5%	21.0%	30.0%	42.7%	52.6%	53.6%	56.4%	52.1%	67.2%	71.7%	74.2%
GEM ann	True	78127	34232	69.5%	23.1%	38.2%	47.2%	51.3%	55.1%	57.9%	58.5%	59.7%	60.0%	77.3%
	False	3878	18415	17.4%	12.8%	15.5%	16.7%	17.1%	16.3%	17.3%	21.7%	17.3%	18.7%	31.3%
GEM cons	True	68279	23094	74.7%	20.3%	32.5%	35.9%	42.7%	50.8%	55.9%	56.9%	59.6%	58.8%	77.7%
	False	3623	8999	28.7%	28.7%	27.9%	23.9%	23.5%	21.3%	21.6%	25.9%	22.8%	19.6%	35.7%
GEM cons ann	True	76677	28738	72.7%	26.3%	47.4%	57.8%	60.0%	61.2%	62.3%	61.2%	60.5%	62.3%	77.8%
	False	3878	10904	26.2%	23.2%	24.3%	24.3%	22.3%	21.9%	20.7%	26.0%	22.1%	24.3%	36.1%
GSNAP	True	85566	33710	71.7%	21.1%	36.5%	46.0%	53.1%	56.8%	58.6%	59.8%	62.2%	63.1%	81.5%
	False	1508	29186	4.9%	1.3%	4.5%	9.2%	12.9%	13.8%	16.2%	19.5%	19.6%	20.2%	30.5%
GSNAP ann	True	86561	34859	71.3%	17.0%	32.3%	42.7%	50.2%	53.1%	54.8%	57.5%	59.1%	57.8%	80.2%
	False	1099	35540	3.0%	1.0%	3.5%	4.5%	5.8%	7.1%	7.1%	10.5%	8.5%	7.1%	14.1%
GSTRUCT	True	87072	32844	72.6%	20.7%	37.8%	32.9%	43.3%	51.0%	54.6%	58.7%	61.2%	59.3%	81.3%
	False	1223	21848	5.3%	1.0%	4.8%	6.8%	11.7%	15.4%	15.6%	17.7%	18.7%	18.4%	31.9%
GSTRUCT ann	True	87729	32486	73.0%	18.9%	29.0%	41.3%	48.9%	53.1%	56.0%	59.0%	61.9%	60.8%	81.4%
	False	1156	21457	5.1%	1.1%	5.1%	7.5%	13.8%	16.4%	20.3%	20.9%	21.6%	17.7%	31.1%
MapSplice	True	73923	35728	67.4%	15.5%	30.6%	40.8%	46.8%	51.3%	53.5%	54.3%	56.8%	57.2%	74.3%
	False	894	8412	9.6%	2.0%	7.2%	12.0%	14.5%	13.4%	9.1%	12.6%	16.9%	11.2%	25.4%
MapSplice ann	True	76680	39790	65.8%	17.9%	32.2%	42.0%	46.1%	50.8%	51.5%	53.2%	54.5%	56.1%	74.6%
	False	1901	32062	5.6%	0.9%	1.8%	3.3%	5.2%	6.1%	4.5%	4.8%	5.2%	6.3%	40.7%
PALMapper	True	82111	33574	71.0%	25.3%	39.2%	46.6%	51.3%	51.3%	54.3%	54.3%	59.0%	56.5%	78.6%
	False	8872	375045	2.3%	1.4%	1.8%	2.7%	3.3%	3.7%	4.8%	6.4%	6.1%	8.4%	18.0%
PALMapper ann	True	81970	36171	69.4%	25.7%	44.2%	49.3%	50.2%	50.0%	53.2%	55.1%	55.5%	53.7%	76.5%
	False	11705	516512	2.2%	1.5%	2.1%	2.9%	3.9%	3.9%	4.0%	4.8%	6.0%	6.0%	13.8%

