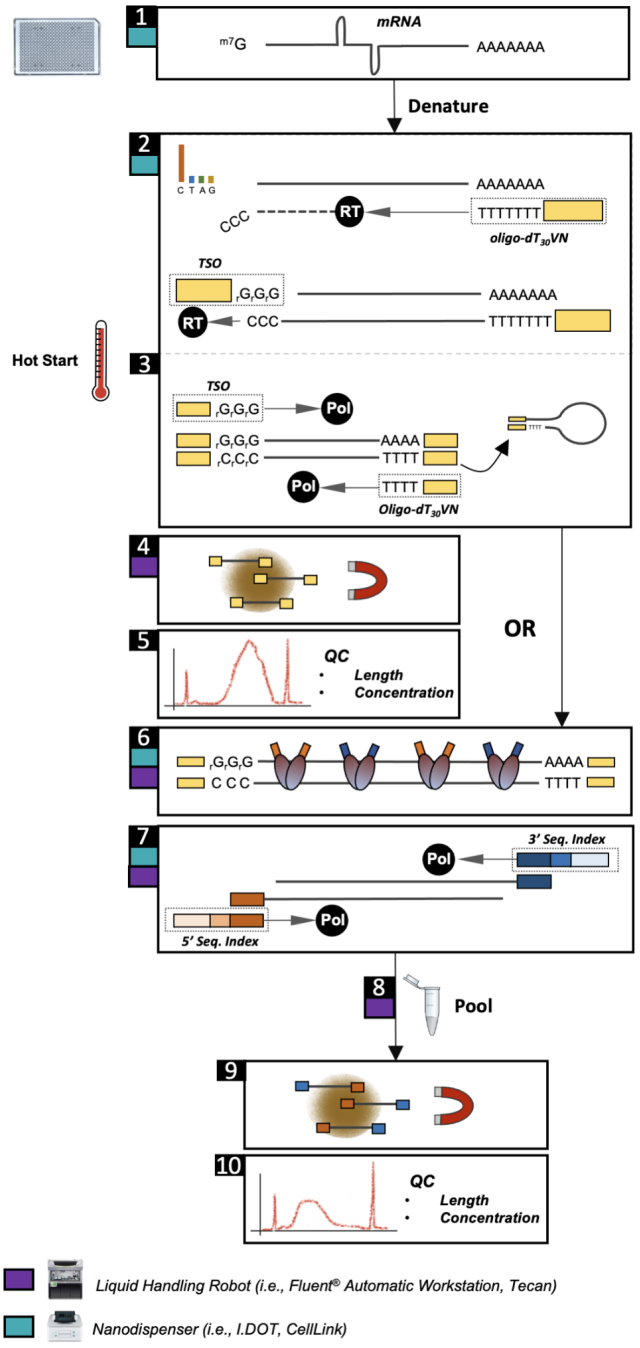


Fig. S1| 10 Steps of FLASH-Seq Workflow. Steps automated with a liquid handling robot or nanodispenser are marked by a purple and teal square, respectively. **Step 1**, cells are individually sorted in plates in the lysis buffer. **Step 2**, after denaturation, mRNAs are reverse transcribed using a template-switching reverse transcriptase (RT) in the presence of an excess of dCTP. **Step 3**, cDNA is amplified by semi-suppressive PCR. **Step 4,** cDNA is purified using magnetic beads. **Step 5**, cDNA concentration and fragment size are measured and the samples are diluted to a final concentration of ~100-200 pg/µL. **Step 6**, the cDNA is tagmented using a hyperactive Tn5 transposase which introduces known adaptor sequences. **Step 7**, the tagmented cDNA is amplified by PCR and sequencing indices are added. **Step 8**, all samples are pooled together. **Step 9,** the library is purified using magnetic beads. **Step 10**, concentration and average fragment size is measured in preparation for the sequencing.

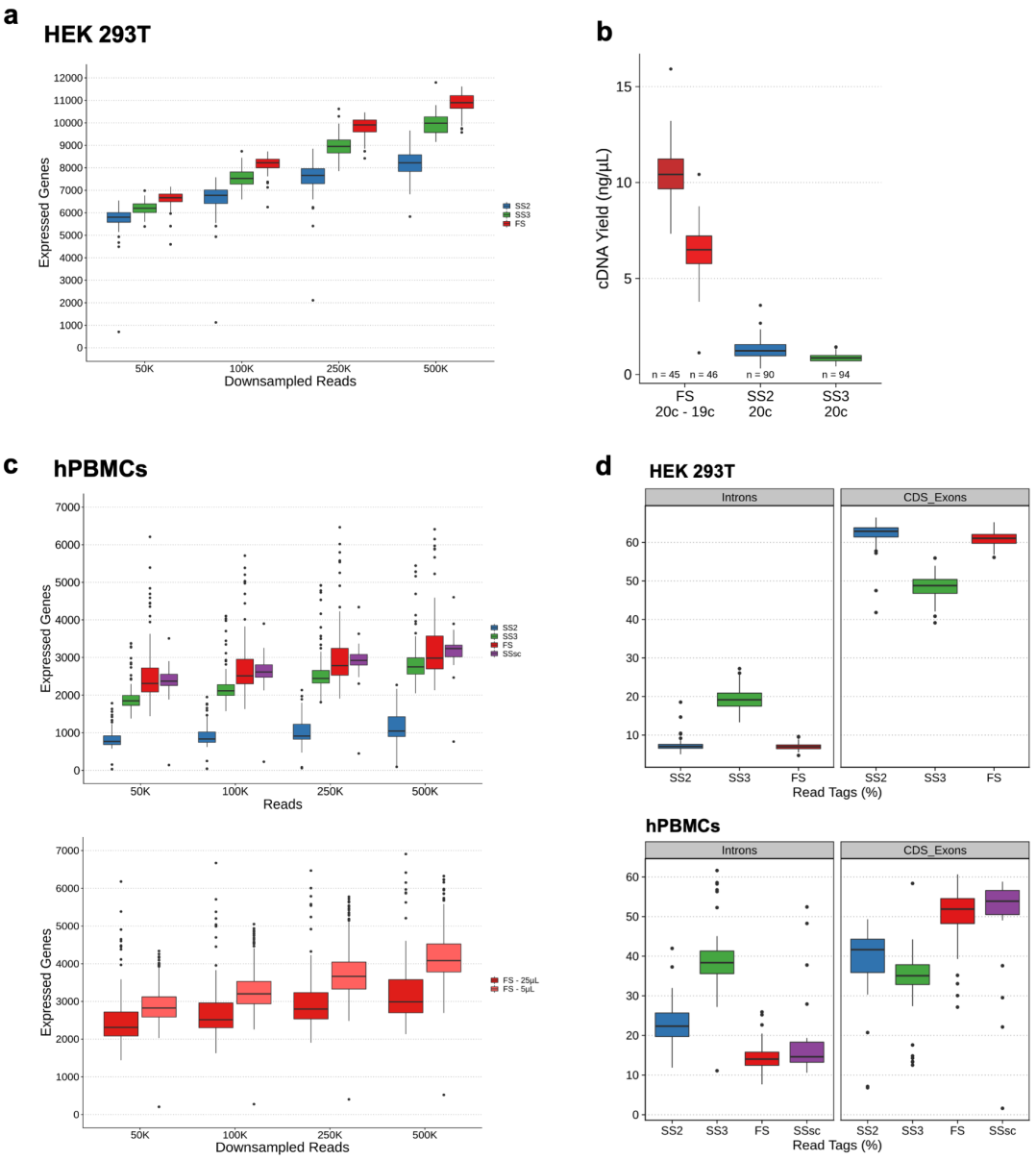


Fig. S2 | FLASH-Seq gene detection using downsampled reads and cDNA yield. a. Number of detected genes in HEK 293T cells processed with SS2 (n = 80), SS3 (n = 42) or FS (n = 85) using 50,000, 100,000, 250,000 or 500,000 downsampled reads. Gene detection threshold was set to >0 read. **b.** cDNA Yield, in ng/µL for SS2 (20 PCR cycles), SS3 (20 PCR cycles) or FS (20 or 19 PCR cycles) **c.** Upper part, number of detected genes in hPBMCs processed with SS2 (n = 47), SS3 (n = 76), FS (n = 136) or SMART-Seq Single Cell Kit (SSsc, Takara Bio, n = 24). FS, SS2, SS3 and SSsc reaction volumes were respectively: 25, 25, 10 and 50 µL. Bottom part, number of detected genes in hPBMCs processed with FS (n = 136, 25 µL) and miniaturized FS (n = 231, 5 µL). Raw reads were downsampled to 50,000, 100,000, 250,000 or 500,000. Gene detection threshold was set to >0 read. **d.** Estimated proportion of read mapped to intronic or CDS-exonic features, in HEK 293T cells and hPBMCs, in read tag percentages.

