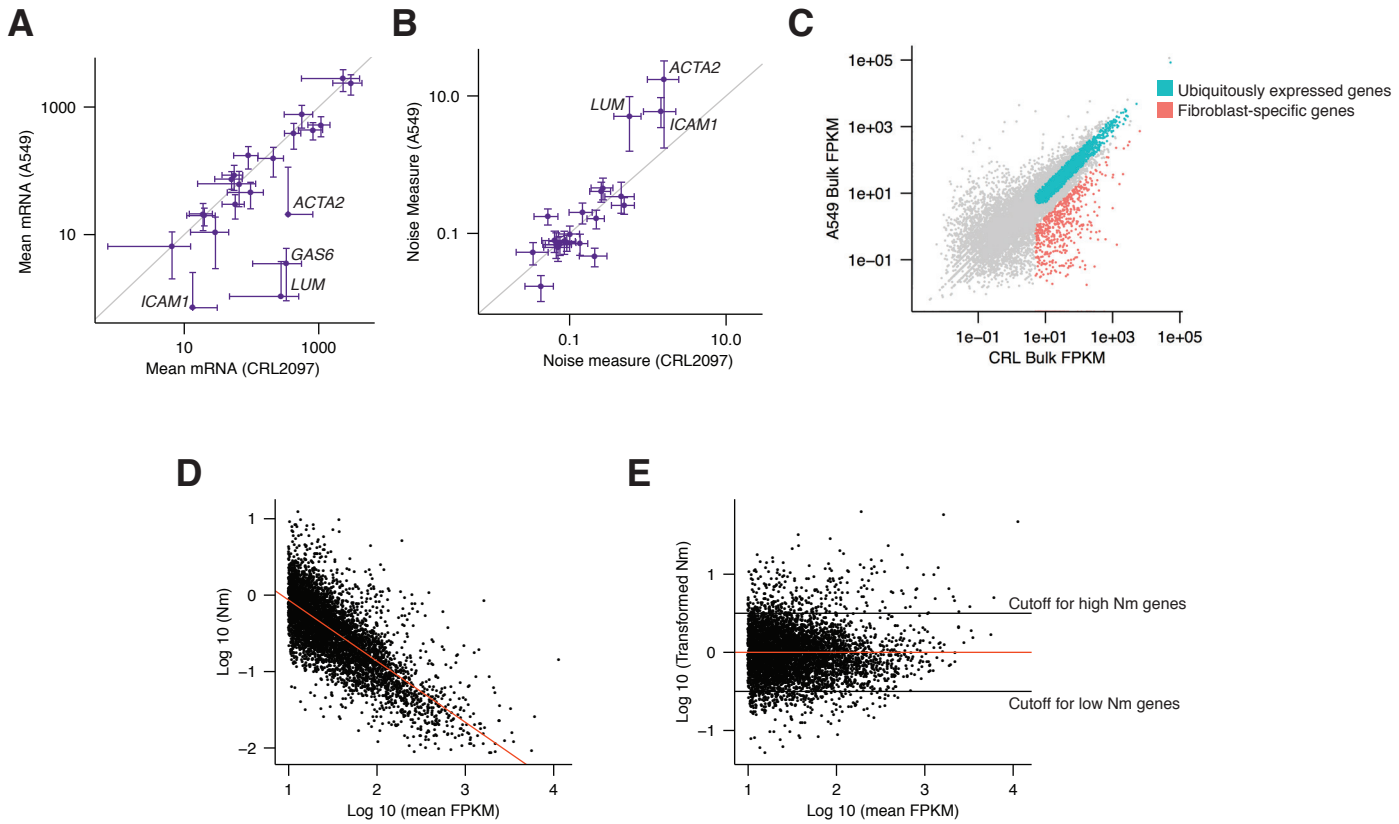
Supplementary Figure 20



Supplementary Fig. 20. Classification of noise level and cell type specificity.
A. Average mRNA counts in cycling hFF and A549 cells. Gray line indicates a 1:1 correspondence. Error bars represent standard error of the mean.
B. Volume-corrected noise measure in cycling hFF and A549 cells. Gray line indicates a 1:1 correspondence. Nm calculated by bootstrapping; error bars represent 95% confidence interval. Data for each gene is a combination of at least two biological replicates, with at least 30 cells per replicate.
C. FPKM measurements from bulk RNA-sequencing in hFF and A549 cells. Each point represents one gene. We classified genes as "ubiquitously expressed" if they had >5 FPKM in both cell types and differed by less than a factor of 2 in FPKM across the two cell types. We considered genes "fibroblast specific" if they had >5 FPKM in fibroblasts and their FPKM was greater than five times higher in fibroblasts that A549 cells.
D. Single-cell RNA-sequencing data in hFF cells. Each point represents one gene. We used the method described in Supplementary Fig. 14 and Methods to calculate Nm for each gene. We observe that higher abundance genes typically have lower Nm values. Red line indicates best fit line.
E. The same data as in D, but transformed to remove the volume-dependence from Nm. Red line here is the transformed fit line from D. We use this transformed data to select volume-matched "low Nm" and "high Nm" genes using a cutoff of Nm=0.5 and Nm=-0.5, respectively. We selected 307 high Nm genes and 257 low Nm genes. Note that these high Nm genes actually have a higher mean abundance (FPKM=196.5) than the low Nm genes (FPKM=55.4), thus showing that the observed differences in noise levels are not due to the overall increase in noise in genes of low abundance.