Nature Methods: doi:10.1038/nmeth.2722

	Junction type	Incorporated	Discarded	Percent incorporated	Percent incorporated, stratified by number of mappings supporting junction									
					1	2	3	4	5	6	7	8	9	10+
PALMapper cons	True	79183	24754	76.2%	31.1%	44.0%	53.6%	57.8%	59.7%	63.1%	63.8%	66.6%	66.8%	83.3%
	False	3811	8450	31.1%	9.9%	17.0%	23.3%	27.1%	29.6%	36.3%	39.1%	39.8%	37.5%	52.0%
PALMapper cons ann	True	79606	26282	75.2%	21.3%	25.8%	35.8%	47.3%	55.0%	59.6%	60.7%	61.1%	64.8%	79.5%
	False	5166	53953	8.7%	3.0%	4.8%	6.5%	7.5%	9.3%	9.7%	9.7%	9.9%	11.1%	16.3%
PASS	True	55961	51872	51.9%	9.6%	19.0%	26.6%	30.9%	34.9%	36.4%	39.0%	40.3%	43.0%	63.4%
	False	2383	122909	1.9%	0.4%	1.2%	2.5%	3.6%	4.9%	5.6%	6.9%	8.2%	7.7%	18.1%
PASS cons	True	56849	50709	52.9%	9.9%	19.3%	26.2%	31.4%	35.7%	38.0%	39.8%	41.7%	43.7%	64.6%
	False	2186	75177	2.8%	0.5%	1.6%	2.9%	4.9%	6.0%	6.0%	8.5%	8.8%	9.3%	20.0%
ReadsMap	True	72489	36558	66.5%	16.1%	31.2%	42.3%	44.5%	48.9%	48.5%	50.5%	51.6%	53.4%	73.5%
	False	16565	926119	1.8%	0.1%	0.4%	0.7%	1.1%	1.7%	2.5%	3.3%	4.3%	5.9%	19.6%
SMALT	True	23924	26802	47.2%	27.9%	37.7%	41.7%	45.9%	47.6%	48.0%	50.6%	51.6%	50.4%	61.6%
	False	53865	127976	29.6%	10.9%	25.0%	34.5%	39.1%	43.3%	45.8%	47.5%	49.2%	49.6%	56.8%
STAR 1-pass	True	75728	34573	68.7%	23.4%	39.2%	48.4%	51.6%	55.1%	59.1%	60.0%	63.2%	62.0%	78.6%
	False	1880	13013	12.6%	3.0%	10.6%	15.3%	24.7%	22.8%	29.7%	37.1%	32.1%	44.2%	59.1%
STAR 1-pass ann	True	81935	34836	70.2%	17.3%	30.0%	39.3%	44.9%	49.7%	51.0%	54.7%	55.7%	57.0%	81.5%
	False	2504	29192	7.9%	2.4%	6.1%	8.7%	11.3%	9.4%	12.5%	11.5%	16.8%	13.5%	35.1%
STAR 2-pass	True	81369	31663	72.0%	17.2%	31.9%	41.8%	49.3%	54.6%	55.2%	59.2%	58.9%	61.1%	79.8%
	False	1687	21062	7.4%	1.8%	5.6%	8.1%	11.1%	13.0%	13.5%	14.7%	16.9%	11.5%	31.1%
STAR 2-pass ann	True	84855	32289	72.4%	19.4%	34.0%	45.0%	51.5%	56.0%	56.5%	59.2%	59.6%	63.1%	80.5%
	False	2254	30601	6.9%	1.8%	4.7%	7.0%	9.7%	8.6%	10.2%	10.7%	16.7%	12.3%	28.9%
TopHat1	True	74198	27192	73.2%	19.1%	36.7%	46.1%	53.2%	57.3%	61.4%	61.3%	62.0%	63.4%	79.3%
	False	1647	9737	14.5%	3.0%	5.9%	8.8%	11.6%	13.7%	16.0%	17.3%	15.9%	17.7%	32.1%
TopHat1 ann	True	81143	27776	74.5%	25.6%	45.2%	53.6%	59.0%	60.9%	63.0%	64.2%	64.7%	65.2%	81.0%
	False	1723	10552	14.0%	3.5%	6.5%	7.6%	11.2%	14.0%	12.7%	18.4%	15.8%	17.5%	30.2%
TopHat2	True	76693	27580	73.6%	21.2%	35.8%	48.9%	53.7%	57.1%	59.4%	60.7%	61.3%	62.3%	80.2%
	False	552	9016	5.8%	2.1%	2.9%	4.8%	4.8%	4.8%	5.2%	10.1%	6.6%	9.5%	9.5%
TopHat2 ann	True	84919	28645	74.8%	25.6%	43.0%	52.2%	57.4%	59.9%	59.5%	61.9%	61.0%	65.2%	81.9%
	False	1494	24895	5.7%	4.1%	4.3%	5.8%	6.5%	4.9%	5.6%	7.5%	6.9%	5.3%	9.6%
Truth	True	92247	31334	74.6%	21.1%	39.2%	52.1%	55.1%	59.4%	61.2%	62.0%	64.4%	64.0%	82.1%
	False	0	0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Number and percentage of exon junctions incorporated into transcript isoforms by Cufflinks. The junctions counted are those present in primary alignments, which were used as input to Cufflinks. Junctions are further classified as true and false by comparison to the simulated gene models.
n.a., not applicable.

Nature Methods: doi:10.1038/nmeth.2722

Supplementary Note

This part of the supplement describes the evaluated alignment protocols, the evaluation metrics, and additional results from analysis of read placement in relation to annotated genes.

Alignment protocols

Each of the sections 1–11 below describes an alignment program or pipeline. For parameter variations based on a common aligner, subheadings designate the individual protocols.

Each protocol made use of genome sequences for human assembly GRCh37 and mouse assembly MGSCv37, as provided at the UCSC Genome Browser website (http://genome.ucsc.edu) in FASTA format (hg19.fa and mm9.fa). The aligners require indices built from the genome FASTA files, as detailed below or in the documentation for the individual programs. These indices are specific to each aligner, but only need to be created once and can be reused for all alignment jobs to the same genome.

Some protocols also made use of gene annotation for the human and mouse genomes. The annotation was obtained in GTF format from Ensembl version 62 (http://www/ensembl.org) and adapted so that reference sequence coordinates corresponded to the genome sequence files from UCSC, using clone fragment and contig information to match the Ensembl and UCSC representations of the genome assemblies.

1. BAGET

An unreleased version of the BAGET pipeline was used. The earlier version 1.0 is available at http://icb.med.cornell.edu/wiki/index.php/BAGET along with a tutorial. BAGET has now been integrated into the r-make tool set (http://physiology.med.cornell.edu/faculty/mason/lab/r-make).

Briefly, BAGET first runs the short read aligner BWA[1] to align the input reads to the genome. Reads that were not aligned in this step are then searched against an index of known exon junctions, also using BWA. Any reads that remain unaligned are scanned for poly(A) tails. After trimming such tails, BAGET attempts to align the reads to the genome again, as above. BWA does not perform spliced alignment, and BAGET therefore relies on the index of known exon junctions to find spliced alignments.