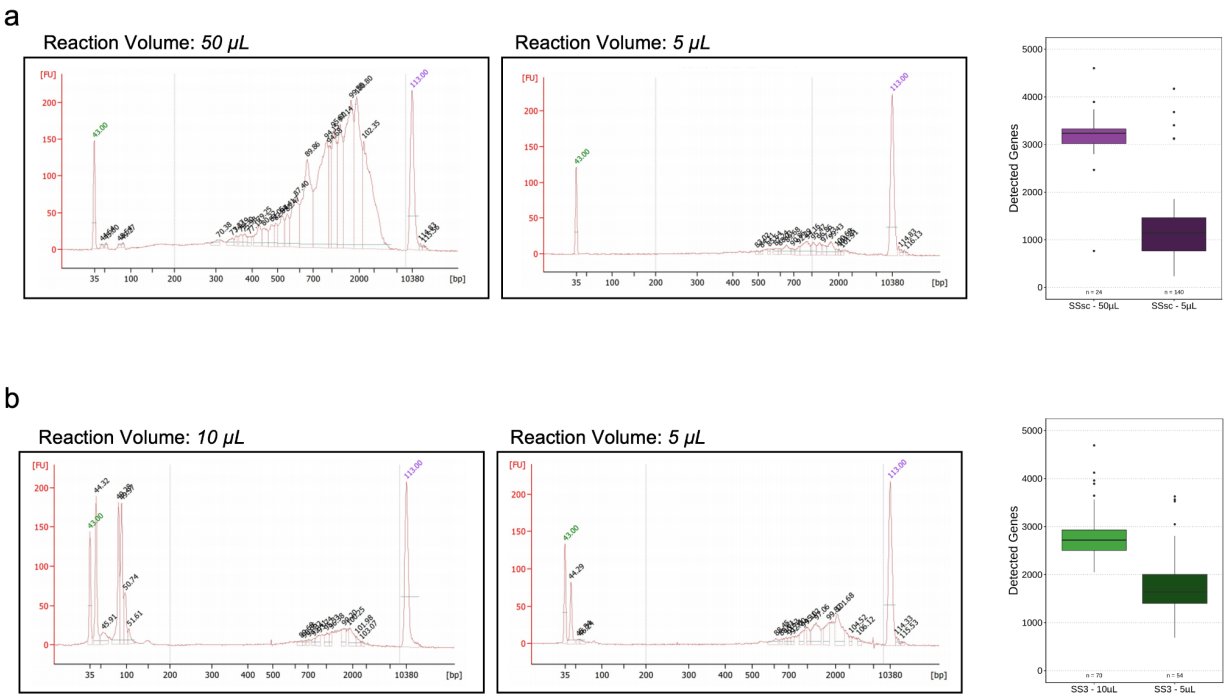


Fig. S3 | **SSsc and SS3 miniaturization to 5 µL. a.** Bioanalyzer traces illustrating the cDNA length distribution of regular (50 µL) or miniaturized (5 µL) SSsc. The right-panel shows the number of detected genes (>0 read) in both conditions. **b.** Bioanalyzer traces showing the cDNA length distribution in regular (10 µL) or miniaturized (5 µL) SS3. The right-panel shows the number of detected genes (>0 read) in both conditions using 500,000 downsampled raw reads.

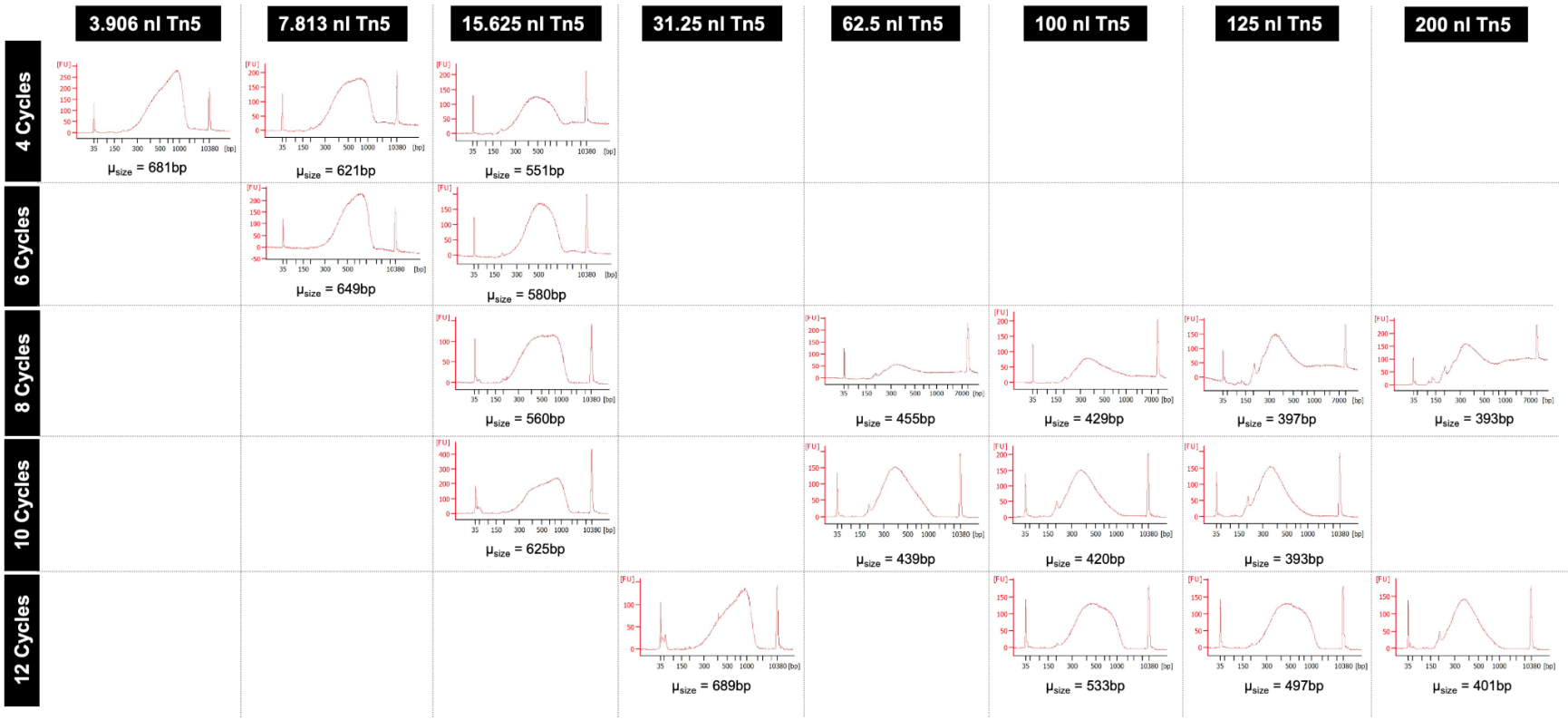


Fig. S4 | Impact of the number of PCR cycles and Tn5 amount on FS-LA cDNA size distribution. Bioanalyzer traces of the final FS-LA libraries. The number of PCR cycles used for cDNA pre-amplification is depicted on the rows (4 to 12 cycles). The amount of Tn5 added to each cell for tagmentation is depicted on the top of the chart. The mean cDNA size distribution between 200 bp and 9000 bp is displayed below each trace.

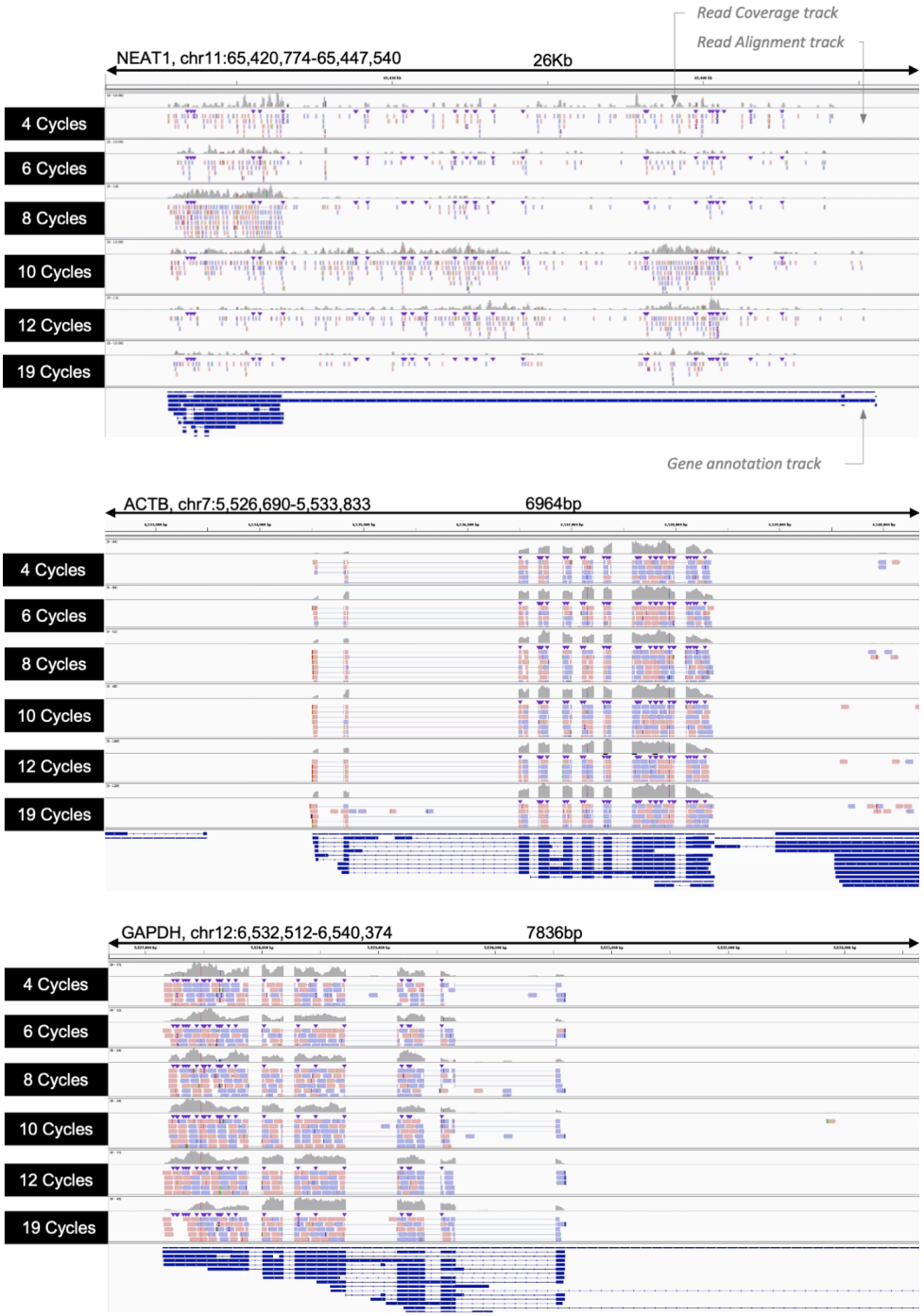


Fig. S5 | Integrated Genome Viewer visualization of selected genes. Each panel contains six representative samples which were downsampled to 500,000 raw reads and mapped onto hg38. As depicted in the first panel, each sample consists of a gene coverage, a read mapping and gene annotation track. In the read mapping track, each blue or red bar corresponds to a single mapped read. The color of the bar depicts the read orientation compared to the reference. Fine lines highlight split reads. In the gene annotation track, fine blue lines correspond to introns while bold blue lines represent exons. Each panel shows a different gene (from top to bottom: lncRNA-NEAT1, ACTB, GAPDH).

