SUPPLEMENTARY MATERIAL

Supplementary Figures

Supplementary Figure 1: Transcriptome structure dynamics during growth. Changes in transcriptome are visualized: (A) as a heat map ordered vertically in the time dimension from bottom (early log phase) to top (stationary phase); the color scale represents $\log_2$ ratio of transcript level changes during growth relative to the reference RNA, and (B) in the Gaggle Genome Browser. Genes in the forward and reverse strands are represented in yellow and orange, respectively. Blue dots represent reference RNA transcript signals (log scale), segmented using regression trees (red line). Differential expression is indicated as log ratios (cell density / reference RNA), upper panel and lower panel show changes in forward and reverse strand, respectively. Each track corresponds to the indicated cell density, measured as the optical density at 600 nm. Transcribed regions that were significantly up-regulated or down-regulated are indicated in red or green lines, respectively.

Supplementary Figure 2: **Untranslated regions of *H. salinarum NRC-1* transcripts**. Distribution of the distances of TSSs to the annotated start codon (start distance, A) and TTSs to the annotated stop codon (stop distance, B) were computed. Comparison of 5' and 3' UTRs detected in *H. salinarum NRC-1* (vertical axis) with published data for 19 genes with experimentally determined TSS and 14 genes with known TTS in *H. salinarum R-1* (C and D) ((Brenneis et al., 2007); horizontal axis). Black lines have intercept = 0 and slope = 1. Vertical axis (start distance and stop distance) shows respectively the distance (upstream is negative) between the TSS and TTS determined in this work and those of (Brenneis et al., 2007). (C) 5'UTR. The red line shows the least-squares fit to the 5' UTR distances (intercept = -9.4 nt; slope = -1.00), indicating excellent agreement, but revealing that we see a slight bias toward longer 5' UTRs equal to one-half the tiling array spacing (10 nt). (D) 3'UTR. We observe that the 3' UTRs detected in our experiments are usually longer, as explained in the main text. (E-H) Predicted TSS for (E) *prrlV2*, (F) *flaA1a*, (G) *fdx* and (H) *gvpD2* are indicated by arrows, comparison with experimentally determined 5' UTR length validates the predicted TSS within

the error of 20nt. Strand specific signals for changes during the growth curve, reference RNA signal and probability of being transcribed are indicated as described in Figure 2 (main text).

Supplementary Figure 3: Intergenic regions in *H. salinarum NRC-1* predicted operons. Horizontal axis shows the lengths of intergenic region in nucleotides, vertical axis shows the frequency for (A) all predicted operons in *H. salinarum NRC-1*, (B) operons classified as conditionally modulated, as described in Experimental Procedures, (C) operons that were not classified as conditionally modulated. Lower panels show the length of intergenic regions (horizontal axis) vs. (D) the minimum tiling score, defined by the difference in the tiling probe levels for genes surrounding the gap and (E) the minimum correlation score along all 719 environmental conditions for each predicted operon. Red dots indicate intergenic regions shorter than 20 nt and green circles represent intergenic regions that were classified as condition-dependent, as described in Experimental Procedures.

Supplementary Figure 4: Correlated transcript and promoter fused GFP dynamics induced by copper (Cu) efflux gene regulation. The gene *yvgX* codes for a Cu specific efflux pump. mRNA expression data for *yvgX* from microarray experiments shows a characteristic pulse response upon system insult with Cu at $t = 0$. A fusion of the *yvgX* promoter to a rapidly decaying GFP variant (smrGFP-ssr1) shows a similar expression profile with an understandable time delay due to protein translation and fluorophore maturation.

Supplementary Figure 5: Validation of ChIP-chip intensities between two microarray platforms. A comparison of the peak intensities (ratios, relative to WCE sample) derived from *MeDiChI* for all three TFs (TFBd, TFBf and TFBa) for which there were biological replicate measurements using the two different microarray platforms (500 nt resolution spotted arrays [*y*-axis] vs. 13 nt resolution Nimblegen arrays [*x*-axis]). Also displayed are the corresponding R^2 between the samples, and the intensities corresponding to *LFDR* = 0.1 (the threshold used to select TFBSs in the manuscript).

Supplementary Tables

All Supplementary Tables are provided as .xls and .pdf files.
In the Excel files (.xls), legends and sub-tables are presented in different worksheets.