2. GEM

The GEM suite comprises several alignment tools, including the GEM contiguous mapper[2] and the GEM splice mapper, that can be combined for RNA-seq analysis. The development snapshot 1.358 was used for this evaluation. Several versions of GEM are available from http://gemlibrary.sourceforge.net.

The workflow applied here implements a progressive alignment scheme where reads are mapped in stages. In the first stage, the GEM contiguous mapper is used to map the entire read. Reads for which a high-quality contiguous alignment are not found are passed to the GEM splice mapper. If a match is not found, a second iteration of contiguous/spliced alignment is attempted after trimming five nucleotides from the 5′ end of the read and 20 from the 3′ end. GEM was applied in three different pipeline configurations that differ in the set of junctions considered for spliced alignment, as outlined below.

2.1. GEM ann

In this protocol, GEM first carries out a *de novo* splice junction discovery step by aligning reads against the genome. This is followed by a second step, where spliced alignments are determined using the set of *de novo* junctions from the first step together with known junctions from the supplied annotation.

2.2. GEM cons

Alignment is carried out as above, but with a conservative subset of *de novo* junctions and without making use of annotation.

2.3. GEM cons ann

As above, but using the conservative subset of *de novo* junctions together with annotated junctions.

3. GSNAP

GSNAP version 2011-08-15 was used[3]. This version can be obtained from http://research-pub.gene.com/gmap.

3.1 GSNAP

In this basic protocol, GSNAP was used without annotation. The following option string was specified for each data set:

```
-B 5 -a paired -N 1 -m 4 -M 1 -i 2 -w 200000 -E 4 -n 100 --pairmax-rna=200000
--gmap-mode=pairsearch,terminal,improve -A sam -O
```

In addition, options `-d`, `--quality-protocol`, `-q` and `-t` and were set as appropriate for each alignment job to specify genome database, quality scale of input data and settings for parallel computing.

3.2 GSNAP ann

GSNAP was executed as above, with the additional option `-s` to supply an index of known splice sites.

4. GSTRUCT

GSTRUCT is a pipeline that makes use of GSNAP as its alignment component. Version 2011-08-15 was used here. The pipeline is not yet available, but a public release is expected soon. Briefly, GSTRUCT considers read alignments from GSNAP with a mapping quality score of 20 or greater, and creates three types of auxiliary information to be used for a re-alignment:

1. Splice sites: Splices found in the first iteration of GSNAP are filtered for consistency against the positive and negative gene extents in that region. These extents are the coverages over the paired-end lengths for paired-end reads that contain a predicted splice site.

2. SNPs: Variant genotypes are called from the first iteration of GSNAP and used with the SNP-tolerance feature of GSNAP in the second iteration.

3. Run lengths: The presence or absence of good alignments from the first iteration of GSNAP is recorded at each genomic position. When the second iteration of GSNAP cannot resolve a multi-mapping read, it prefers the one that overlaps a good alignment from the first iteration.

4.1 GSTRUCT

GSTRUCT was applied on the results from running GSNAP without annotation (see 3.1).

4.2 GSTRUCT ann

GSTRUCT was applied on the results from running GSNAP with splice site annotation (see 3.2).

5. MapSplice

An unreleased version of MapSplice[4] was used, internally called 8_8. This version was based on the most recent MapSplice 1 release 1.15.2, available from http://www.netlab.uky.edu/p/bioinfo.

5.1 MapSplice

This protocol corresponds to the standard method of running MapSplice and does not make use of gene annotation. MapSplice is designed to operate without annotation by default.

5.2 MapSplice ann

This protocol made use of gene annotation by running MapSplice with increased sensitivity (which would also cause it to detect more spurious junctions), followed by post-processing to filter out splice junctions with low read support that were not present in the annotation.

6. PALMapper

PALMapper has been described[5] and its source code, tutorials and further information are available from http://raetschlab.org/suppl/palmapper. The program was used in a variant-aware alignment pipeline, where the RNA-seq data is first aligned to the genome in order to detect possible variations in the genome sequence. These genome variants are used in a final alignment run and serve to improve read placement. Additionally, information on splice junctions collected during the initial run or from gene annotation can be used to improve the final alignment run.

PALMapper was run in two stages. The initial stage, for the detection of variants and junctions, allowed up to six edit operations and imposed restrictions on anchor length of split reads (`-min-spliced-segment-len`) and edit operations in the vicinity of splice sites (`-QMM`). Variant calls and junction information were recorded for later use. At this initial stage PALMapper was run with the following parameters:

```
palmapper -M 6 -G 5 -E 6 -l 25 -L 30 -K 12 -C 35 -I 200000 -NI 1 -SA 100 -CT 50 -a -S -report-
splice-sites 0.95 -filter-max-mismatches 0 -filter-max-gaps 0 -filter-splice-region 5 -
polytrim 40 -min-spliced-segment-len 10 -QMM 7 -acc <ACCSPLICEPATH> -don <DONSPLICEPATH> -
report-junctions <JUNCTIONSFILE> -qpalma-indel-penalty 5 -discover-variants -report-variants
<VARIANTSFILE> -no-gap-end 10 -non-consensus-search -report-splice-sites-top-perc 0.01
```

