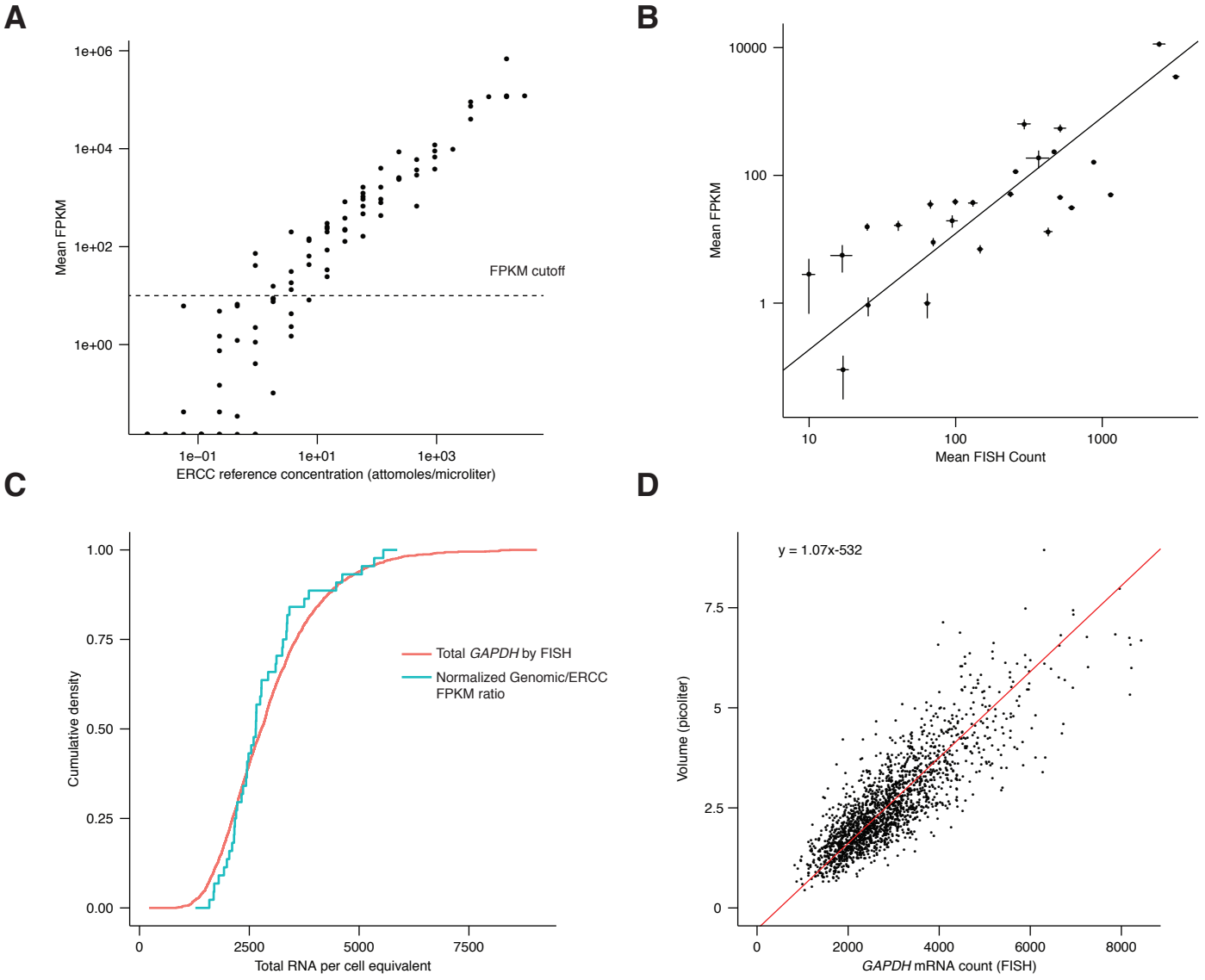
Supplementary Figure 14



Supplementary Fig. 14. Calibration of single-cell sequencing data.

A. Mean FPKM and known concentrations for each of the ERCC reference transcripts. Each point represents a single ERCC transcript and is an average over all 96 samples. Below 10 FPKM, we begin to see significant dropouts, and therefore choose and FPKM of 10 as our cutoff for “reliable” measurements. All other data in the manuscript is taken from genes having greater than 10 FPKM.

B. Mean count as measured by RNA FISH vs. mean FPKM from single-cell RNA sequencing. Each point represents a single gene and is an average over 44 single cells for single-cell sequencing, and an average over at least two biological replicates with at least 30 cells apiece for RNA FISH. These data suggest that an FPKM of 1 corresponds to approximately 23.2 transcripts per cell, as measured by RNA FISH in our cells, although the relationship between RNA FISH counts and FPKM scales nonlinearly (FPKM ~ (FISH)1.7, see Methods). We used this fitted relationship between RNA FISH count and FPKM to transform FPKM into transcript counts. Error bars represent SEM.

C. Comparison of “total RNA” distributions from single-cell sequencing and RNA FISH. Data represent a collection of total RNA measurements from single cells. We assume that total *GAPDH* mRNA counts by RNA FISH are proportional to total RNA. For sequencing data, we use the ratio of reads mapped to genomic loci to reads mapped to ERCCs as a proxy for total RNA. We scaled this ratio to have the same mean as the distribution of total RNA by RNA FISH. After scaling, the distributions are similar, suggesting that our method for measuring total RNA via the ratio of genomic to ERCC transcripts is sound.

D. Mapping between total RNA count (here, total *GAPDH* mRNA in single cells) and volume, as measured by RNA FISH. Each point represents a single cell. We use this mapping to convert total RNA from sequencing experiments to actual volume. The red line is the best fit, as computed by principle components analysis.