Supplementary Table 1: (A) Significant Transcription factor binding sites (TFBS, LFDR < 0.1) and (B) Multi transcription factor binding loci

Supplementary Table 2: Transcription start sites, termination sites and untranslated 5' and 3' regions for annotated genes and operons that were expressed in *H. salinarum NRC-1* during growth and/or reference conditions.

Supplementary Table 3: Overlapping transcripts in *H. salinarum NRC-1*. Overlaps greater than 20 nt were considered as significant, given the error of transcript boundary detection using tiling arrays.

Supplementary Table 4: Revision of gene start codons based on detected TSS, Peptide Atlas information and *H. salinarum R-1* annotation.

Supplementary Table 5: Newly transcribed elements in *H. salinarum NRC-1* genome. Genome location, estimated length of the transcript and transcript levels (log ratio) during growth.

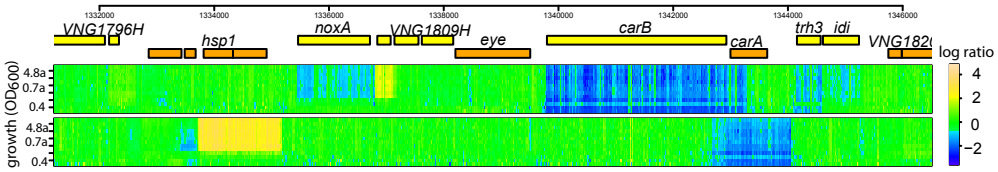
Supplementary Table 6: Conditional operons in *H. salinarum NRC-1*. Minimum correlation, tiling score values and combined score are reported. Manually verified conditional operons are also indicated.

Supplementary Table 7: Transcription factor binding sites (TFBS) internal to genes that are associated with a transcript boundary.

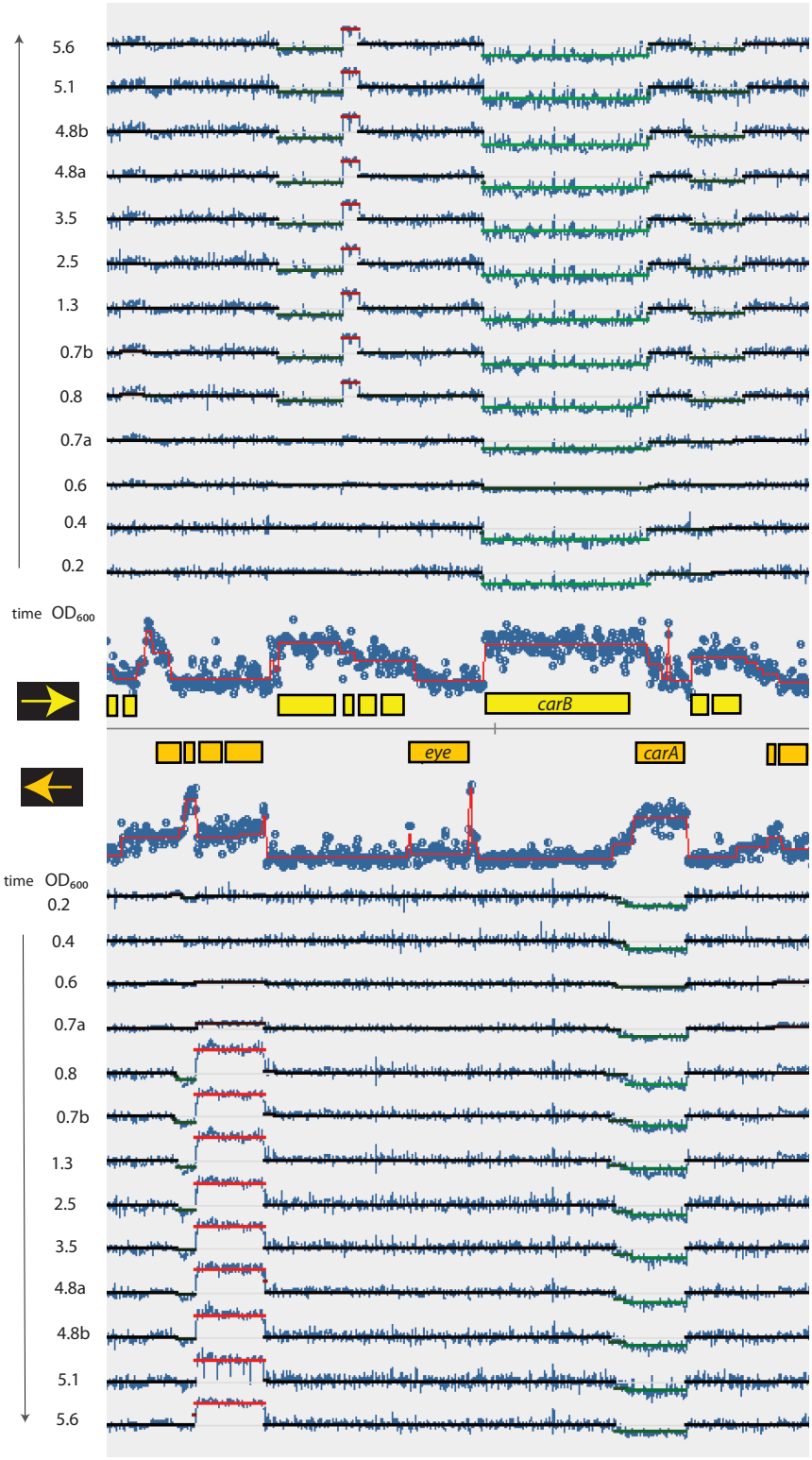
Brenneis, M., Hering, O., Lange, C., and Soppa, J. (2007). Experimental characterization of Cis-acting elements important for translation and transcription in halophilic archaea. PLoS Genet *3*, e229.