A sensitive alignment regime was applied for the final alignments, allowing for up to 10 edit operations and a maximum of two splice junctions per read. As variant and junction information collected in the first run were used for this alignment, read truncation was not enabled; instead a higher number of edit operations was allowed, leading to a possible accumulation of mismatches and indels at the ends of reads. At this subsequent stage PALMapper was run with the following parameters:

```
palmapper -M 10 -G 2 -E 10 -l 20 -L 20 -K 12 -C 30 -I 20000 -NI 2 -SA 5 -CT 50 -a -S -filter-
max-mismatches 0 -filter-max-gaps 0 -filter-splice-region 5 -junction-remapping
<JUNCTIONSFILE> -score-annotated-splice-sites <JUNCTIONSFILE> -acc <ACCSPLICEPATH> –don
<DONSPLICEPATH> -report-splice-sites-top-perc 0.005 -QMM 7 -use-variants <VARIANTSFILE> -max-
dp-deletions 1 -use-variants-editop-filter
```

Three strategies were used to post-process the alignments:

1. Alignment filtering by the Simple Alignment Filter Tool (SAFT; http://raetschlab.org/suppl/saft), which filters all alignments based on the number of edit operations, and spliced alignments based on the number of reads supporting splice junctions and minimal segment length. These criteria were set as detailed in the table below.

Protocol	Data set	Allowed edit operations	Junction-supporting reads required	Minimal segment length for spliced alignments
PALMapper cons	K562	0	3	18
PALMapper cons	Simulation 1	1	3	18
PALMapper cons	Simulation 2	4	5	18
PALMapper cons ann	Simulation 1	6	4	6
PALMapper cons ann	Simulation 2	6	6	6

2. Analysis and treatment of ambiguous read placement by the Multi-Mapper Resolution (MMR) Tool (http://raetschlab.org/suppl/MMR) to determine the best alignments for read pairs. This tool implements a strategy to select alignments by iteratively minimizing the variation of coverage in a window around the possible mapping locations. MMR options were set to "`-I 3 -F 1 -p -i 400000`" for K562 data and "`-I 2 -F 1 -p -i 400000`" for simulated data.

3. Alignment pair optimization to determine the best pairs of single-end alignments. This algorithm considers all proper pairs of alignments and iteratively selects pairs with maximal summed single-end alignment scores. The alignment score considers matches, mismatches, indels and base-call quality scores[6]. Multiple pairs were reported such that no single-end alignment was included in more than one pair.

Nature Methods: doi:10.1038/nmeth.2722

There were four protocols evaluated based on PALMapper. Software versions were: PALMapper 0.4rc3, SAFT 0.1 and MMR 0.1.

6.1 PALMapper

Variant-aware alignment without annotation was followed by MMR.

6.2 PALMapper ann

Variant-aware alignment with annotation, followed by MMR.

6.3 PALMapper cons

This more conservative protocol comprises variant-aware alignment without annotation, followed by SAFT filtering and alignment pair optimization.

6.4 PALMapper cons ann

Variant-aware alignment with annotation followed by SAFT filtering and alignment pair optimization.

7. PASS

The PASS spliced alignment pipeline[7] version 1.64 was run in two different ways. Annotation was not used. PASS can be downloaded from http://pass.cribi.unipd.it.

7.1 PASS

Default parameters for Illumina data were used. With these settings, truncation of low-quality bases is enabled and the maximum number of allowed mismatches per mapping is fixed. Read truncation is based on a learning step that correlates the number of mapped reads with the base call quality scores of excluded bases.

7.2 PASS cons

Default parameters for Illumina data were used as above, except for the variable `SAM_REDUNDANCY_PAR`, which was set to add the options: `-unpaired_coverage 1 -unpaired_score 60`
These options serve to increase specificity by filtering out alignments at genomic regions of low coverage.

8. ReadsMap

The ReadsMap program is part of the Transomics pipeline from Softberry (http://www.softberry.com). ReadsMap production release 1.0 (internal version number 6.0.0) was applied with default parameters, without providing gene annotation or mate pair information. The default parameters are suitable for mapping reads with mismatches, but mapping reads with indels requires other options. Poor-quality tails were not truncated from reads and partial mappings were not reported.

Regions marked as repeats in the reference genome sequence were ignored, except for the first and last 30 bp of such regions. For this purpose, the masking information in the genome sequence from UCSC was used, where repeats correspond to elements identified by RepeatMasker or Tandem Repeats Finder (with a period of 12 or less). Most reads originating from such repeats were therefore not mapped.

9. SMALT

SMALT version 0.5.1 was used, and is available at http://www.sanger.ac.uk/resources/software/smalt.

The indices of the reference genomes were built with the following options:

```
smalt index -k 13 -s 7 hg19k13s7 hg19.fa
smalt index -k 13 -s 7 mm9k13s7 mm9.fa
```

All human reads were aligned with the following options:

```
smalt map -x -p -f samsoft -o mapped.sam hg19k13s7 mate1.fq mate2.fq
```

All mouse reads were aligned with the following options:

```
smalt map -x -p -f samsoft -o mapped.sam mm9k13s7 mate1.fq mate2.fq
```

Although SMALT does not perform spliced alignments, it can report up two complementary alignments per read. This feature is activated with the –p option, which was used here. When two complementary alignments are reported, one will be labeled as secondary (see the SMALT manual). The SAM format output from SMALT was post-processed to merge compatible primary and secondary alignments of the same read into spliced alignments. Briefly, gaps between primary and secondary alignments were filled with intron (N) operations, and priority given to the primary alignment when the same part of the read was included in both alignments.

