



Supplemental Figure S2. Barcodes have negligible effects on reporters expression. In these experiments, we performed transient electroporation of low complexity TRIP libraries (approximately 20,000 distinct barcodes) in Kc167 cells and measured the abundance of barcodes in the DNA and in the RNA 3 days after electroporation, with or without heat shock. Since the reporters are not integrated, the observed effect cannot be attributed to the local chromatin context, but only to the experimental noise and to the sequence of the barcodes. **(a)** The DNA/DNA scatter plot of barcode abundance for heat-shock promoter with and without heat shock shows a greater than 5000-fold dynamic range. Yet the measurements are accurate (correlation 0.998). **(b)** The RNA/RNA scatter plot of the same reporter with and without heat shock shows a correlation of 0.91, demonstrating that experimental noise is dominated by the cDNA counts. **(c)** The RNA/DNA scatter plot of the same reporter without heat shock shows that expression correlates with copy number in 99.5% of cases (highlighted in black). The 0.5% outliers whose expression is higher than expected are highlighted in blue. **(d)** The RNA/DNA scatter plot after heat shock shows the same trend. The highlighted barcodes were picked from panel **c**. These barcodes also have higher expression than expected in this experiment, showing that their sequence most likely influences the expression of the reporter. **(e)** The RNA/DNA scatter plot for promoter III shows a similar pattern. **(f)** In order to identify shared sequence motifs in the sequence of outlier barcodes, we trained a support vector machine (SVM) on the 4-mers extracted from the barcodes of the *Hsp70b* promoter (*i.e.* from panel **c**). The ROC curve shows that the trained model was unable to identify the outlier barcodes of

promoter III. We also compared the propensity of barcodes to form RNA secondary structures using pKiss (Janssen and Giegerich 2015) but the differences between outliers and non outliers were not statistically significant. In summary, a small fraction of the barcodes (less than 2% in every experiment) influences the expression of the reporters, but we could not identify any sequence feature at the origin of the effect, probably due to the reduced sample size of outliers.