Sup Figure 1

A

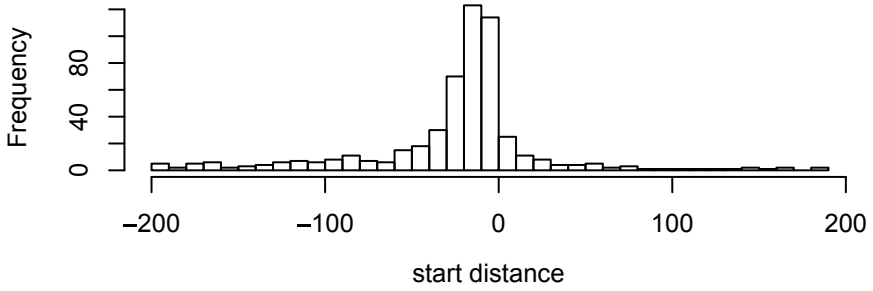


B

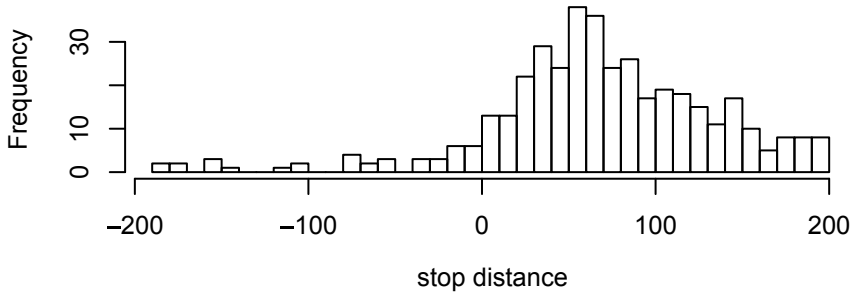


Sup Figure 2

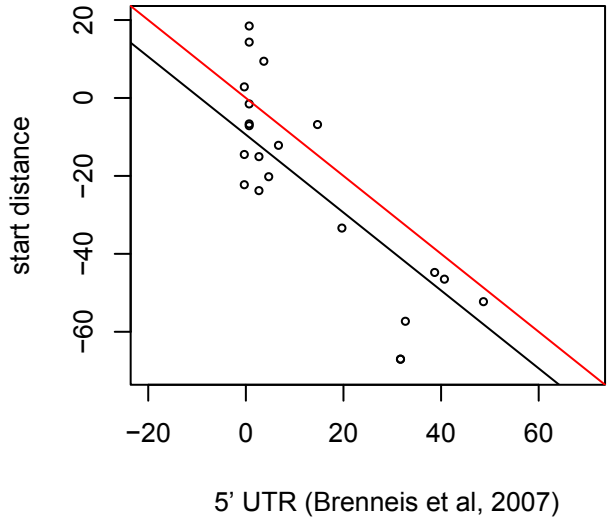
A



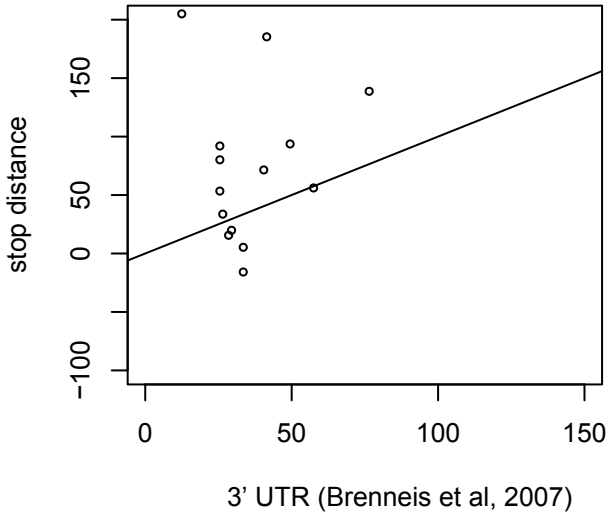
B