10. STAR

STAR version 1.9 was used[8]. Although this version has not been released, the more recent version 2.1.1 available from http://code.google.com/p/rna-star/ only differs with regard to input/output formatting and minor bug fixes.

10.1 STAR 1-pass

In this most basic protocol, STAR was used in single-pass mode and without annotation. STAR uses genome index files that must be saved in unique directories. The human genome index was built from the FASTA file hg19.fa as follows:

```
genomeDir=/path/to/hg19
mkdir $genomeDir
STAR --runMode genomeGenerate --genomeDir $genomeDir --genomeFastaFiles hg19.fa \
 --runThreadN <n>
```

The option `--runThreadN` should be set to specify the number of processor threads to use. The mouse genome index was built from mm9.fa using the same options. Alignment jobs were excuted as follows:

```
runDir=/path/to/1pass
mkdir $runDir
cd $runDir
STAR --genomeDir $genomeDir --readFilesIn mate1.fq mate2.fq --runThreadN <n>
```

10.2 STAR 1-pass ann

In this protocol, STAR uses a splice junction database to improve accuracy. Splice junction coordinates are supplied at the index generation step in a tab-delimited file, as detailed in the STAR manual. The genome index was created as described under 10.1 above, with two additional options:

```
--sjdbFileChrStartEnd /path/to/junctions.txt --sjdbOverhang 75
```

Alignment jobs were then executed as follows:

```
runDir=/path/to/1pass_ann
mkdir $runDir
cd $runDir
STAR --genomeDir $genomeDir --readFilesIn mate1.fq mate2.fq --runThreadN <n>
```

10.3 STAR 2-pass

In the STAR 2-pass approach, splice junctions found in a first alignment run are used to guide the final alignment. The first pass is performed as described under 10.1 above. A new index is then created using splice junction information contained in the file SJ.out.tab from the first pass:

```
genomeDir=/path/to/hg19_2pass
mkdir $genomeDir
STAR --runMode genomeGenerate --genomeDir $genomeDir --genomeFastaFiles hg19.fa \
 --sjdbFileChrStartEnd /path/to/1pass/SJ.out.tab --sjdbOverhang 75 --runThreadN <n>
```

The resulting index is then used to produce the final alignments as follows:

```
runDir=/path/to/2pass
mkdir $runDir
cd $runDir
STAR --genomeDir $genomeDir --readFilesIn mate1.fq mate2.fq --runThreadN <n>
```

10.4 STAR 2-pass ann

In this version of the 2-pass protocol, annotated splice junctions are provided in the first alignment step. The first pass is therefore executed as described under 10.2 above. New index files are then created using splice junction information contained in the file SJ.out.tab from the first pass:

```
genomeDir=/path/to/hg19_2pass_ann
mkdir $genomeDir
STAR --runMode genomeGenerate --genomeDir $genomeDir --genomeFastaFiles hg19.fa \
 --sjdbFileChrStartEnd /path/to/1pass_ann/SJ.out.tab --sjdbOverhang 75 --runThreadN <n>
```

The resulting index is then used to produce the final alignments as follows:

```
runDir=/path/to/2pass_ann
mkdir $runDir
cd $runDir
STAR --genomeDir $genomeDir --readFilesIn mate1.fq mate2.fq --runThreadN <n>
```

11. TopHat

The spliced alignment program TopHat[9,10] uses the short read aligner Bowtie[11,12] as its alignment engine. Two versions of TopHat and Bowtie were evaluated, both available from http://tophat.cbcb.umd.edu.

11.1 TopHat1

This protocol followed the recommendations in a recent publication by the TopHat developers[13], using TopHat version 1.3.2 with default options, except for options specifying quality scale of input data, library type and number of processor threads. The options were as follows.

For mouse data:

```
-o tophat.out -p 8 mm9 mate_1.fq mate_2.fq
```

For K562 data:

```
-o tophat.out -p 8 --solexa1.3-quals --library-type=fr-firststrand hg19 mate_1.fq mate_2.fq
```

For simulated data:

```
-o tophat.out -p 8 --solexa1.3-quals hg19 mate_1.fq mate_2.fq
```

Bowtie version 0.12.7.0 was used for read alignment.

11.2 TopHat1 ann

TopHat was used as specified under 11.1 above, with the added option `-G` to supply a gene annotation file in GTF format.

11.3 TopHat2

The most recent TopHat and Bowtie versions available at the time of this study were used (2.0.3 and 2.0.0.6, respectively) with the options specified under 11.1 above.

11.4 TopHat2 ann

TopHat 2.0.3 and Bowite 2.0.06 were used with the options specified under 11.2 above.

Evaluation metrics

This section discusses the metrics used to evaluate aligners in this study. Some of these metrics, or highly similar ones, have also been employed in earlier comparisons of spliced aligners[4,8,9,14,15], as noted in several instances below.

General definitions

Unless otherwise mentioned, metrics were computed on the set of primary alignments in the output from each protocol, so as not to bias the evaluation due to differences among protocols in the number of alignments reported per read.

In assessing accuracy on simulated data, we have applied the concepts of precision and recall to a range of features, including insertions, deletions, splices and transcript isoforms, as detailed below. In general terms, precision is defined as the proportion of predicted features that are correct, and recall as the proportion of actual features that are correctly predicted. Note that precision is also known as positive predictive value (PPV) and equivalent to 1 – false discovery rate (FDR). Sensitivity is an alternative term for recall. For an extensive discussion of precision and recall in the context of short read alignment, see Lindner and Friedel[16].

