ONLINE SUPPLEMENTAL MATERIAL

Cell fixation and preservation for droplet-based single-cell transcriptomics

Jonathan Alles[1], Samantha D. Praktiknjo[1], Nikos Karaiskos[1], Stefanie Grosswendt[1], Salah Ayoub[1], Luisa Schreyer[1], Anastasiya Boltengagen[1], Christine Kocks[1,*], and Nikolaus Rajewsky[1,*]

[1] Systems Biology of Gene Regulatory Elements, Berlin Institute for Medical Systems Biology (BIMSB), Max Delbrück Center for Molecular Medicine in the Helmholtz Association (MDC), 13125 Berlin, Germany

* To whom correspondence should be addressed. Tel: +49-30-9406 2999; Fax: +49-30-9406 3068; Email: rajewsky@mdc-berlin.de

Correspondence may also be addressed to christine.kocks@mdc-berlin.de

The authors wish it to be known that, in their opinion, the first 3 authors should be regarded as joint First Authors.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure S1. Computational cell selection and cell concentration-dependence of the numbers of transcripts and genes detected per cell.

(A to C) Data correspond to the Drop-seq experiments shown in Figure 2 and Supplemental Figure S2.

(A) Identification of cell barcodes associated with single-cell transcriptomes in a pool of amplified single-cell libraries. Drop-seq involves Poisson-limited dilution of cells implying that the great majority of beads (more than 95% under our Drop-seq conditions) is not exposed to cells, only to ambient RNA. To identify the cell barcodes associated with cellular transcriptomes, cell barcodes are plotted in decreasing order of reads against the cumulative fraction of reads. An inflection point (vertical dotted line) was observed, roughly in agreement with the number of cells expected on the basis of barcoded bead input used for cDNA amplification: 1000 (Live and Fixed), 500 (Fixed 2 (1 week)), and 1000 (Fixed 2 (3 weeks); human-mouse cell doublets were removed computationally.

(B, C) Dependence of median number of genes or UMIs per cell on the cell concentration used in Drop-seq. Plotted are the cell numbers (as estimated from the plots in Suppl. Figure S2A above) in the four libraries against the median number of genes or UMIs observed. We noticed that the less cells were present in a Drop-seq library, the more genes or UMIs per cell we obtained. This was a larger trend observed over many more libraries (unpublished data). A linear model was fitted over the four libraries; shaded area is the 95% confidence interval on the fitted values. (B) Human (HEK) cells. (C) Mouse (3T3) cells. The libraries were computationally downsampled to comparable read depths.

Supplemental Figure S2. Fixed cell samples can be stored for weeks to give reproducible Drop-seq results.

(A and B) Drop-seq analysis of human (HEK) and mouse (3T3) cell mixtures, analyzed at a final combined concentration of 50 cells/µl. Scatter plots show the number of human and mouse transcripts (UMIs) associating with each cell. Each dot represents a cell identified as human- or mouse-specific (blue or red, respectively), using a threshold of 90% reads mapping to the human or mouse genome, respectively and a cut-off of more than 3500 UMIs expressed per cell (cells expressing less than 3500 UMIs are grey). Both Drop-seq experiments yielded single-cell transcriptomes that allowed clear species separation and a low % of cell doublets.

(A) Fixed cells stored at -80˚C for one week contained 1.2% human-mouse doublets.

(B) Fixed cells from the same batch as in (A) stored at -80˚C for 3 weeks contained a similar number of human-mouse doublets (1.2%).