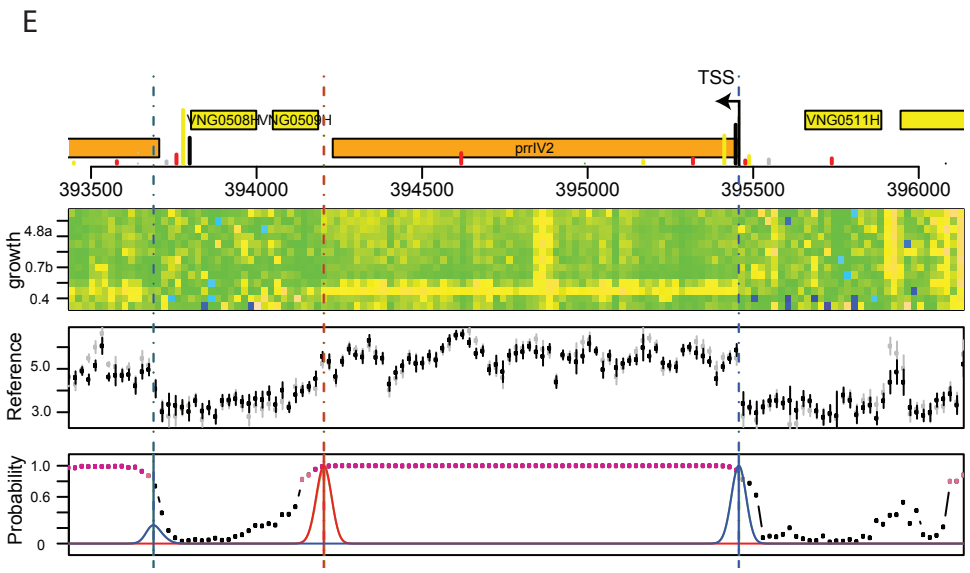


C



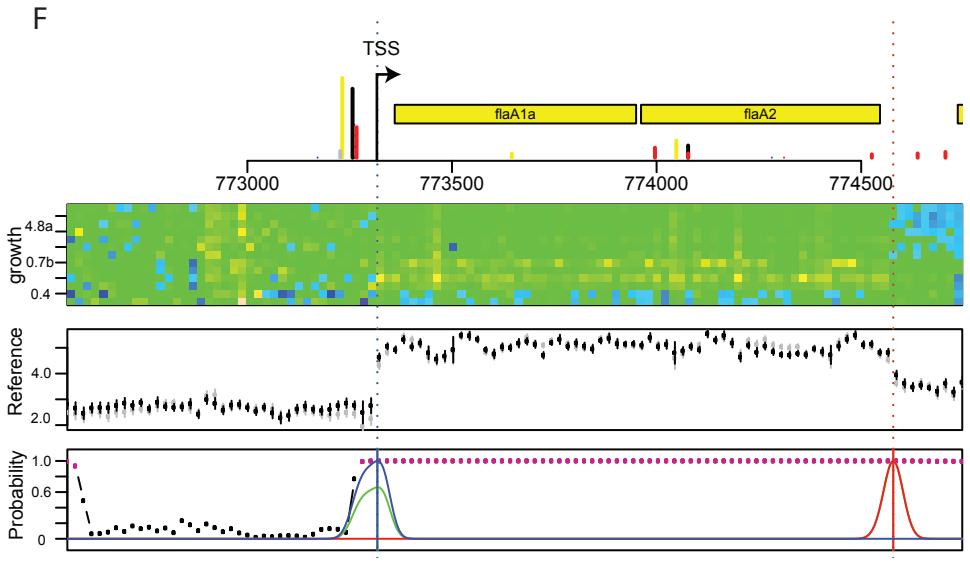
D





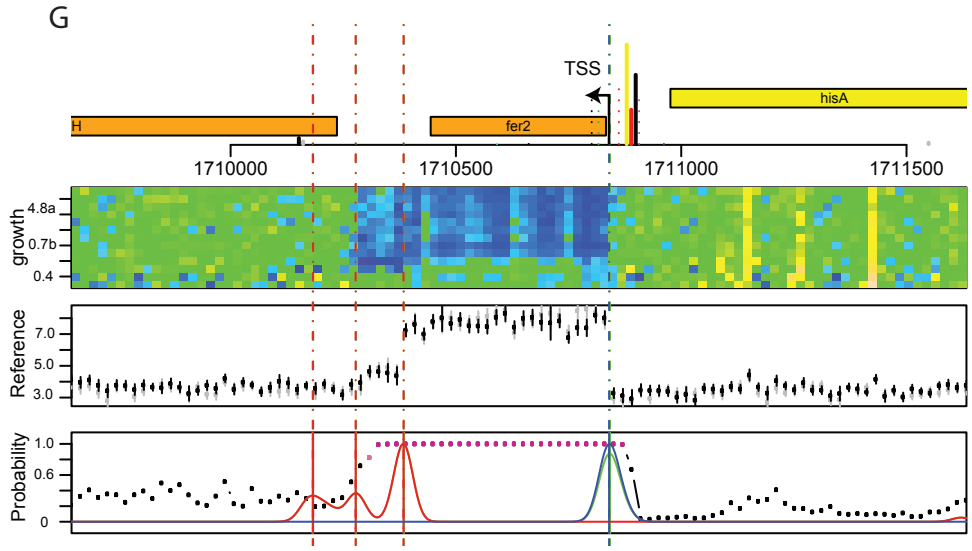
VNG0510G: proteasome-activating nucleotidase

Predicted 5' UTR length (error of 20 nt): 8 nt