In assessing spliced alignment performance, we distinguish between detection of *splices* in individual reads and detection of unique *splice junctions* on the genomic sequence. The latter are often supported by multiple splices depending on expression level and sequencing depth.

Alignment yield

We measured the proportion of sequenced (or simulated) reads that were mapped and the frequency of ambiguous mappings (i.e. reads with more than one reported alignment). While a high frequency of mapped reads is desirable, this must be balanced against the risk of reporting erroneous alignments. It should also be noted that high-throughput sequencing data often contains a proportion of reads that originate from adapter or primer sequences used during library construction, and reads with error rates that preclude mapping. A good aligner would therefore be expected to report alignments for most but not all reads, when applied to high-quality output from current sequencing instruments.

Yield metrics were summarized both at the level of individual reads and read pairs (**Figs. 1** and **2a** and **Supplementary Table 3**). Alignment programs are expected to report consistently mapped pairs: if one read can be uniquely mapped, it should generally be possible to place its corresponding paired read uniquely as well (**Fig. 1**, dark blue bars).

When a read pair matches well to multiple genomic locations and a single placement cannot be selected with high confidence, an aligner may output multiple alignments for the read. In those cases, the rules of this evaluation still require that a single alignment per read be labeled as most likely (primary). This is also the practice recommended in the SAM alignment file format specification[17].

It should be noted that several aligners apply strategies to place multi-mapping reads uniquely by using information from other reads (**Supplementary Note 1**), so that even if a read matches multiple locations equally well at the sequence level, it may still be possible to prioritize the correct location. An advanced alignment program would therefore be expected to produce unique mappings for most reads.

Some of the tools evaluated here reported a very high frequency of ambiguous mappings (**Fig. 1** and **Supplementary Fig. 1**). Such levels of uncertainty in the alignment output can result in suboptimal results in downstream analyses (**Supplementary Fig. 19**), where tools have difficulty choosing among the many alternative read placements. Reporting of many alignments per read can also result in very large output files, which are difficult to store and process.

Mismatch and truncation frequencies

An aligner should be able map reads with multiple mismatches, which may represent true differences between the sequenced transcriptome and the reference genome, or constitute errors introduced during

sample preparation and sequencing. We computed the number of mismatches (substitutions) per primary read alignment and visualized the resulting distributions (**Fig. 2a** and **Supplementary Fig. 4**). Some of the alignment protocols evaluated here showed a low tolerance for mismatches. In this context, it should be noted that many programs have an option to increase the tolerance for mismatches at the expense of longer running time. However, the programs assessed here were executed with settings chosen by the developers, and the evaluated protocols should therefore correspond to best-practice workflows. All programs were run by the respective developer teams, except for TopHat, which was executed by the evaluation team according to the protocol published by the authors[13].

The distribution of mismatches in alignments would be expected to follow to the base caller quality score distribution, such that a read with low mean quality score contains more mismatches relative to the genomic sequence. We observed that protocols with a low tolerance for mismatches also failed to align a large proportion of reads with low mean base call quality score (**Supplementary Fig. 2**).

A very high frequency of mismatches in the output may also be an indication of poor performance. One would typically expect few mismatches if the data is of high-quality. If a particular alignment program outputs a significantly lower of number of mismatch-free mappings than others, this may indicate that suboptimal alignments are being reported.

Truncation frequency

The frequency of mismatches in alignments should be interpreted in the context of truncation behavior (**Fig. 2**). Some aligners can truncate the ends of reads, and thus output a partial alignment when unable to map an entire sequence. This is a particularly important feature for spliced alignment programs, as a proportion of reads in any RNA-seq data set will contain splices near the read termini, such that one exon is covered only by a few bases. It is often impossible to align such read ends confidently. A good spliced aligner would therefore be expected to output a moderate proportion of truncated alignments.

Basewise accuracy

The use of simulated data facilitates exact computation of accuracy metrics, of which basewise accuracy is the most fundamental. Here, we measured the proportion of all simulated bases that were correctly mapped, and the proportion incorrectly mapped (**Supplementary Tables 2** and **5**). Related metrics were used in the study by Grant et al.[14]. We additionally computed accuracy separately for unspliced reads and those containing splice junctions (**Supplementary Tables 6–7**). The performance on the latter group is of particular interest to this evaluation, and these reads tend to be more difficult to align. Note that when computing basewise accuracy, ambiguity in indel placement must be accounted for, as discussed in earlier work[14] and described in Methods.

Read placement accuracy

In addition to basewise accuracy, it is important to measure performance at the read level. Read frequencies may be more relevant than base frequencies for several downstream applications. For example, to quantify gene expression levels it may be sufficient to assign reads to correct loci, even if some bases are incorrectly placed or alignments are truncated.

Here, we computed the proportion of simulated reads that were perfectly mapped, the proportion with a subset of bases correctly placed, and the proportion of reads that were mapped with no base correctly placed. The last category will typically consist of reads that were assigned to the wrong locus, but we noted that one program placed a substantial proportion of reads at approximately the correct location due to a programmatic error. Hence, we separately tallied reads for which the alignment overlapped the correct location, but had no base correctly placed (**Fig. 3** and **Supplementary Tables 5–7**).

Accuracy among unique and ambiguous mappings

By comparing accuracy between unique and ambiguous mappings, a level of confidence can be established for each category (**Supplementary Table 4**). For example, if the accuracy is very low among ambiguous mappings, it may be advantageous to exclude those from downstream analyses. A good aligner should map the great majority of reads uniquely, and achieve high accuracy for the set of uniquely mapped reads.