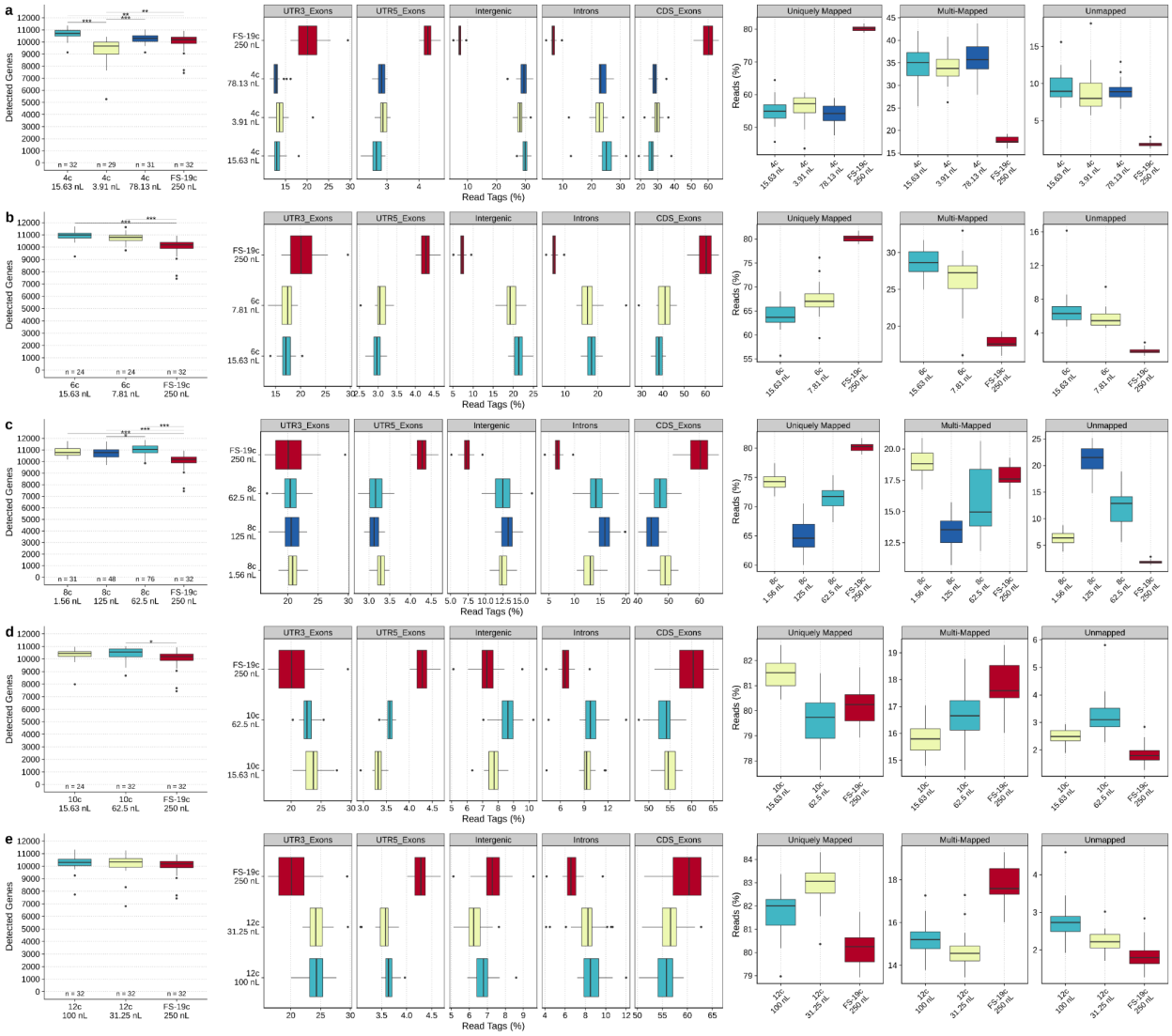


Fig. S6 | Overview of FS-LA gene detection and mapping statistics. Each panel from right to left: First, the number of detected genes in HEK 293T cells using FS (19 PCR cycles) or FS-LA, 250,000 downsampled reads, gene detection threshold >0 read. Significance level was evaluated using Dunn's test. P-values were corrected for multiple testing using Bonferroni correction (pval < 0.05). Second, the proportion of read mapped to exonic, intronic or intergenic features measured using ReSQC, in read tag percentages. Third, the percentage of uniquely mapped, multi-mapped or unmapped reads. Each panel compares regular FS (19 PCR cycles, in red) with **a.** 4 PCR cycles FS-LA **b.** 6 PCR cycles FS-LA **c.** 8 PCR cycles FS-LA **d.** 10 PCR cycles FS-LA **e.** 12 PCR cycles FS-LA.

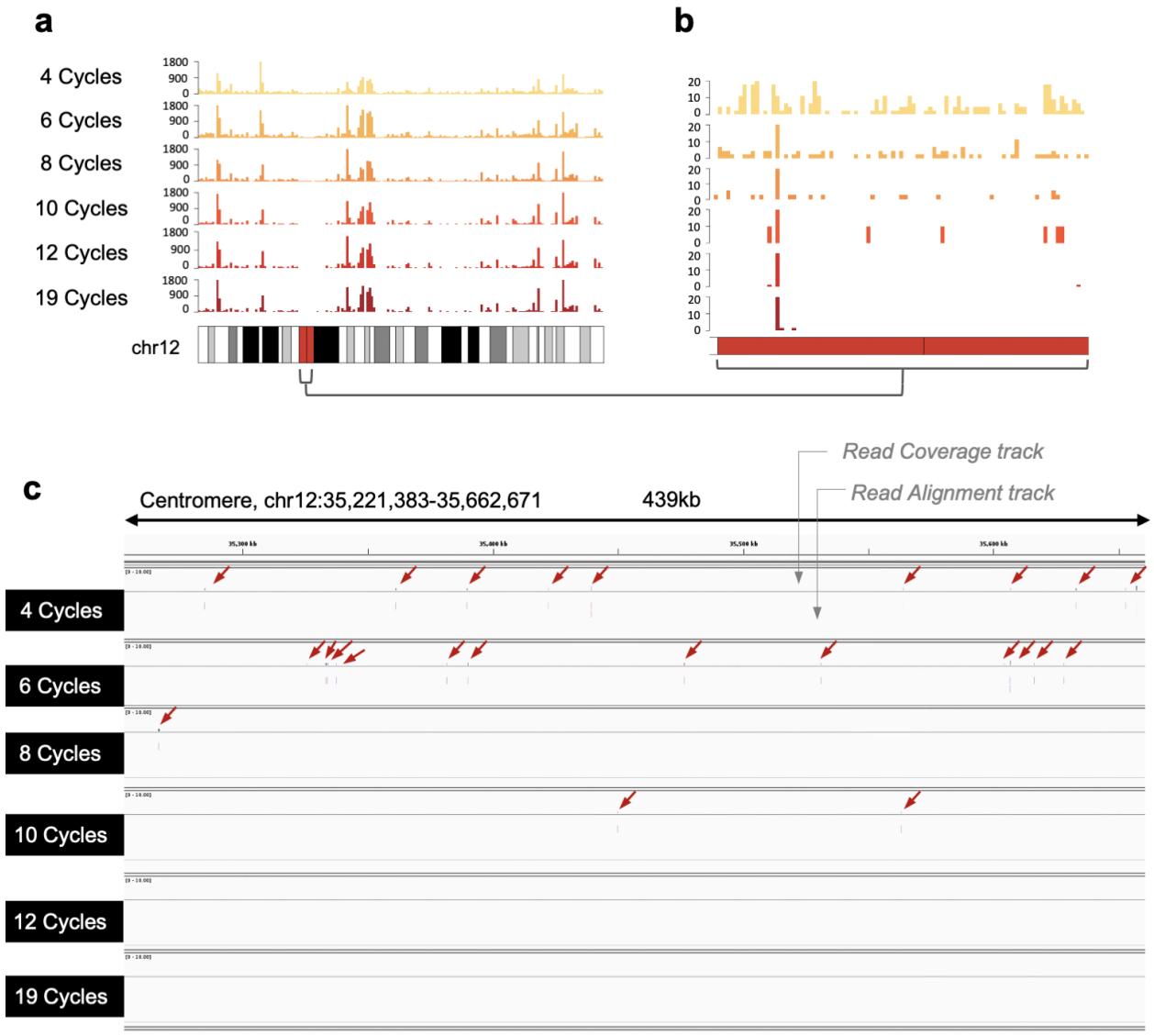


Fig. S7 | Intergenic reads generated during low-amplification. a. Density of mapped reads on chromosome 12, averaged to 750,000 bp windows for 6 representative samples, amplified with 4, 6, 8, 10 or 12 PCR cycles (FS-LA) or 19 PCR cycles (FS) and downsampled to 500,000 raw reads. **b.** Zoom on the centromeric region. **c.** Integrated genome viewer screenshot showing an example of unexpected read mapping on a 439Kb region of chromosome 12 centromere. Mapped reads are marked by a red arrow. Each line represents a different sample processed with 4, 6, 8, 10, 12 or 19 PCR cycles.