Experimentally determined 5'UTR length: 1 nt



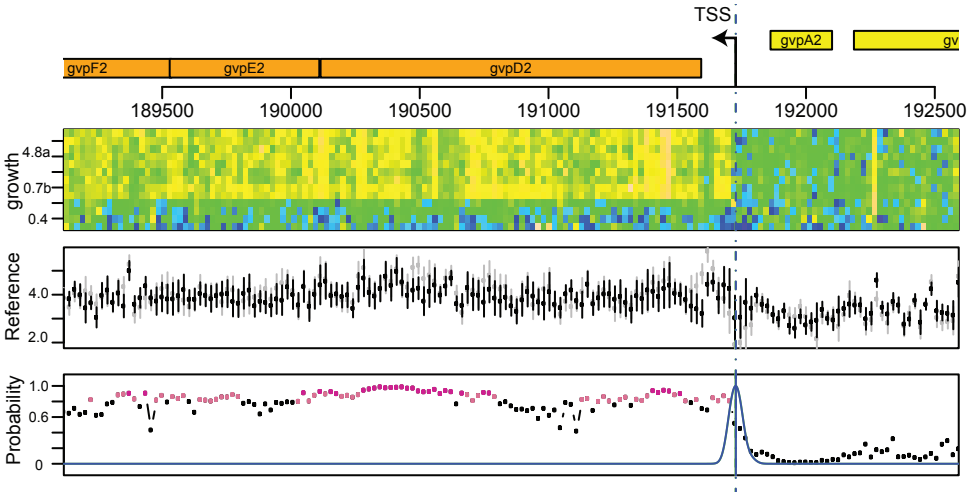
VNG1008G: flagellin A1 precursor, flaA1a

Predicted 5' UTR length (error of 20 nt): 39 nt
Experimentally determined 5'UTR length: 39 nt