Indel frequency and accuracy

It is difficult to implement sensitive detection of insertions and deletions (indels) within the context of the fast search algorithms used by short read aligners[3,12], and the capability to detect indels therefore differs markedly among mappers. Here, we captured these trends by counting the number of insertions and deletions in the primary alignments from each protocol. The results were expressed as indel frequencies, defined as the number of indels per thousand sequenced reads. Indel frequencies are tabulated in **Figure 4a** and **Supplementary Figure 5**, which also use bar charts to depict the size distribution of indels from each program. These distributions reveal that some protocols lack the ability to detect longer indels.

We additionally computed the accuracy of indel detection on simulated data. Precision and recall (defined above) were computed for indels of different length, thus extending the approach of Grant et al.[14]. The resulting matrices were visualized using heatmaps (**Fig. 4b**). These figures illustrate the differences in accuracy among protocols, and how this is affected by indel size.

Spatial distribution of mismatches, indels and splices over read sequences

Depending on the search algorithms used by aligners, biases may result in the distribution of alignment features (mismatches, indels and splices) over the read sequences. We plotted these distributions, averaged over all primary alignments, for each protocol (**Supplementary Fig. 7**). The frequency of mismatches would typically be expected to increase towards the ends of reads, reflecting a concomitant decrease in sequence quality (**Supplementary Fig. 8**). This trend was not apparent for all protocols, indicating a problem with the placement of substitutions.

In contrast, gaps (indels and splices) should primarily reflect differences between the genome and transcriptome, as opposed to sequencing artifacts (for current Illumina sequencing data). The distribution of these features should therefore be roughly even over the read length. A reduction in gap frequency towards the ends of reads may reasonably be expected, as confident gap placement, particularly intron placement, can be difficult or even impossible near read termini (see the section on *Truncation frequency* above).

Coverage of annotated genes

We explored a range of metrics reflecting how reads were placed in relation to annotated genes: number of exon hits (alignments covering only exonic features), spliced exon hits (as the previous category, but aligning with a splice operation), partial exon hits (alignments covering exonic and non-exonic features), intron hits, intergenic hits, number of genes with proper exon hits, proportion of exon hits and the number of alignments associated with specific types of features (protein-coding, pseudogene, etc.). Scatter plots were used to uncover trends in the coverage statistics. A selection of these are shown in **Supplementary Figures 9–11**. In order to aid the interpretation of the data in various plots, a trend line was plotted alongside the data points based on linear regression.

This analysis served in part to confirm that aligners behave similarly on simulated data compared to real data when high-level metrics are considered (representative behavior on simulated data was also confirmed using the more fundamental metrics described above). Additionally, we searched for cases where particular protocols constituted outliers, indicating exceptional or aberrant performance. We reasoned that trends in different coverage statistics, if consistent across many datasets, can give indirect indications about the relative performance of the methods.

For example, if a method reports more spliced alignments than others, and the remainder of the statistics show no anomalies, this is indicative of better relative performance. Of course, this interpretation is inherently subjective, as it is only valid if the reported spliced mappings are actually correct, something which cannot be established in the case of real datasets. In spite of this caveat, exploration of feature coverage statistics can provide enough insight to nominate the best performing methods. While unlikely to provide a clear-cut ranking of the methods, such conclusions are established independently from the simulation benchmarking results, and hence can reinforce them.

Nature Methods: doi:10.1038/nmeth.2722

Splice frequency and junction characteristics

A metric of particular interest for the evaluation of spliced aligners is the frequency of splices present in alignments. Splice frequency was defined as the number of reported splices divided by the number of sequenced reads. As an indication of whether reported splices are likely to be correct, we separated splices matching annotated introns from novel splices (**Fig. 5a**). A further dimension was added to this analysis by counting the number of alignments supporting each reported junction (**Fig. 5b**). For a well annotated genome, high rates of novel junctions supported by few read alignments indicates a significant false discovery rate. This type of analysis was also employed in the publication describing the aligner STAR[8].

To further characterize the novel junctions, we distinguished four categories depending on whether the splice sites where annotated and belonged to the same gene (**Supplementary Figs. 14–15**). This revealed that different aligners tend to predict different types of novel junctions. We additionally studied the size distribution of splices on both real and simulated data (**Supplementary Fig. 13**). Unexpected shapes of those curves, such as sudden bumps or stair-like appearance, are indicative of problems with spliced alignment. The erratic nature of such trends can be confirmed by comparisons between results on real and simulated data, and by considering the true distributions produced by the simulator.

Splice accuracy

For results on simulated data, precision and recall of splices was computed. A splice was considered correct if placed so that its genomic start and end (donor and acceptor) coordinates agreed with those of the true alignment. This analysis was carried out for all splices in primary alignments (**Fig. 5a**), as well as for the subset located between positions 20 and 57 in the 76 nt reads (**Supplementary Fig. 16**). This subset can be aligned with higher confidence due to the existence of at least 20 nt flanking sequence on each side of the splice. It is therefore of interest to see whether the relative performance of aligners differs for this group of more tractable splices. Note that these figures show FDR (1-precision) rather than precision, for consistency with the curves in **Figure 5c–e** (described below). Splice recall was further stratified based on true read coverage of corresponding junctions (**Supplementary Fig. 17**). Several aligners use information from multiple reads in same locus to place splices in individual read alignments. This can lead to a bias, such that splices are preferentially detected at high-coverage junctions. This has been investigated in a similar manner in earlier comparisons of spliced aligners[4,9,15].