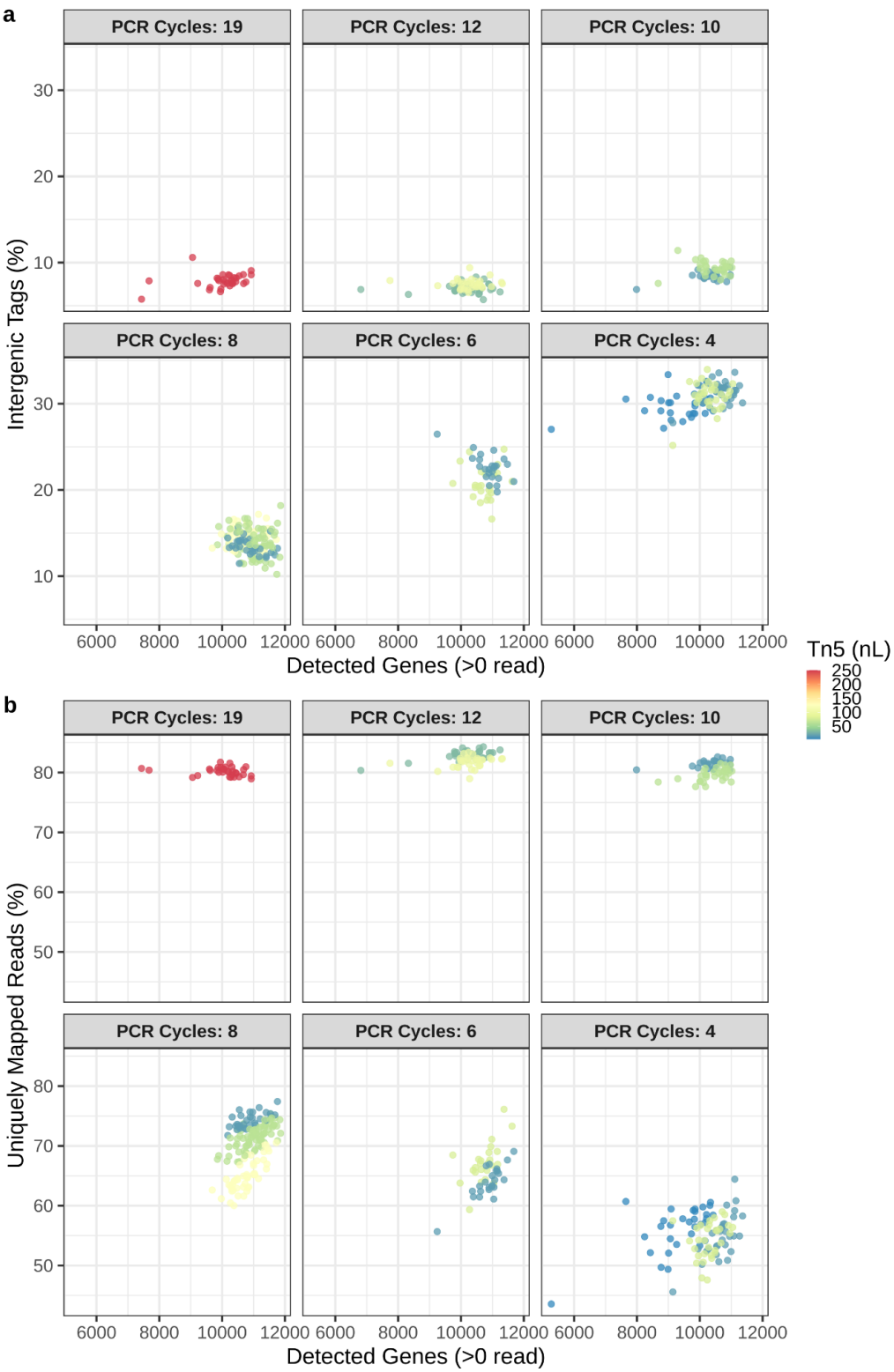


Fig. S8 | Relation between the number of detected genes a. with the percentage of intergenic reads **b.** the percentage of uniquely mapped reads**.** Color scale represents the amount of home-made Tn5 used for tagmentation, in nL.

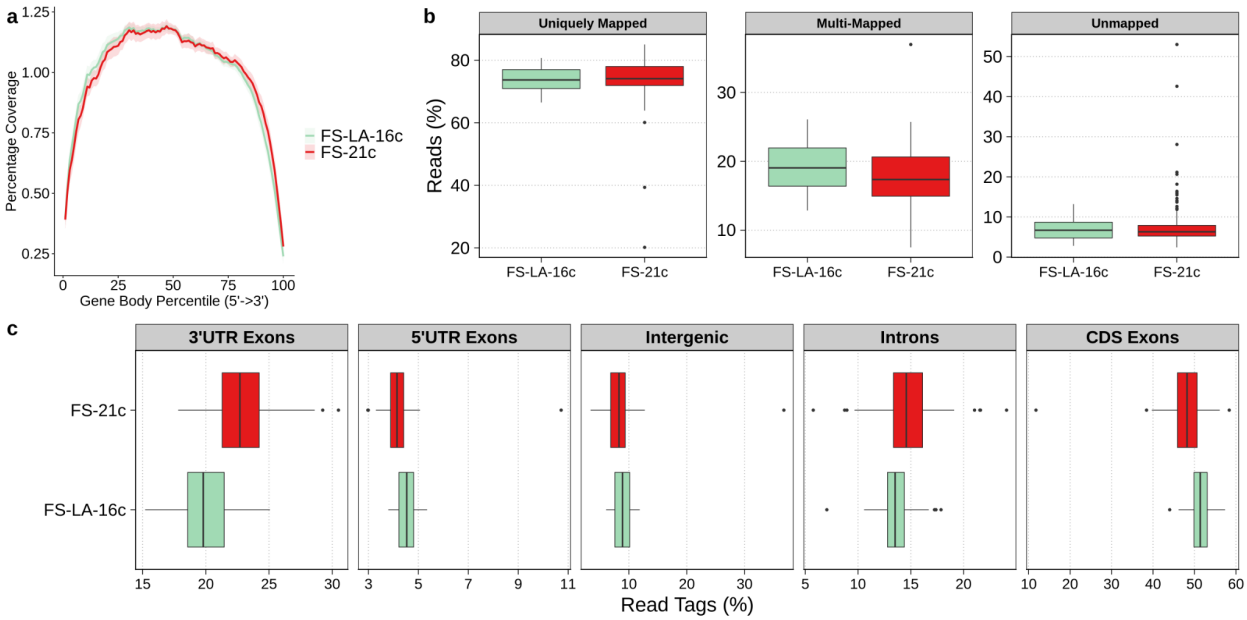


Fig. S9 | Overview of low amplification hPBMCs mapping statistics. a. FS gene body coverage of hPBMCs processed with FS (21 cycles, n = 180) or FS-LA (16 cycles, n = 107). **b.** STAR mapping statistics showing the percentage of uniquely mapped, multi-mapped and unmapped reads. **c.** Distribution of mapped reads between introns, intergenic regions or 3'-UTR / 5'-UTR / Coding sequence (=CDS) exons. Displayed in percentage of read tags and computed using ReSQC.