VNG2293G: ferredoxin (2Fe-2S), fdx

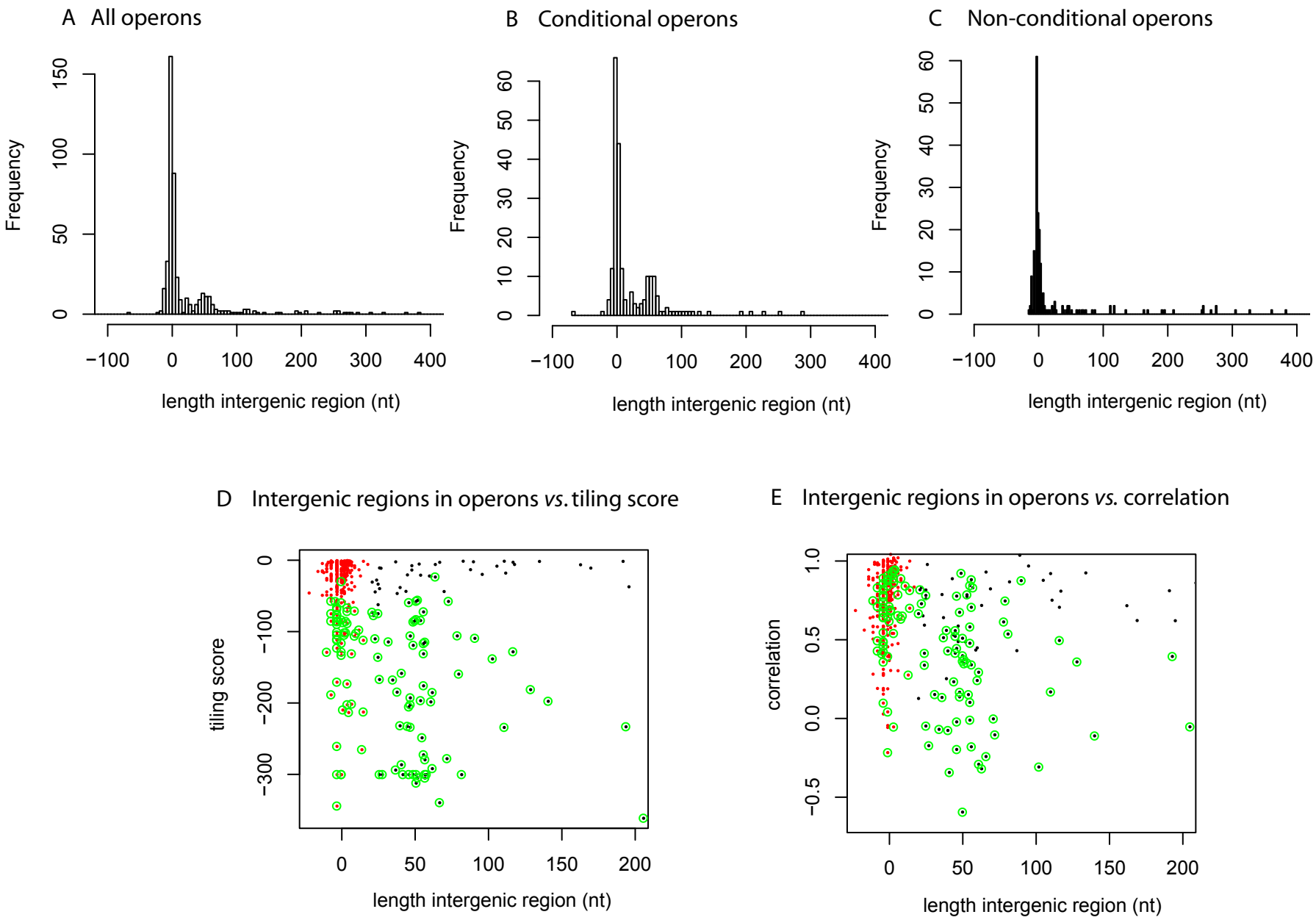
Predicted 5' UTR length (error of 20 nt): 2 nt
Experimentally determined 5'UTR length: 1 nt

