Cell Reports, Volume *34*

Supplemental Information

A snapshot of translation in *Mycobacterium tuberculosis* during exponential growth and nutrient starvation revealed by ribosome profiling

Elizabeth B. Sawyer, Jody E. Phelan, Taane G. Clark, and Teresa Cortes

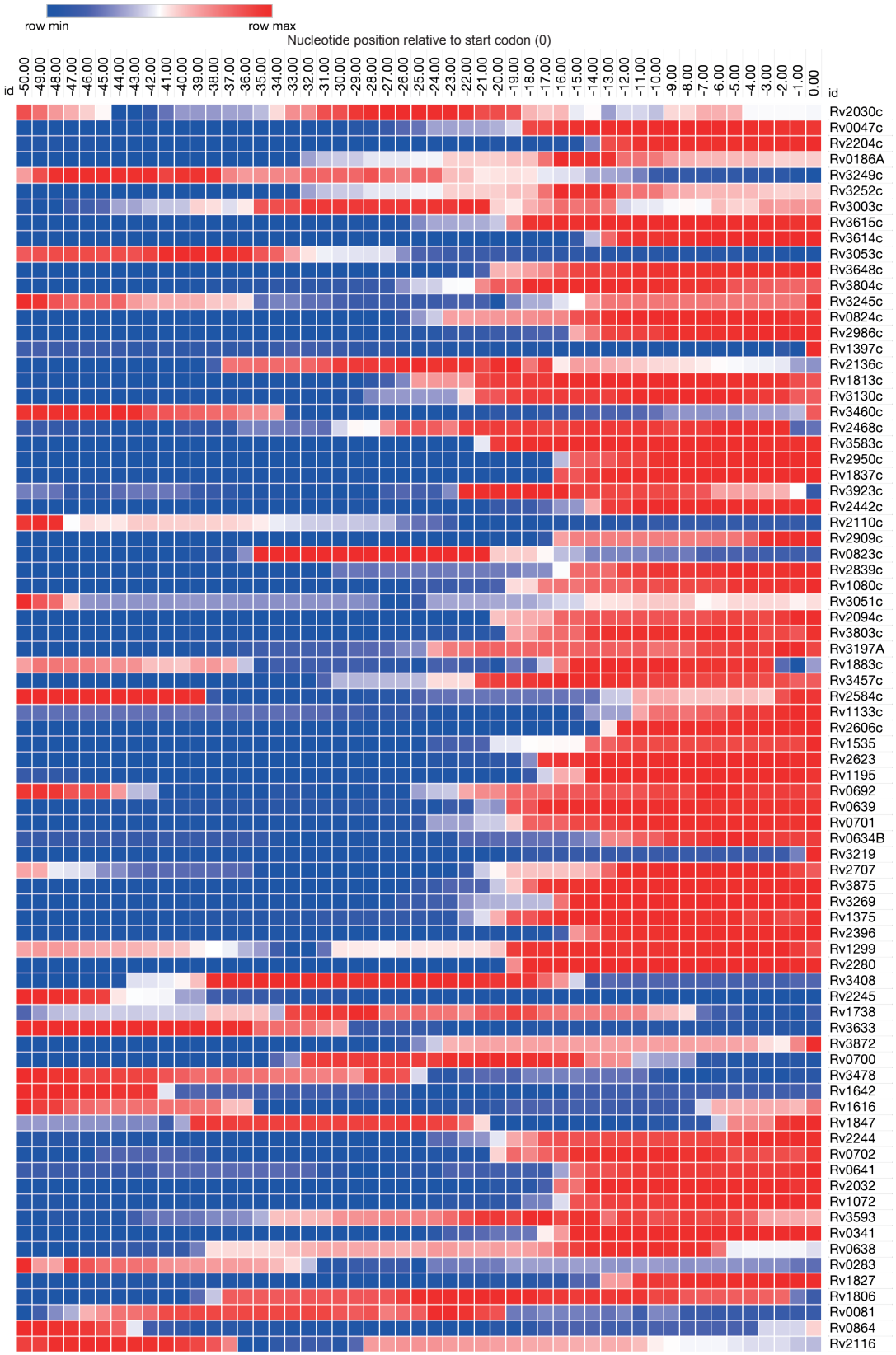

Figure S1. Heatmap of RPF density in the 5' UTRs of the 79 most highly expressed UTRs. Regions of highest density (red) mostly relate to ribosome recruitment, but some genes show read-through from the preceding gene or possible sORFs upstream of the CDS. Related to Figure 1C.

