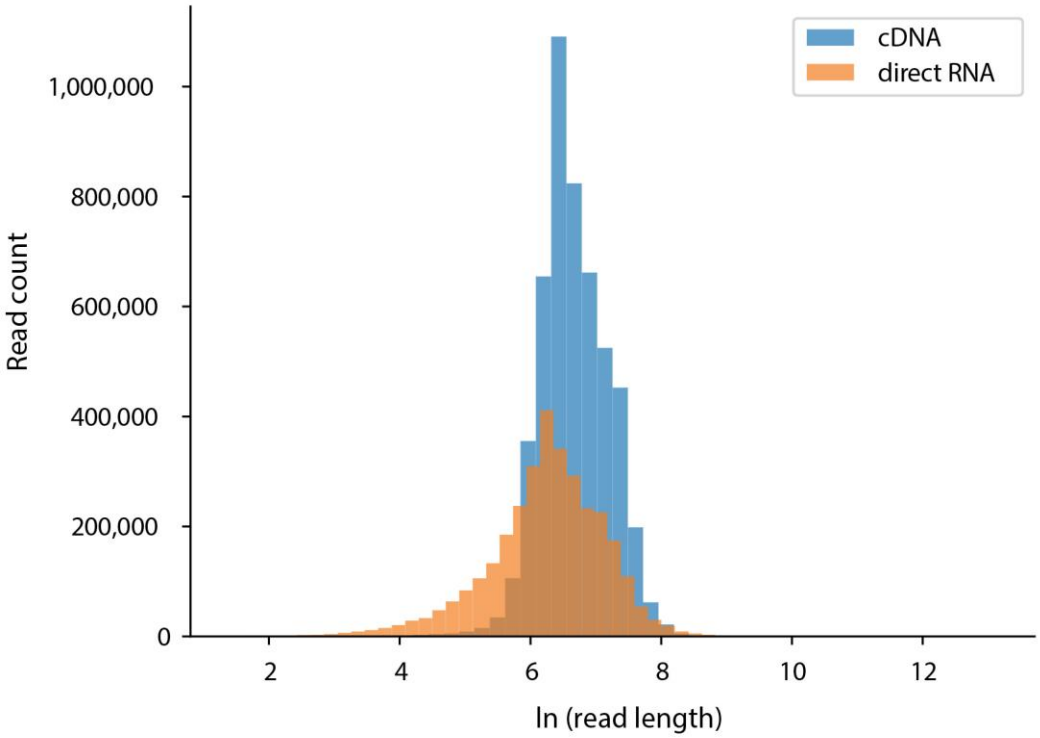
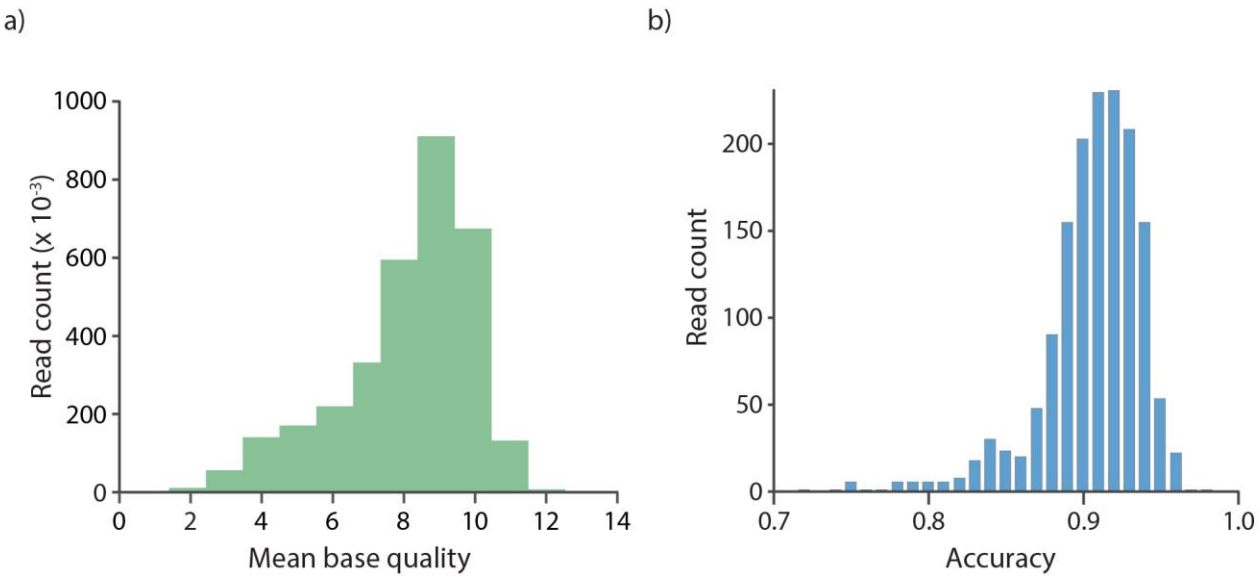