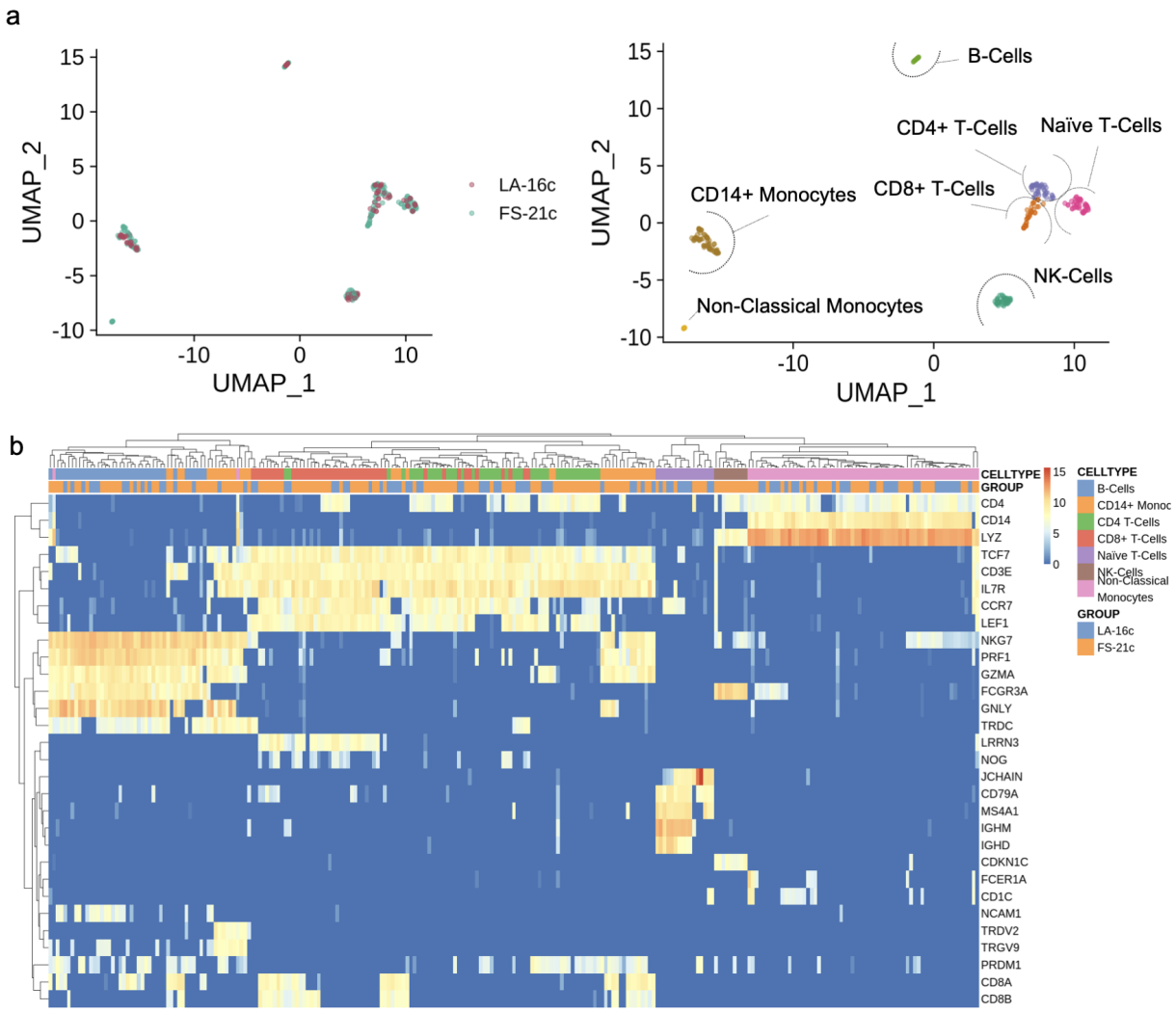


Fig. S10 | Dimension reduction and cell type assignment of hPBMCs processed with FS or FS-LA. **a.** Dimension reduction of hPBMCs processed with FLASH-Seq (FS-21c, n = 160) or FLASH-Seq low amplification (LA-16c, n = 95) protocol, highlighted by the method (left) or assigned cell types (right). Only cells with >100,000 raw reads and <25% unmapped reads were selected for this analysis. **b.** Heatmap of selected marker genes expressed by cells processed with either FS-21c or FS-LA-16c.

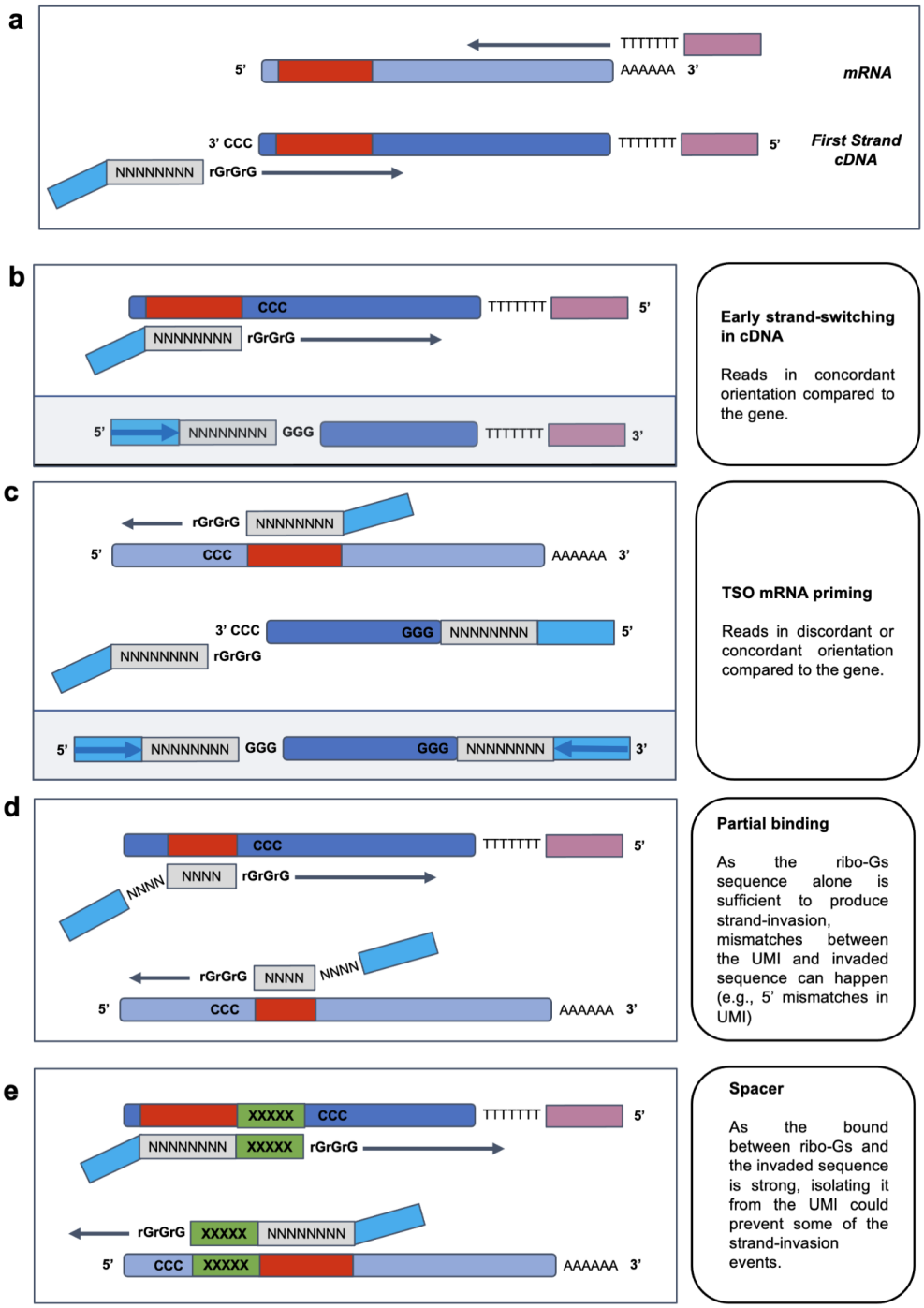


Fig. S11 | Strand-Invasion mechanisms. a. Regular RT and first-strand cDNA synthesis in SS3 protocol. **b.** cDNA Strand-invasion. The UMI and rGrGrG motif bind to a sequence inside of the cDNA **c.** TSO mRNA priming**.** The TSO replaces oligo-dT to prime the RT. **d.** Example of 5' partial match between the UMI and the cDNA or mRNA. **e.** Example of TSO invasion when a spacer sequence is used to separate the rGrGrG motif from the UMI.

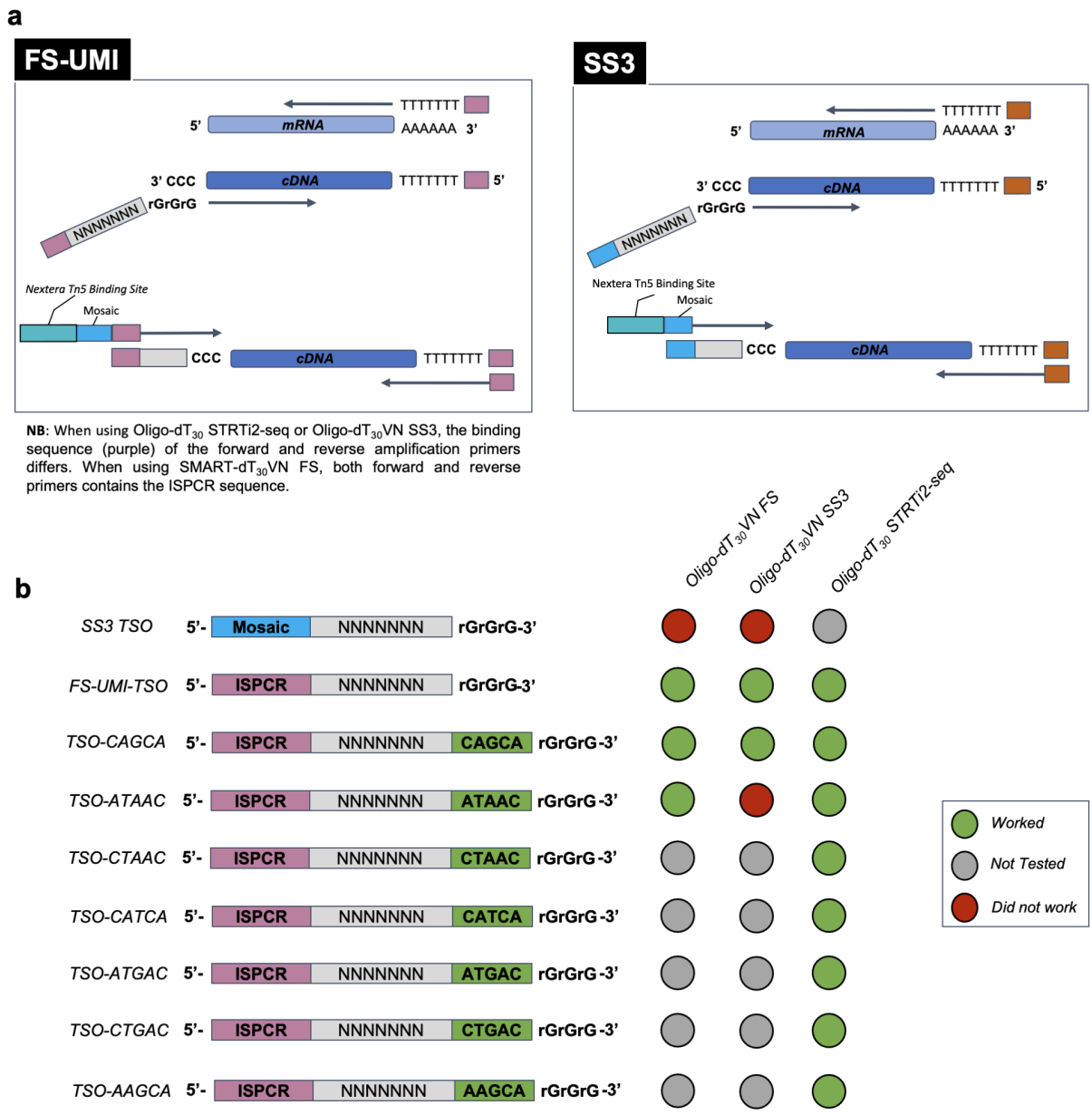


Fig. S12 | Schematic representation of the tested oligo-dT / TSO combinations. a. Reverse transcription and template-switching for the FS-UMI and SS3 protocols. **b.** Combinations of TSO and oligo-dT tested in this study.