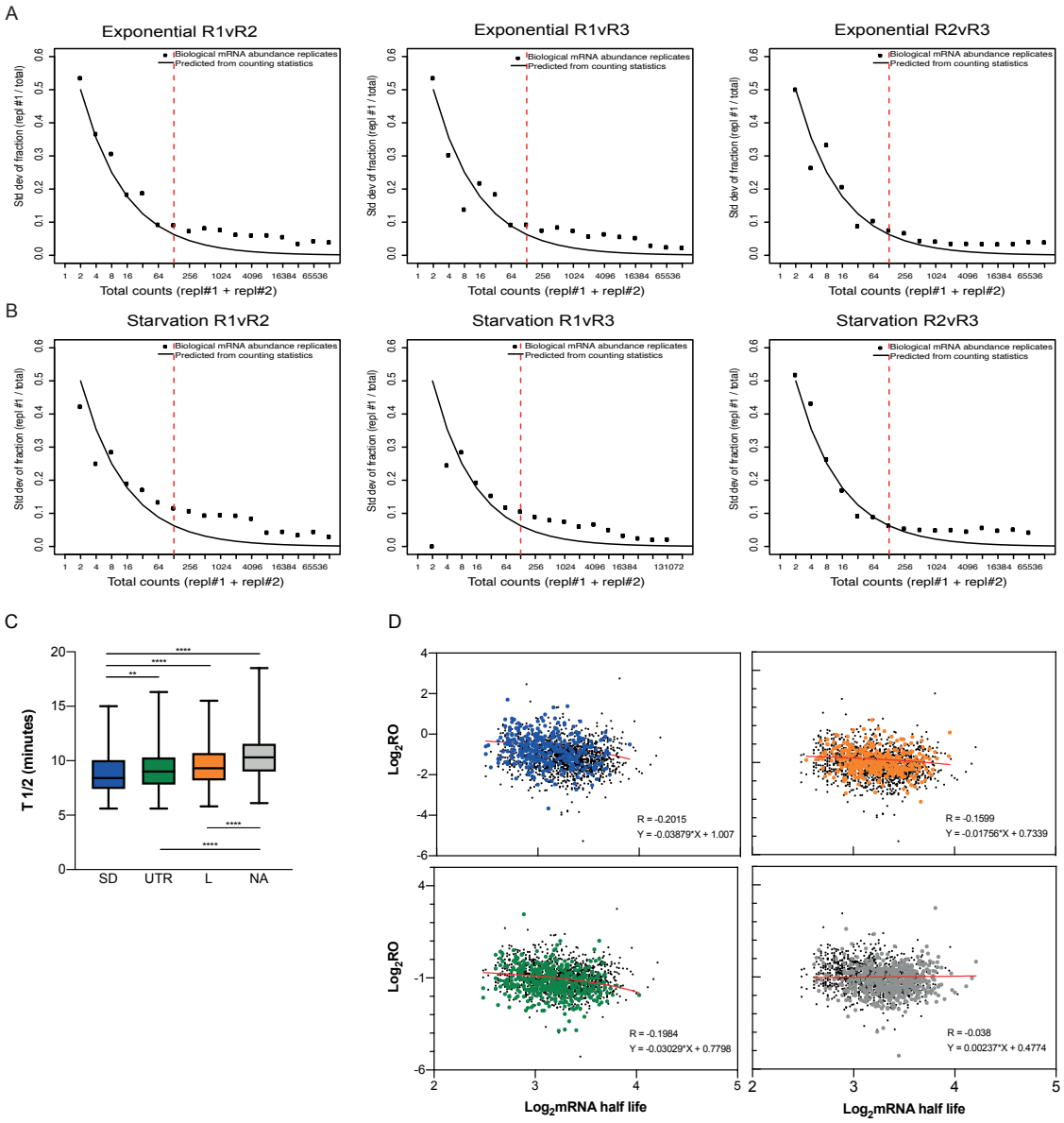

Figure S2. Effects of counting statistics on error in quantification and effect of mRNA half-life in RO. Pairwise comparisons of the three biological replicates for exponential (A) and nutrient starvation (B) were used to measure reproducibility. Fully independent biological replicates of mRNA abundance were used to measure reproducibility. For each pairwise comparison, the fraction of the total number of sequencing reads for each gene was calculated from one of the replicates. Subsequently, genes were binned based on the total number of reads and for each bin the standard deviation was calculated. The standard deviation was predicted for counting statistics using simple binomial partitioning of the total numbers of reads between the two replicates. A threshold of 128 total counts was chosen as a point where the inter-replicate variation approached its infinite-counts asymptote and counting statistics contributed little. (C-D) Effect of mRNA half-life in RO. (C) Box-plots of half-lives for Shine-Dalgarno (SD), UNSD, leaderless (L), and non-assigned (NA) mRNAs during exponential growth; Kruskal-Wallis test with Dunn's correction, **p < 0.0087 (data from ((Rustad et al., 2013)). (D) Correlation between RO and mRNA half-life for Shine-Dalgarno (blue), leaderless (orange), UNSD (green) and non-assigned (grey) genes.; black dots represent all genes, with those belonging to the UTR category overlaid in colour. Spearman's r values and equations from simple linear regression are indicated for each gene category. Related to STAR Methods and Figures 1E, 1F, 1G, 1H and 4B.