Supplementary Figure 1

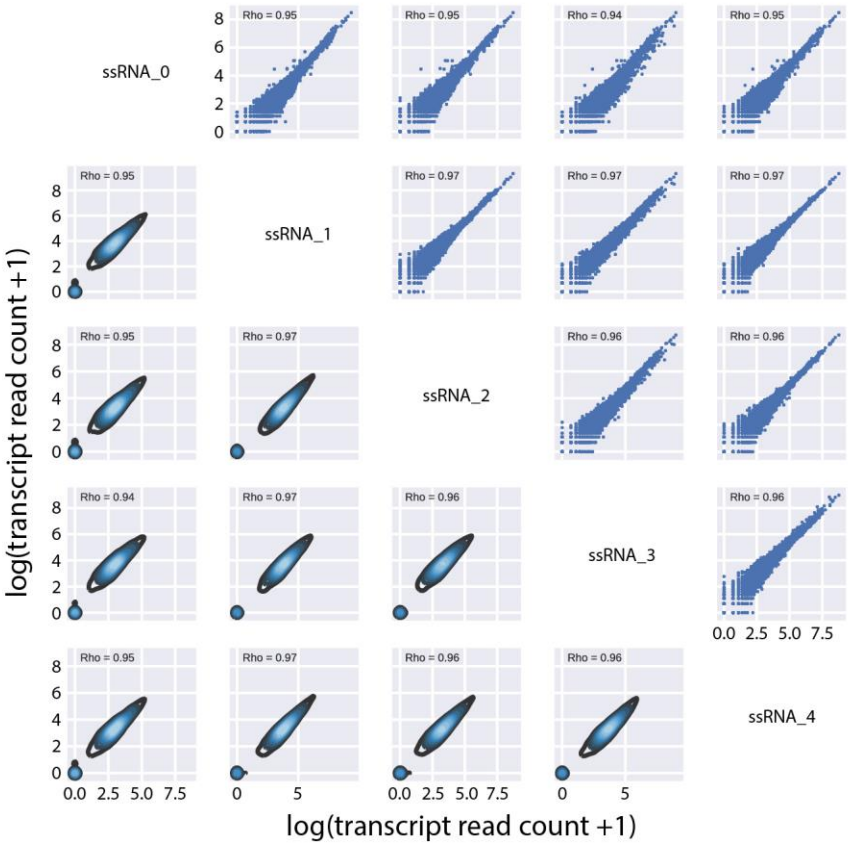
Read-length distributions for direct RNA and nanopore cDNA datasets