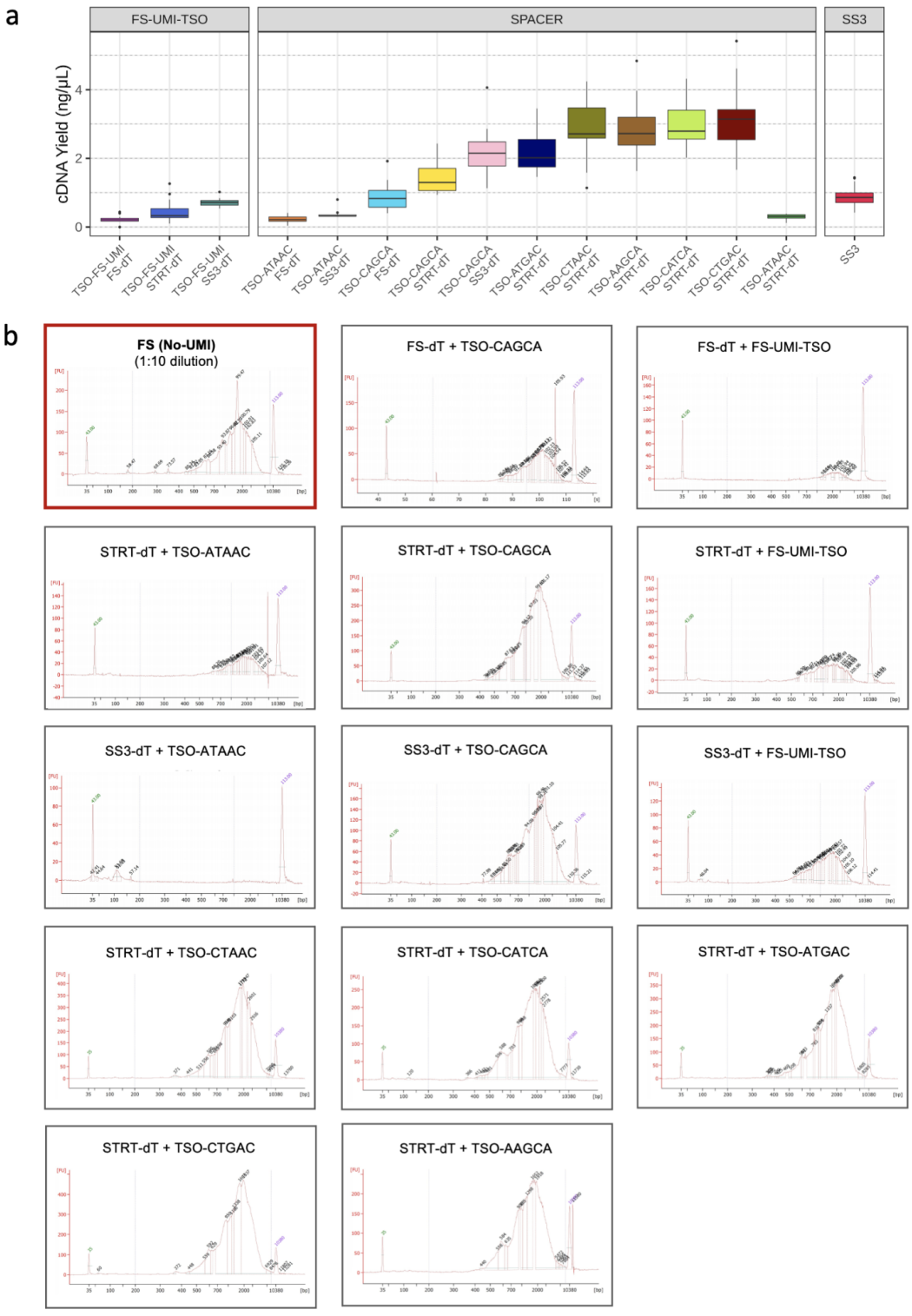


Fig. S13 | cDNA yield and Length distribution of the UMI-TSO / Oligo-dT combinations. a. cDNA yield **b.** Bioanalyzer traces of a representative sample of each UMI-TSO / oligo-dT combination. Control FS performed using regular TSO (= devoid of UMI) is highlighted in red.

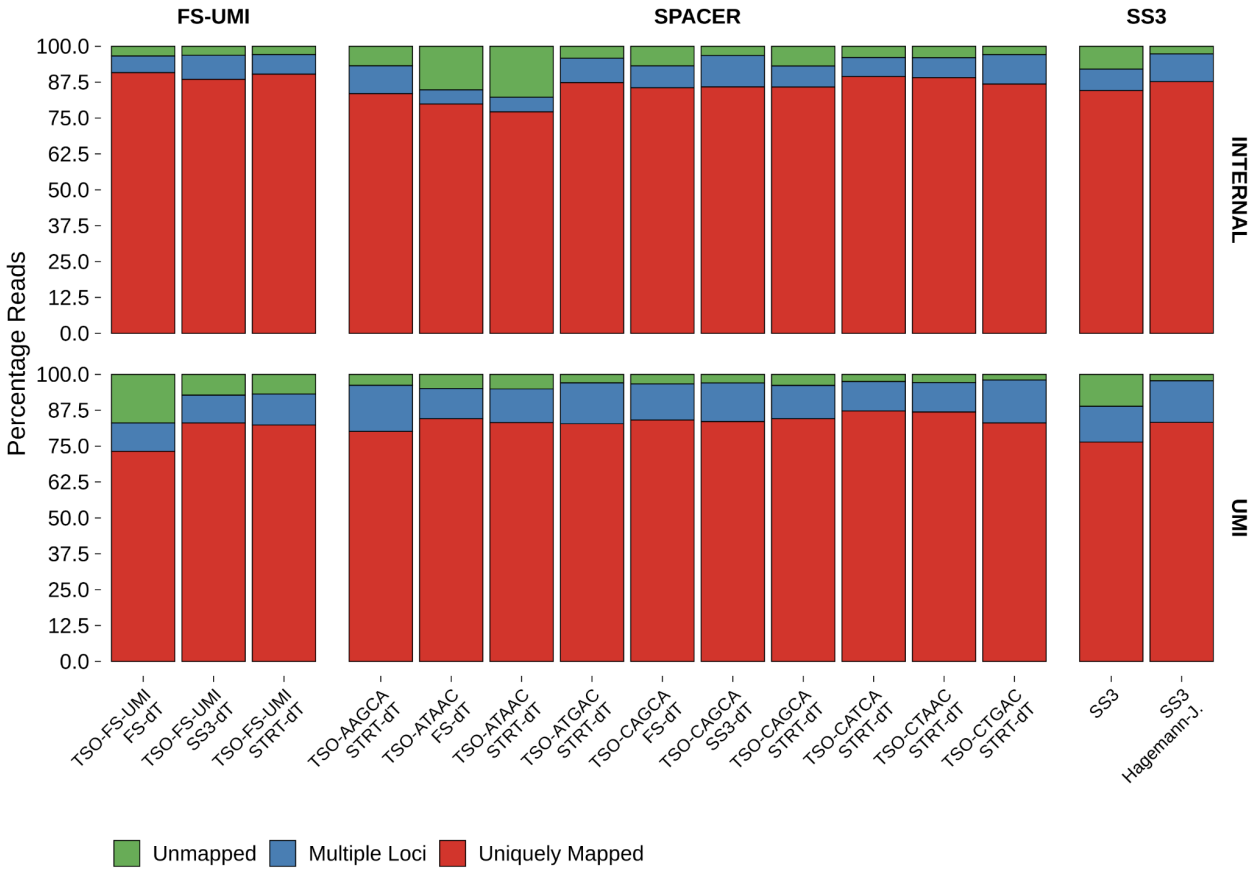


Fig. S14 | Internal and UMI reads mapping statistics of the different UMI-TSO / oligo-dT combinations. STAR mapping statistics showing the percentage of uniquely mapped, multi-mapped and unmapped reads using either internal or UMI-reads.