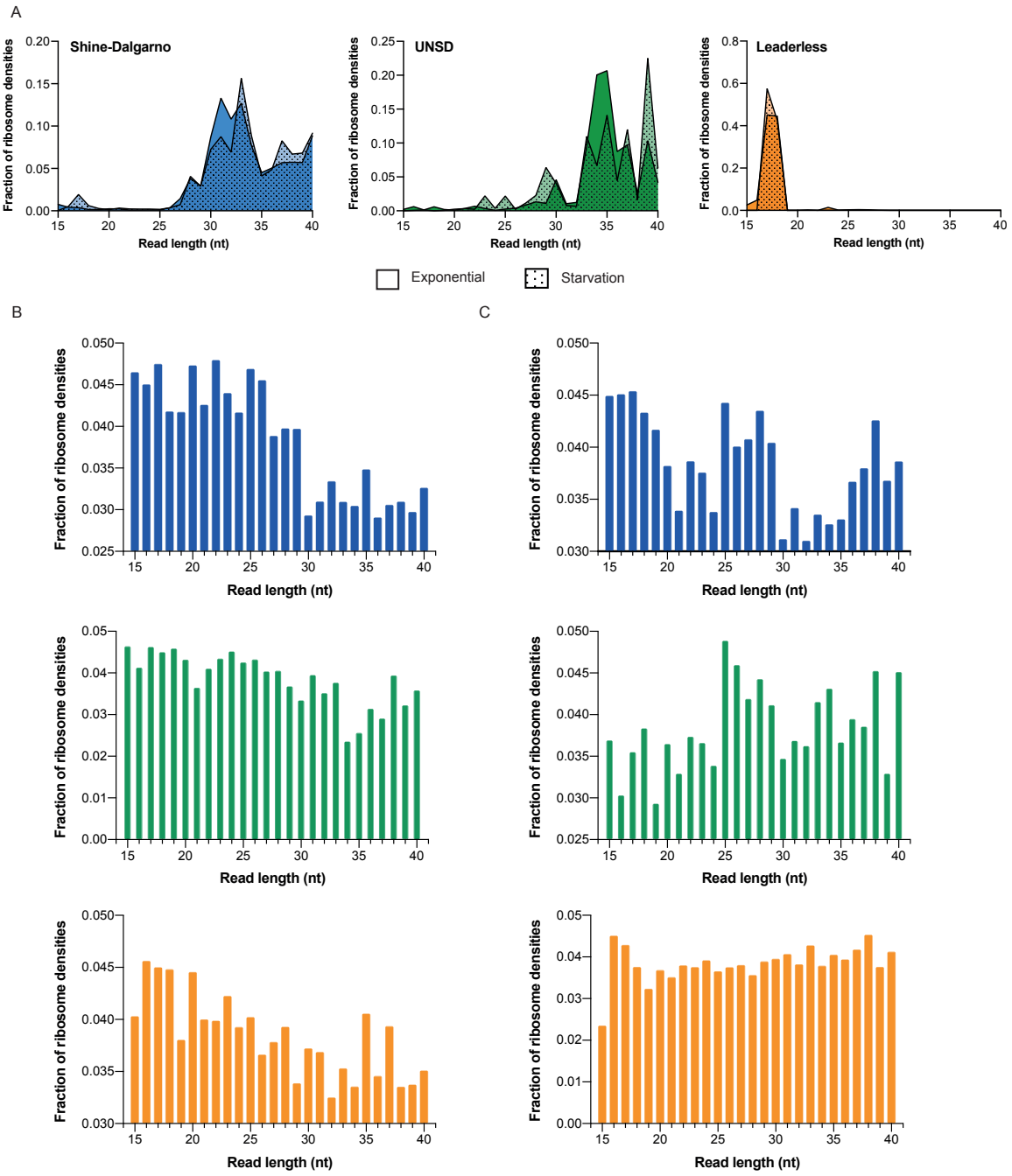

Figure S3. Read length distributions obtained for initiating and elongating ribosomes. (A) Read length distributions obtained for initiating ribosomes. Exponential growth shown in plain colour and starvation patterned; leaderless, orange; Shine-Dalgarno, blue; UNSD, green. (B) RPF length distribution of 3' assigned ribosome densities for elongating ribosomes within the open reading frame in Shine-Dalgarno (blue), UTR (green) and leaderless (orange) genes in exponential (A) and starvation (B). Only genes included in the metagene analysis were considered. Related to Figures 2A, 2B and 3B, 3C and 3D.