Supplementary Figure 2

Analysis of direct RNA method

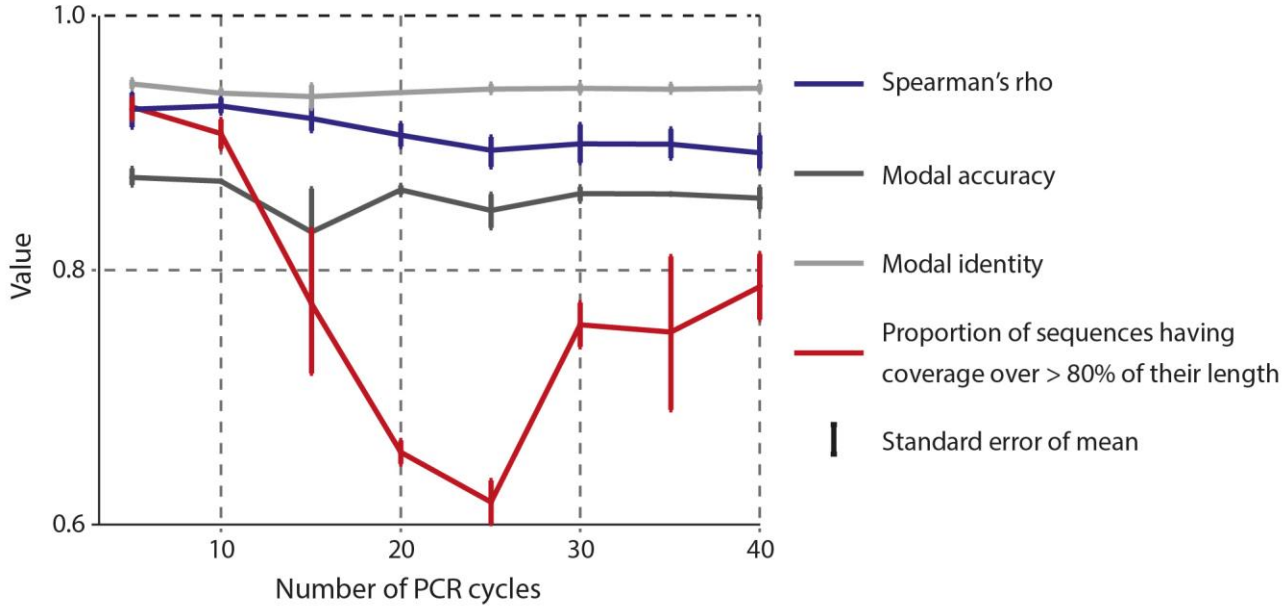
a) Distribution of mean quality values for all reads in the direct RNA yeast dataset. b) Distribution of read accuracies from the retrained direct RNA basecaller.



Supplementary Figure 3

Technical replicates of the direct RNA method.

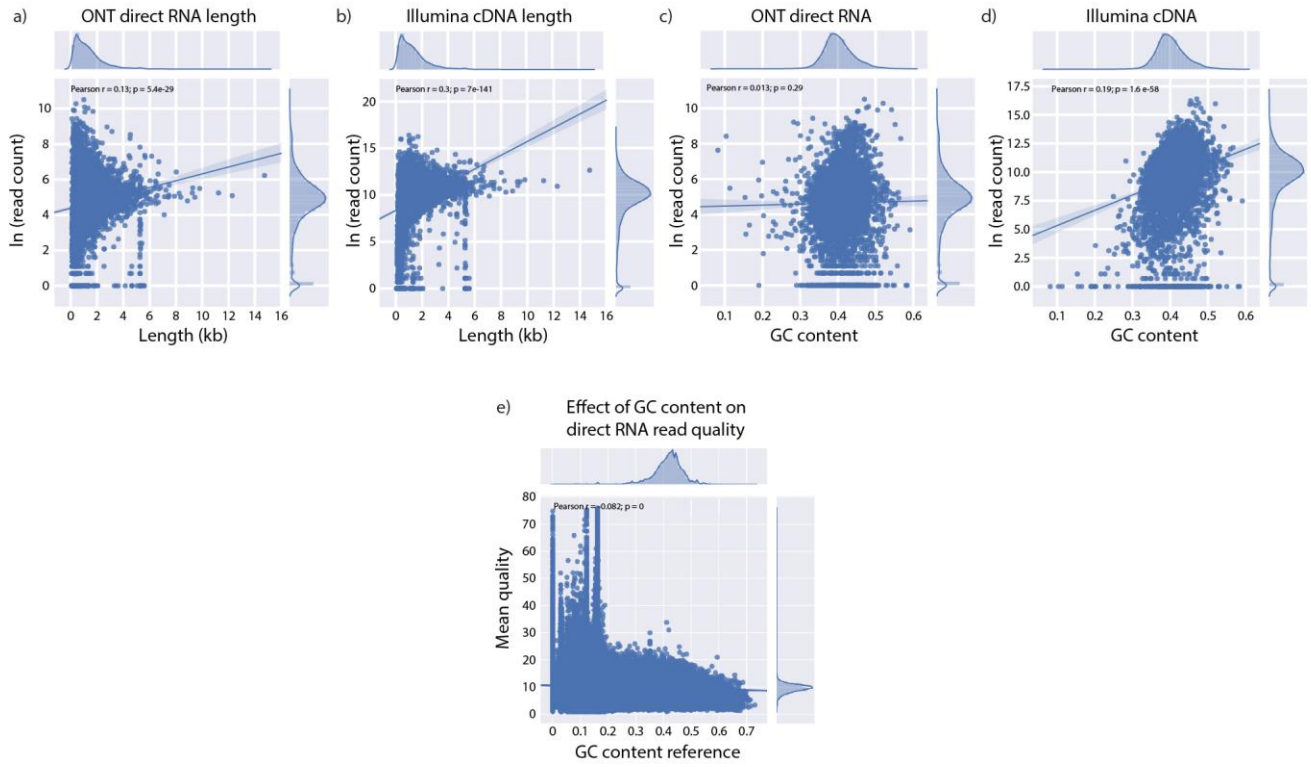
The correlation between read counts after mapping to the yeast transcriptome for 5 technical replicates of the Direct RNA method. The five technical replicates were separate library preparations of yeast run on separate MinION Chips. Above the diagonal are pairwise scatter plots and below the diagonal are pairwise density plots (Rho from Spearman's rank correlation is shown over each plot). Each scatter or density plot includes all transcripts in the annotation: n = 6713 transcripts.