Junction frequency and accuracy

Precision and recall was also computed for junction calls on the simulated data (**Supplementary Table 2**). A junction was considered correct if its genomic start and end coordinates matched those of a junction in the simulated transcriptome. The distinction between splice and junction metrics is important: a method may align the great majority of spliced reads correctly (high splice accuracy), and still distribute a small proportion of reads over many false junctions (low junction accuracy).

We noted that indeed many such false low-coverage junctions were reported. To demonstrate this behavior, we counted the number of junctions at different thresholds for the number of alignments required to call a junction. The results were visualized by plotting counts of true versus false junction calls at each threshold, yielding figures that can be interpreted in a similar manner to receiver operator characteristic (ROC) plots (**Fig. 5c–e**). Similar approaches have been used in previous aligner comparisons[8,12]. Here, methods with high junction accuracy can be identified by curves that are above and to the left of those of other methods.

Transcript reconstruction accuracy

A common aim of RNA-seq studies is to identify the complete transcript isoforms present in the assayed samples. Due to the fragmentary nature of RNA-seq library construction and data acquisition, isoform reconstruction is a difficult problem. Several algorithms designed for this task have been implemented[18-20], of which Cufflinks is the most widely established. To assess the suitability of alignment results for transcript reconstruction, we ran Cufflinks on the output from each alignment protocol, and computed precision and recall for reconstruction of individual exons as well as spliced transcripts (see Methods for details). As transcript reconstruction may be impossible for isoforms with low read coverage, recall was stratified by expression level for simulated data.

Nature Methods: doi:10.1038/nmeth.2722

Coverage of annotated genes

We assessed how RNA-seq reads were placed in relation to annotated gene structures from the Ensembl database. The results are briefly summarized in Results and some further observations are detailed here.

Relative to the frequency of exonic alignments, BAGET and SMALT mapped a high proportion of reads to intronic sequence, whereas the opposite trend was apparent for ReadsMap and to some extent the TopHat2 protocol using annotation (**Supplementary Figs. 9–11**). For BAGET and SMALT, the likely explanation is that priority is given to reads aligned in an unspliced manner to the genome. The annotation-based TopHat2 protocol takes the opposite approach – first aligning reads to the known transcriptome – and may thereby underrepresent intronic mappings. ReadsMap avoids repeat elements (**Supplementary Fig. 12**), which are prevalent in introns and represent challenging mapping targets due to the many homologous sequences present throughout the genome.

The occurrence of read alignments partially overlapping exons was also exceptionally high in the output from BAGET and SMALT. It is likely that such mappings result from failure to identify splice junctions, as suggested by a negative correlation with counts for spliced alignments at exons (**Supplementary Figs. 9–11**). TopHat2, GSNAP, GSTRUCT, STAR, MapSplice and the most conservative PALMapper protocol typically reported the fewest alignments partially overlapping exons, close to the expected result for simulated data.

For GSNAP, the performance on most gene coverage metrics was dependent upon the provision of gene annotation, while the related, more advanced GSTRUCT pipeline performed similarly with and without annotation. The same trend was apparent for STAR, where the basic (1-pass) version benefited greatly from using annotation, and the more advanced (2-pass) version behaved similarly to GSTRUCT.

References

1. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
2. Marco Sola, S., Sammeth, M., Guigó, R. & Ribeca, P. The GEM mapper: fast, accurate and versatile alignment by filtration. *Nat. Methods* **9,** 1185–1188 (2012).
3. Wu, T. D. & Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26,** 873–881 (2010).
4. Wang, K. *et al.* MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.* **38,** e178 (2010).
5. Jean, G., Kahles, A., Sreedharan, V. T., De Bona, F. & Rätsch, G. RNA-Seq read alignments with PALMapper. *Curr Protoc Bioinformatics* **32,** 11.6.1–11.6.37 (2010).
6. De Bona, F., Ossowski, S., Schneeberger, K. & Rätsch, G. Optimal spliced alignments of short sequence reads. *Bioinformatics* **24,** i174–80 (2008).
7. Campagna, D. *et al.* PASS: a program to align short sequences. *Bioinformatics* **25,** 967–968 (2009).
8. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).
9. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–1111 (2009).
10. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).
11. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10,** R25 (2009).
12. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).
13. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* **7,** 562–578 (2012).
14. Grant, G. R. *et al.* Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics* **27,** 2518–2528 (2011).
15. Bonfert, T., Csaba, G., Zimmer, R. & Friedel, C. C. A context-based approach to identify the most likely mapping for RNA-seq experiments. *BMC Bioinformatics* **13 Suppl 6,** S9 (2012).
16. Lindner, R. & Friedel, C. C. A comprehensive evaluation of alignment algorithms in the context of RNA-seq. *PLoS ONE* **7,** e52403 (2012).

17. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–2079 (2009).
18. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28,** 511–515 (2010).
19. Li, J. J., Jiang, C.-R., Brown, J. B., Huang, H. & Bickel, P. J. Sparse linear modeling of next-generation mRNA sequencing (RNA-Seq) data for isoform discovery and abundance estimation. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 19867–19872 (2011).
20. Mezlini, A. M. *et al.* iReckon: Simultaneous isoform discovery and abundance estimation from RNA-seq data. *Genome Res.* **23,** 519–529 (2013).