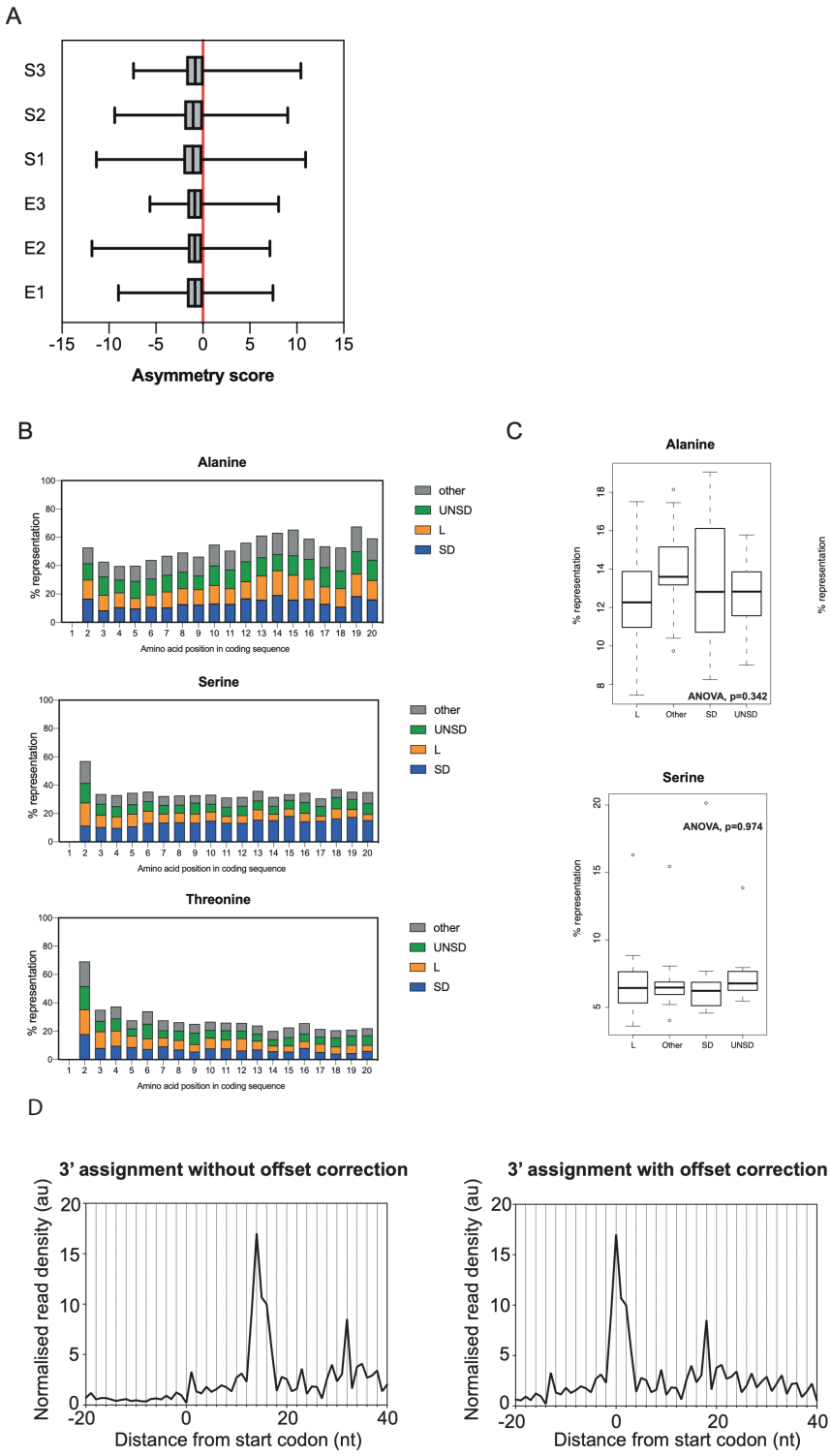

Figure S4. Effects of use of Chloramphenicol and offset correction. (A) Asymmetry scores for the exponential (E1, E2 and E3) and starvation (S1, S2, S3) Ribo-seq libraries. (B) Percentage representation of Alanine, Serine and Threonine codons across the first 20 amino acids of the coding regions for the four different gene categories (L, leaderless; SD, Shine-Dalgarno; UNSD, UTR no Shine-Dalgarno and other, not assigned).
(C) Boxplots comparing the median levels of % representation across the four gene categories for alanine, serine and threonine codons. Statistical test used for comparison and the associated p-values are indicated. (D) A 14 nt offset correction after 3' assignment places the start codon in the P-site of the ribosome, indicated by high ribosome density at the start codon as the initiation complex becomes elongation-competent. Related to STAR Methods and Figures 2A, 2C, 3B, 3C and 3D.