

# Supplementary Table 2

## List of 20 most depleted six-mers in the WT coding region

Six-mer	Log2 Fold Change	Zscore in Recoded	Remarks
AATAAT	-1.87	-40	-10 motif of 24 WT constitutive genes, for example <i>yifB</i>
AGAATG	-1.86	-37	Any base (A/T/G/C) upstream to this six-mer creates a -10 motif found in WT constitutive genes
ATAATA	-2.03	-36	Anti-sense -10 motif of 13 WT constitutive genes, for example <i>osmB</i>
ATAATG	-1.90	-40	Anti-sense -10 motif of a WT constitutive gene <i>yfdV</i>
GAGAAT	-1.99	-39	-10 motif of a WT constitutive gene <i>fabI</i>
GACAAT	-1.55	-34	-10 motif of a WT constitutive gene <i>yghW</i>
AGGAGG	-2.80	-36	Ribosome binding site motif
GAGGAG	-2.49	-36	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Catalog">http://parts.igem.org/Ribosome_Binding_Sites/Catalog</a> )
GGAGAC	-2.44	-33	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
GAGGGA	-2.93	-33	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Community_Collection</a> )
GGAGGG	-2.87	-33	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
GGAGGA	-2.33	-35	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Catalog">http://parts.igem.org/Ribosome_Binding_Sites/Catalog</a> )
AAGGAG	-2.38	-34	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Catalog">http://parts.igem.org/Ribosome_Binding_Sites/Catalog</a> )
GAGGAA	-1.81	-34	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Catalog">http://parts.igem.org/Ribosome_Binding_Sites/Catalog</a> )
ATGGAG	-1.51	-33	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
CTTGGA	-3.54	-34	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
TTGGAG	-3.31	-35	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
TTTGGA	-2.79	-33	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
TTGGAA	-2.76	-34	Ribosome binding site motif ( <a href="http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson">http://parts.igem.org/Ribosome_Binding_Sites/Prokaryotic/Constitutive/Anderson</a> )
ATAGAC	-3.61	-33	Part of the DNA binding sequence of the <i>MetJ</i> (represses genes involved in biosynthesis and transport of methionine)

## List of 20 most over-represented six-mers in the WT coding region

Six-mer	Log2 Fold Change	Zscore in Recoded	Remarks
GCTGGA	+1.77	101	Part of E.coli's Chi sequence
GCTGGT	+1.74	89	Part of E.coli's Chi sequence
CTGGTG	+1.99	103	Part of E.coli's Chi sequence
CTGGCG	+2.32	158	Part of E.coli's Chi sequence
CTGCTG	+1.95	98	Part of E.coli's Chi sequence
TGCTGG	+1.99	124	Part of E.coli's Chi sequence
GTGCTG	+1.97	92	Part of E.coli's Chi sequence
CGCTGG	+2.34	156	Part of E.coli's Chi sequence
GCTGGC	+1.94	119	Part of E.coli's Chi sequence
GCGCTG	+2.21	130	Part of E.coli's Chi sequence
CCGCTG	+1.81	83	Part of E.coli's Chi sequence
TGGCGC	+1.66	82	Part of E.coli's Chi sequence
TGATTG	+3.36	162	Part of consensus sequence of <i>murR</i> (DNA binding protein repressing genes involved in catabolism of cell wall sugars)
GTGATT	+2.93	110	Part of consensus sequence of <i>murR</i> (DNA binding protein repressing genes involved in catabolism of cell wall sugars)
CTGATT	+2.75	117	Part of consensus sequence of <i>murR</i> (DNA binding protein repressing genes involved in catabolism of cell wall sugars)
TGATTA	+2.44	94	Part of consensus sequence of <i>murR</i> (DNA binding protein repressing genes involved in catabolism of cell wall sugars)
GATTGC	+2.33	87	Part of consensus sequence of <i>murR</i> (DNA binding protein repressing genes involved in catabolism of cell wall sugars)
GCGATT	+2.13	81	Part of consensus sequence of <i>Cra</i> DNA binding protein (catabolic repression)
ATTGAT	+2.44	85	-----
ATTGCC	+2.46	106	-----

### Supplementary Table 2:

#### Counting all six-mers in the coding region of *E. coli* shows depletion of promoter motifs and ribosome-binding sites

We counted occurrences for each possible six-mer in all *E. coli* genes. In addition, we did the same counting but for the 1000 alternative versions we recoded (by encoding the same amino acid sequence and while preserving the overall codon bias). Next, we compared the Z-score for each six-mer representing the rank of the WT count to the 1000 recoded genomes. Shown in the table are the 40 six-mers with the most extreme Z-scores, 20 from each side (20 most depleted six-mers in the upper part of the table, and 20 most over-represented six-mers in the lower part of the table). Overall, promoter motifs (specifically 'minus-10') and ribosome binding site motifs are depleted in the coding region of *E. coli*. On the other hand, Chi sequences are over represented in the coding region.