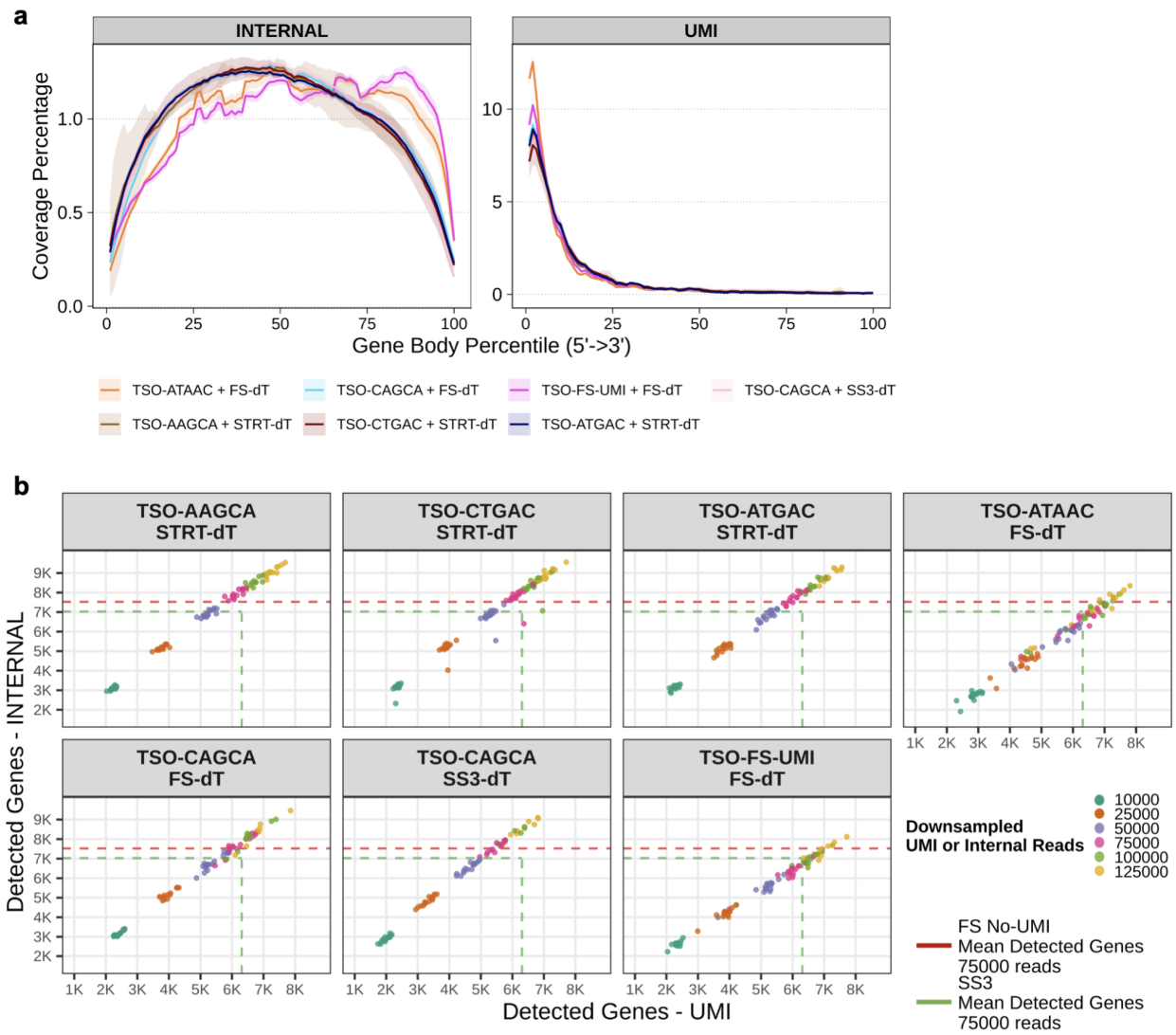


Fig. S15 | Addition of UMI to FS protocol and evaluation of oligo-dT - TSO combinations (remaining conditions of Fig 3). **a.** Gene body coverages. **b.** Relationship between the number of detected genes using UMIs and internal reads. UMIs and internal reads were downsampled to 10K, 25K, 50K, 75K, 100K, 125K reads. Dashed red line represents the mean number of detected genes using FS (n = 85) downsampled to 75K raw reads. Dashed green lines represent the mean number of detected genes from Hagemann-Jensen *et al* using UMI or internal reads, downsampled to 75K raw reads.

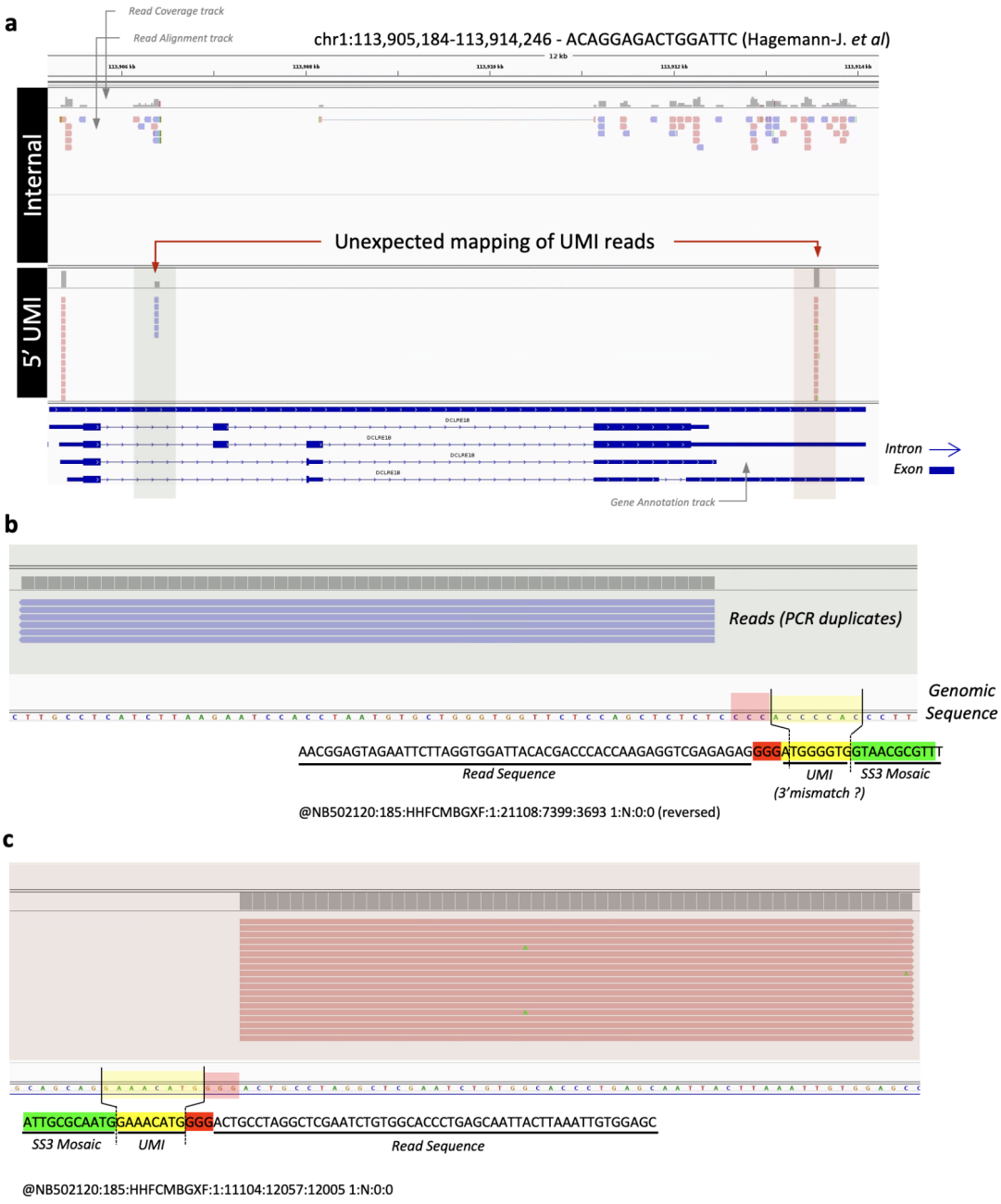


Fig. S16 | Strand-invasion example. **a.** Integrated Genome Viewer snapshot depicting the reads of the HEK 293T cell 'ACAGGAGACTGGATTC' (Hagemann-J. *et al*) mapping to gene DCLRE1B. The upper part of the picture shows the internal reads (= no UMI). The bottom part of the picture displays the 5' UMI reads. Two piles of pcr duplicated 5' UMI reads are located inside the gene body rather than at the 5' end. **b.** Zoom on the first read pile (green in a.). These reads are located in an intronic sequence and are in discordant orientation compared to the gene. The read sequence is displayed at the bottom. The read's GGG motif and UMI show almost a perfect complementarity with the genomic sequence. **c.** Same as b. but zooming on the second pile of PCR duplicated reads (orange in a.). A perfect match between the GGG motif and UMI the genomic sequence is observed.