Supplementary Figure 4

Effect of increasing number of PCR cycles

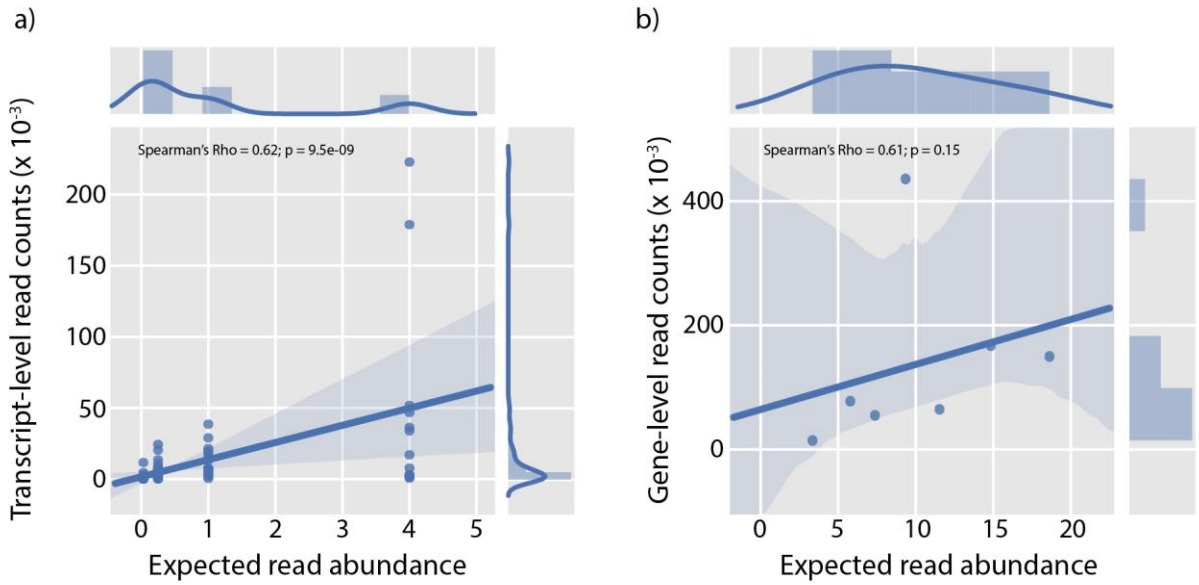
The effect of number of PCR cycles on bias, read length and deviation from expected read counts for ERCC spike-ins. Three independent replicates were performed at each cycle number totaling 24 separate nanopore cDNA sequencing runs. Error bars denote s.e.m..



Supplementary Figure 5

Direct RNA versus Illumina: comparison of bias.