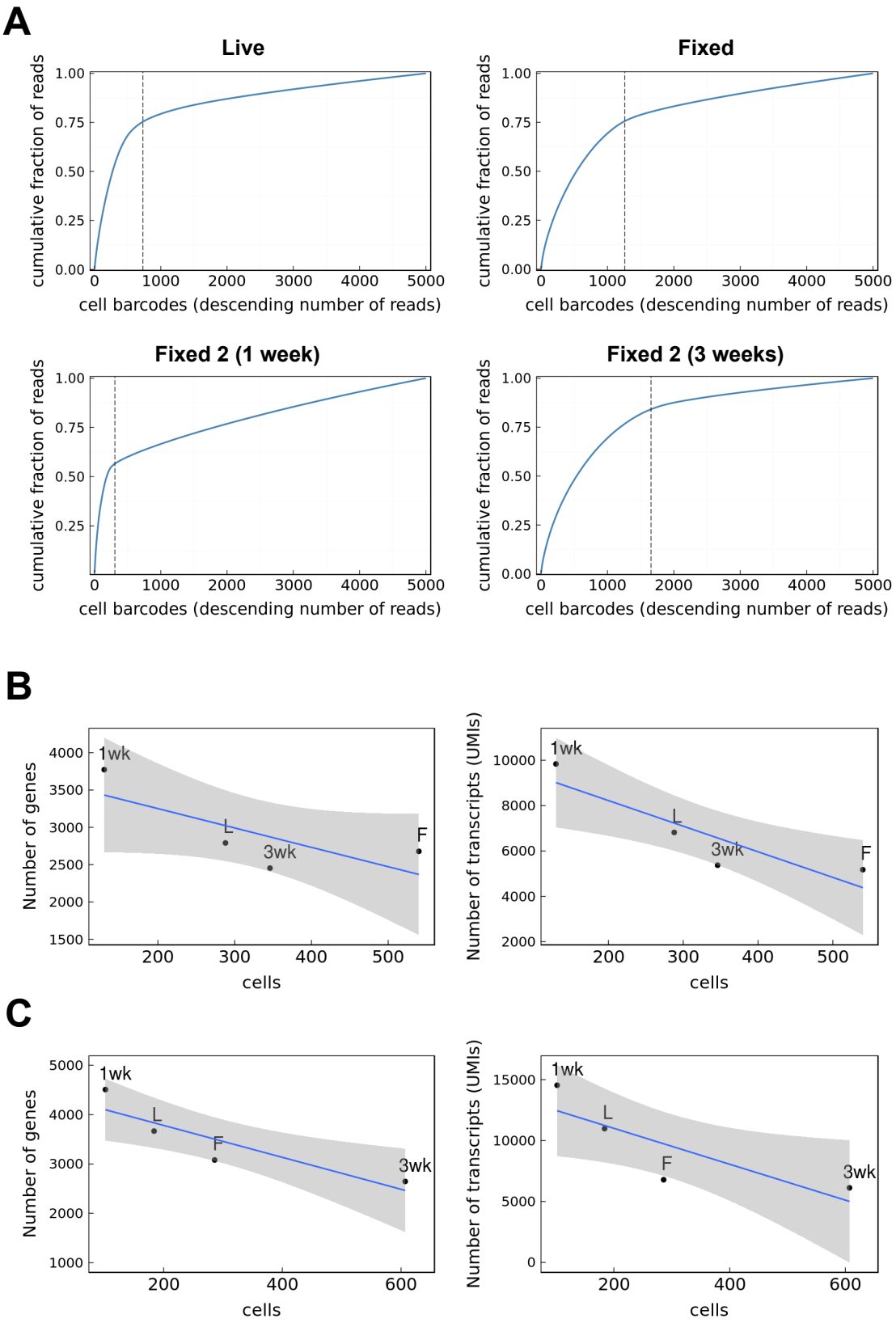
(C and D) Drop-seq sensitivity. Libraries were sequenced to a mean depth of 224,000 aligned reads per cell (Fixed 2 (1 week)) or 84,500 aligned reads per cell (Fixed 2 (3 weeks)). Violin plots show the distribution and the median of the number of genes (C) or transcripts (UMIs) (D) detected per cell, for the top 100 cells of the libraries depicted in Suppl. Figure S2A and S2B (Fixed 2 (1 week): 41 human and 57 mouse cells; Fixed 2 (3 weeks): 24 human and 72 mouse cells).

(E) Correlation between gene expression measurements from the two independent Drop-seq experiments described above. Dots represent gene expression levels. Human (HEK) cells, light grey; mouse (3T3) cells, dark grey. Drop-seq expression counts were converted to average transcripts per million (ATPM) and plotted as log2 (ATPM + 1).