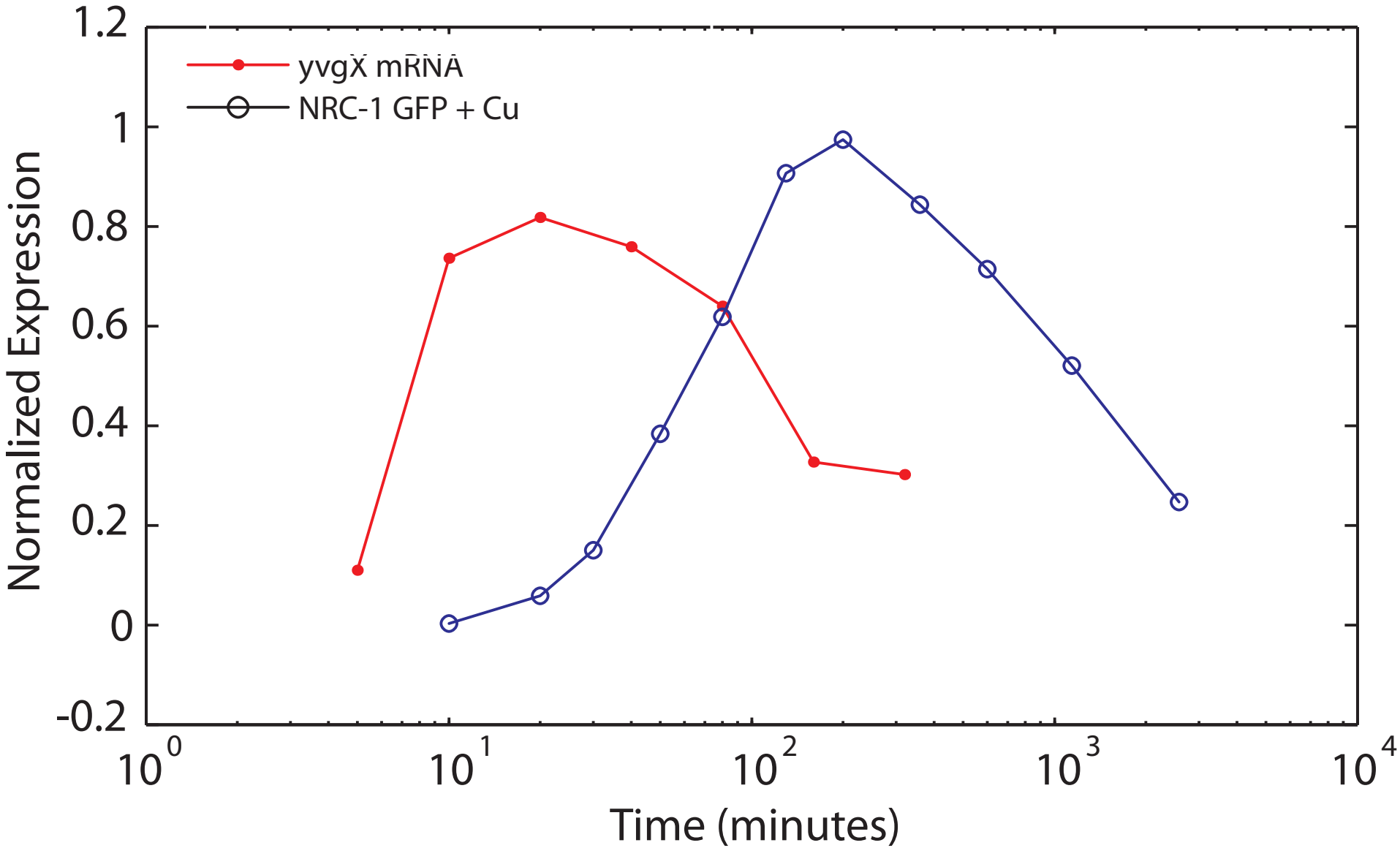


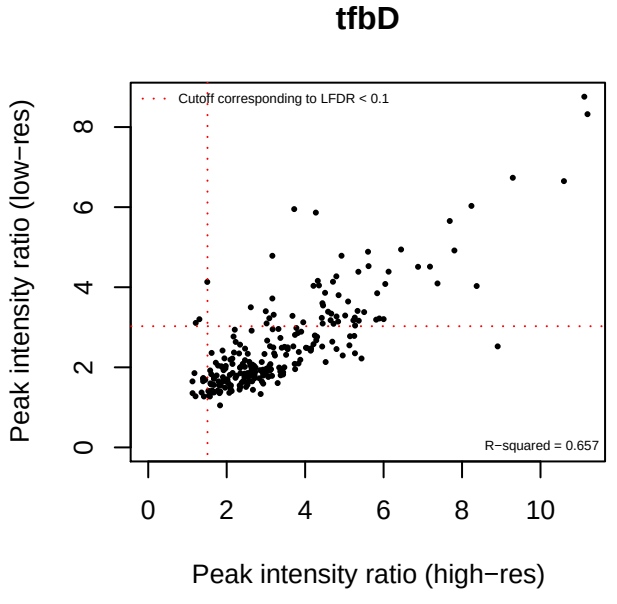
VNG6240G: gas vesicle protein gvpD2

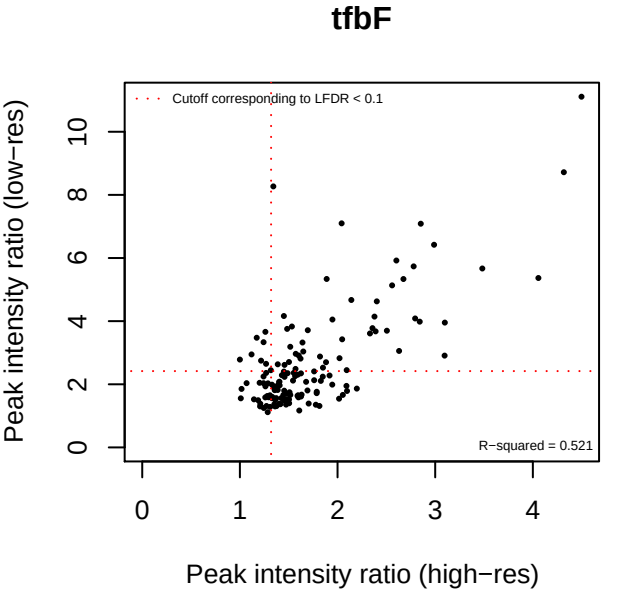
Predicted 5' UTR length (error of 20 nt): 117 nt
Experimentally determined 5'UTR length: 116 nt

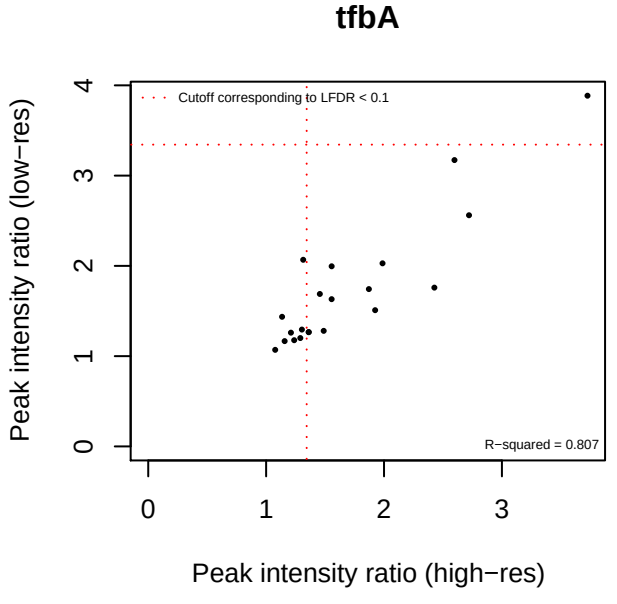
Sup Figure 3


A All operons
Frequency
0
50
100
150
−100
0
100
200
300
400
length intergenic region (nt)
B Conditional operons
Frequency
0
10
20
30
40
50
60
−100
0
100
200
300
400
length intergenic region (nt)
C Non-conditional operons
Frequency
0
10
20
30
40
50
60
−100
0
100
200
300
400
length intergenic region (nt)
D Intergenic regions in operons vs. tiling score
tiling score
0
−100
−200
−300
0
50
100
150
200
length intergenic region (nt)
E Intergenic regions in operons vs. correlation
correlation
−0.5
0.0
0.5
1.0
0
50
100
150
200
length intergenic region (nt)


yvgX mRNA
NRC-1 GFP + Cu
Normalized Expression
Time (minutes)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
10^0
10^1
10^2
10^3
10^4


tfbD
Cutoff corresponding to LFDR < 0.1
R−squared = 0.657
Peak intensity ratio (low−res)
Peak intensity ratio (high−res)
0
2
4
6
8
10


tfbF
Cutoff corresponding to LFDR < 0.1
R−squared = 0.521
Peak intensity ratio (low−res)
Peak intensity ratio (high−res)
0
1
2
3
4
10


tfbA
Cutoff corresponding to LFDR < 0.1
R−squared = 0.807
Peak intensity ratio (low−res)
Peak intensity ratio (high−res)
0
1
2
3
4