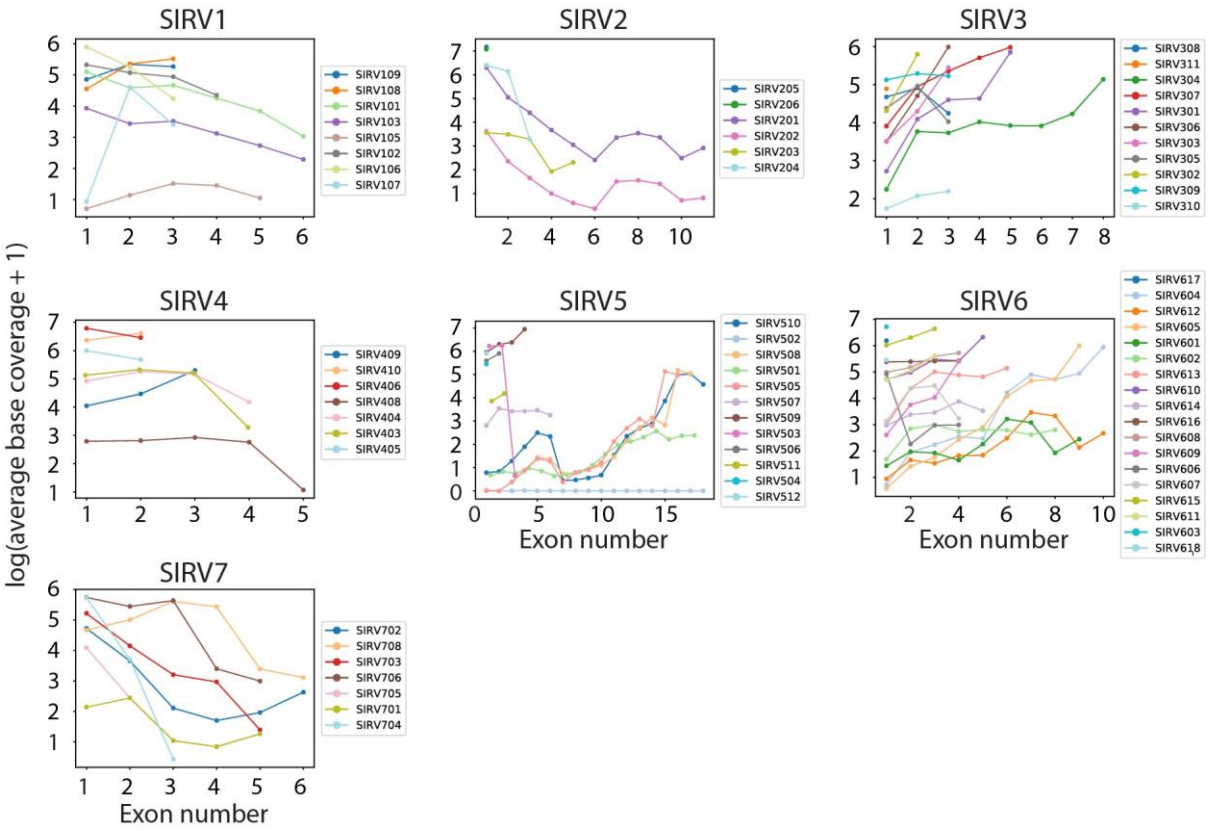
Correlation between read counts and transcript length for a) direct RNA (Pearson's r = 0.13, p = 5.4e-29) or b) Illumina (Pearson's r = 0.3, p = 7e-141) yeast datasets. Correlation between read counts and GC content for c) direct RNA (Pearson's r = 0.013, p = 0.29) or d) Illumina (Pearson's r = 0.19, p = 1.6e-58) yeast datasets. In each of (a-d), all transcripts were included: n = 6713 transcripts. e) Correlation between mean quality of aligned read portions and the GC content of aligned reference portions for direct RNA yeast dataset (Pearson's r = 0.082, p = 0, n = 2,777,523 alignments). The correlation coefficients and the corresponding two-sided p-values were calculated using the stats.pearsonr function from the scipy Python package.



Supplementary Figure 6

Gene-level and transcript-level correlations to SIRV control.

Reads aligned using the spliced-alignment strategy and correlations calculated a) at the transcript level (Spearman's Rho = 0.62, p = 9.5e-9, n = 69 transcripts) or b) at the gene level (Spearman's Rho = 0.61, p = 0.15, n = 7 genes) for the SIRV E2 dataset. The correlation coefficients and the corresponding two-sided p-values were calculated using the stats.spearmanr function from the scipy Python package.



Supplementary Figure 7

Coverage of individual exons in the SIRV E0 dataset.