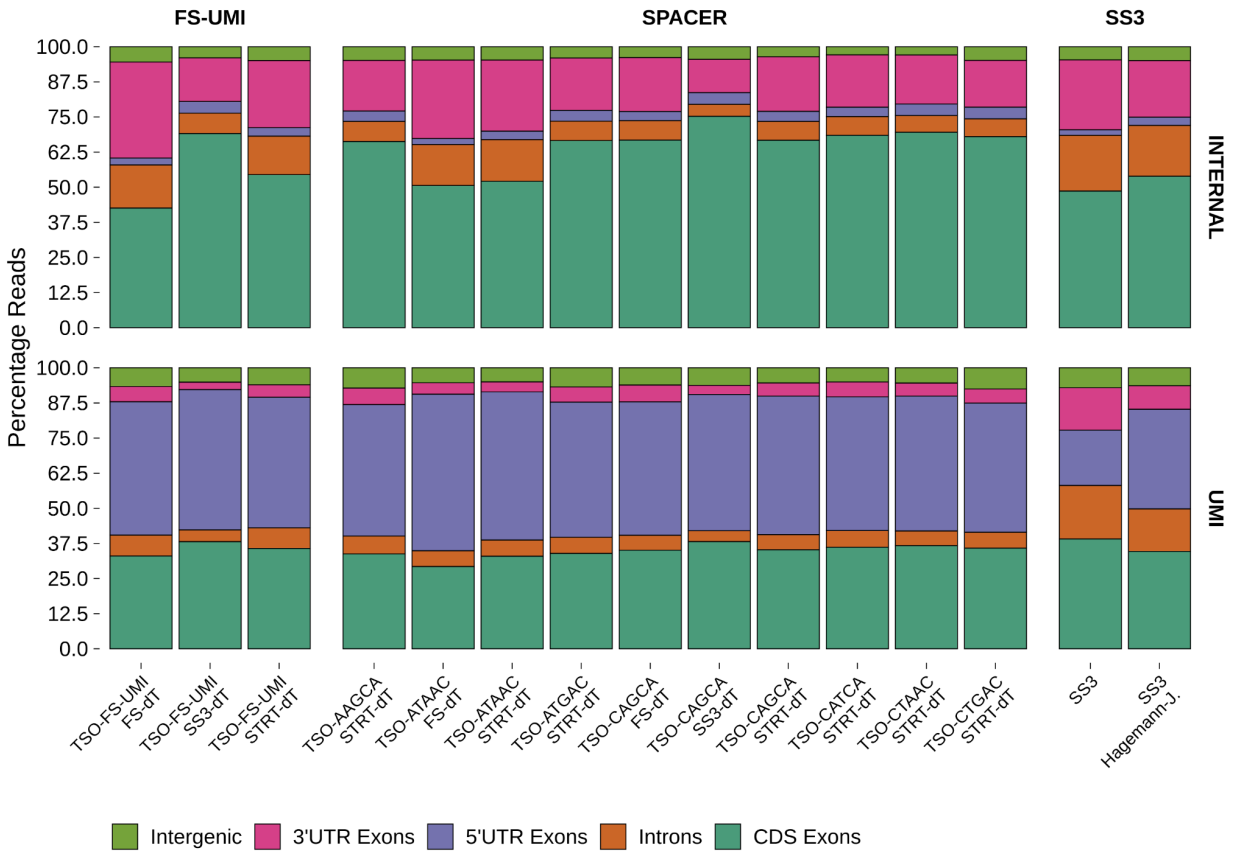


Fig. S17 | Mapped read distribution in FS-UMI and SS3 protocols. Distribution of mapped internal or UMI-reads between introns, intergenic regions or 3'-UTR / 5'-UTR / Coding sequence (=CDS) exons. Expressed in percentage of read tags and computed using ReSQC.

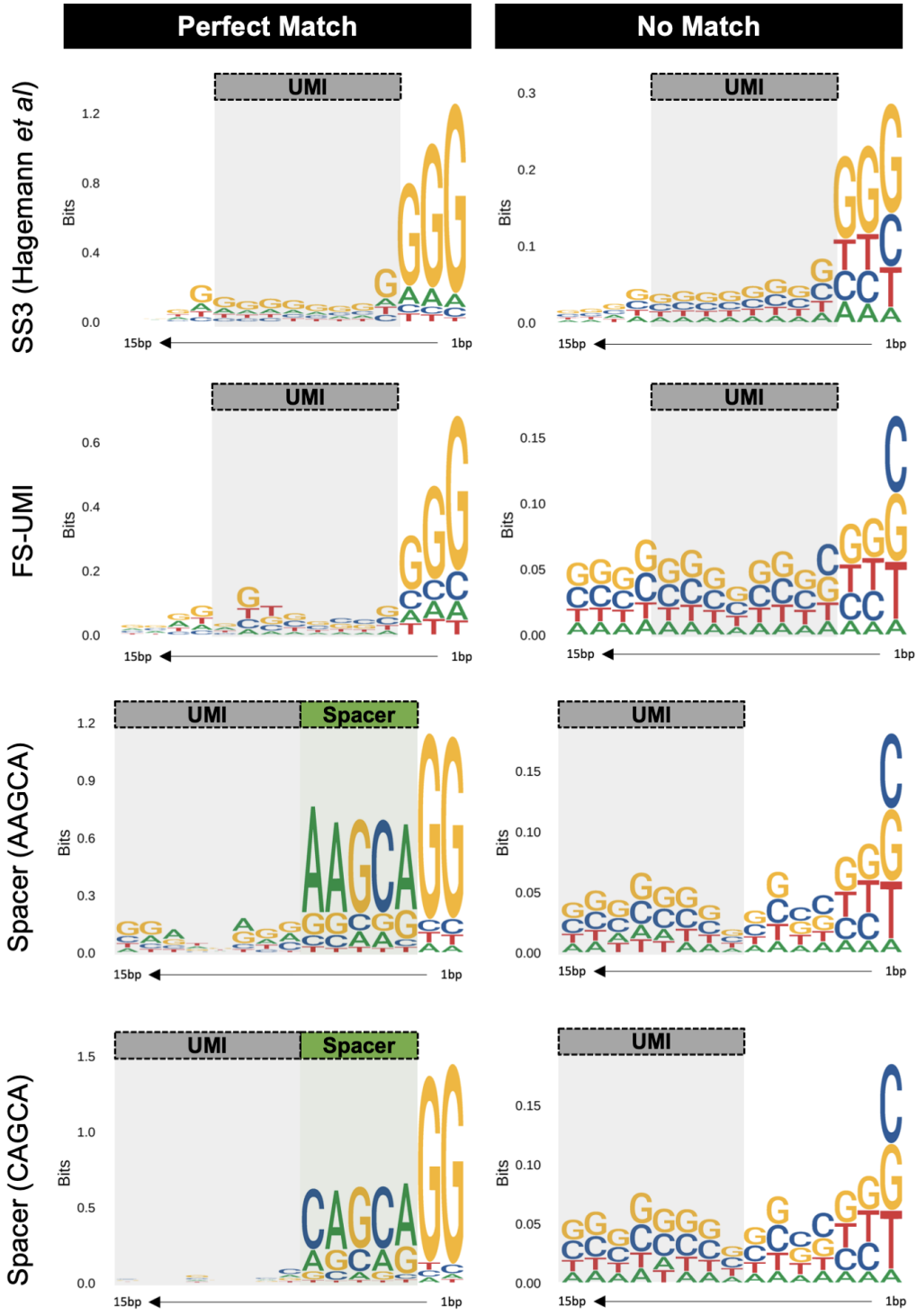
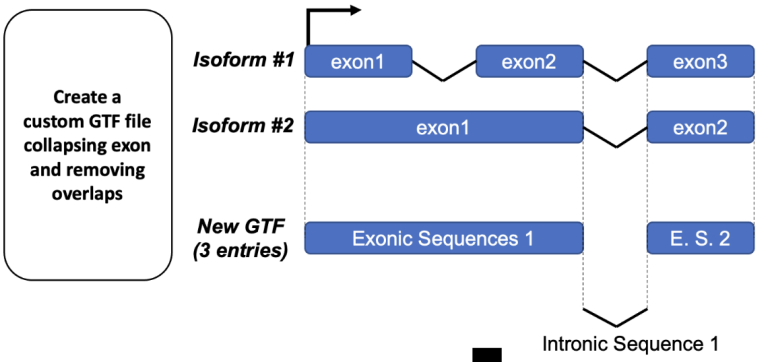
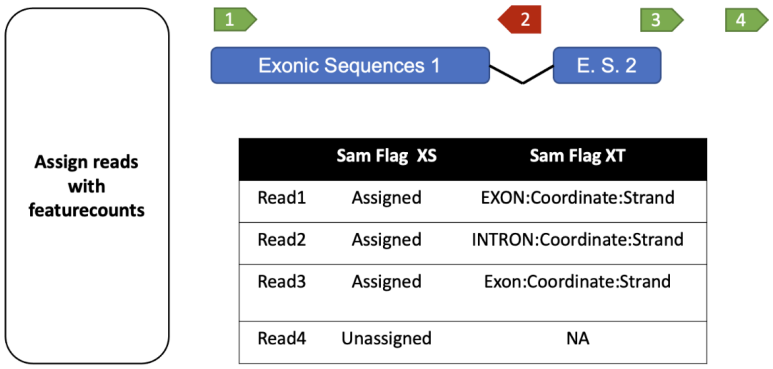
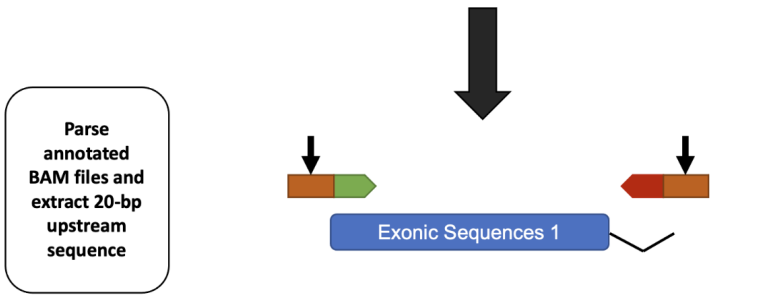


Fig. S18 | Nucleotide distribution of the genomic sequence adjacent to the read start (in bits). The sequences were split between those displaying a perfect match between the UMI and adjacent sequence (left) and those which did not (right). The expected position of the UMI is highlighted in gray. The expected position of the spacer is highlighted in green. Each row corresponds to one of four representative TSO + oligo-dT conditions.





	Sam Flag XS	Sam Flag XT
Read1	Assigned	EXON:Coordinate:Strand
Read2	Assigned	INTRON:Coordinate:Strand
Read3	Assigned	Exon:Coordinate:Strand
Read4	Unassigned	NA



Pos	UMI	Upstream	Sam Flag XS	Sam Flag XT
Chr:start-end:strand	ATTCGTCG	GTGCGTAAAGTGTCACACTT	Assigned	EXON:start-end:Strand
Chr:start-end:strand	GGCTGAAC	AGTGTGTACGGCTGAACGGG	Assigned	INTRON:start-end:Strand

Fig S19 | Schematic representation of the strand invasion detection script. First a custom GTF containing collapsed exon-intronic sequences of hg38 protein coding genes is created. Mapped 5' UMI reads are assigned to a feature using featurecounts. The SAM Flag XT contains the feature coordinates, strand and if the read is mapped unambiguously to an exonic or an intronic sequence. The BAM file is parsed using a custom R script and the 20-bp sequence adjacent to the read start is extracted. This sequence is compared with the UMI sequence, looking for a perfect match.