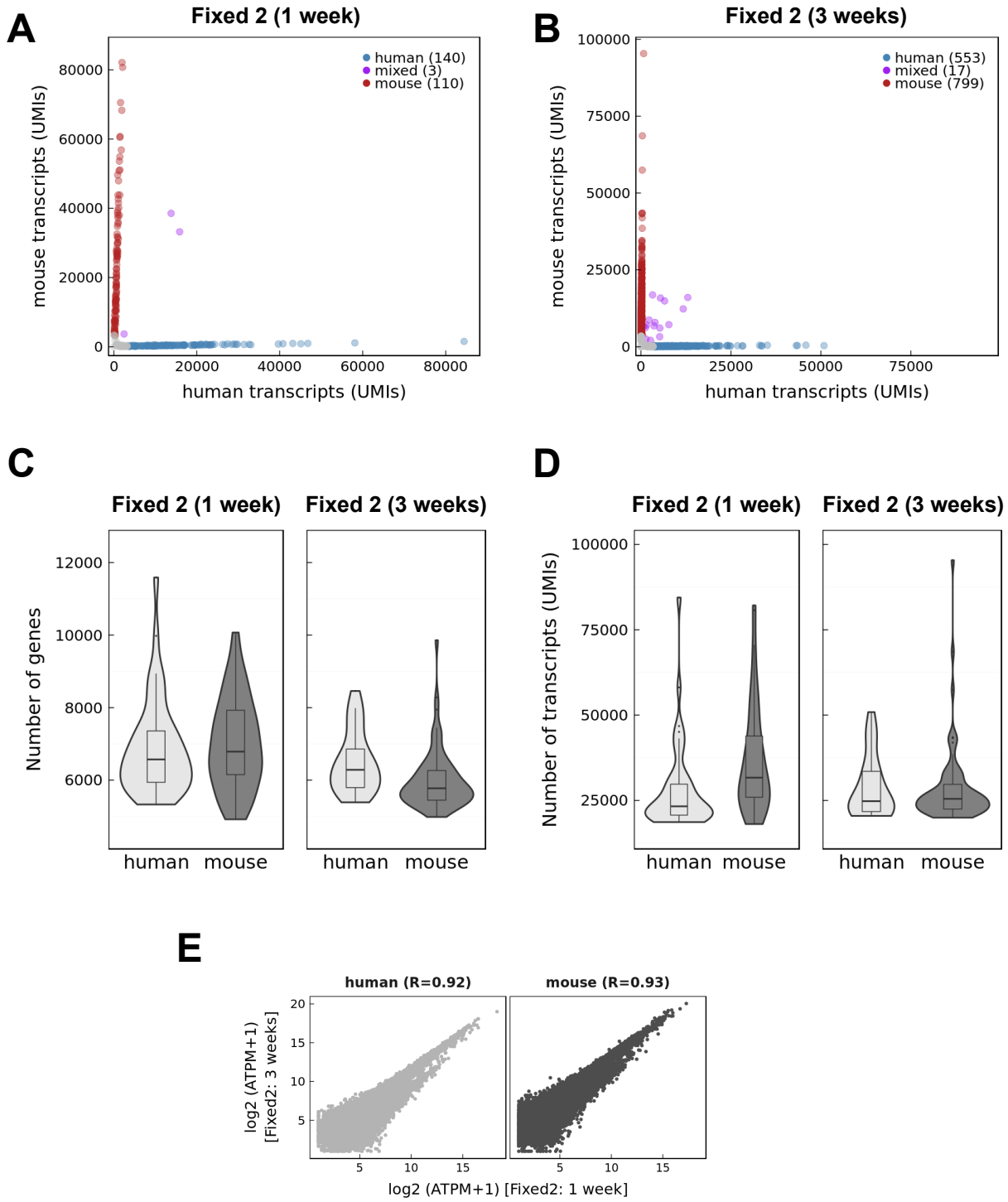
Supplemental Figure S3. Fixation causes only minor loss of non-mitocondrially-encoded RNAs by leakage from fixed cells.

Plot depicting the percentage of reads mapping to genes encoding non-mitochondrially encoded mRNAs. Stressed or broken cells exhibit loss of non-mitochondrially encoded mRNAs owing to cell leakage relative to mRNAs from mitochondrially encoded genes, which are relatively more protected from leaking out of broken cells (14). Our fixation protocol causes only minor loss of reads corresponding to cytoplasmic mRNAs. (A) Human (HEK) cells. (B) Mouse (3T3) cells.

Suppl. Figure S1



Suppl. Figure S2



Suppl. Figure S3