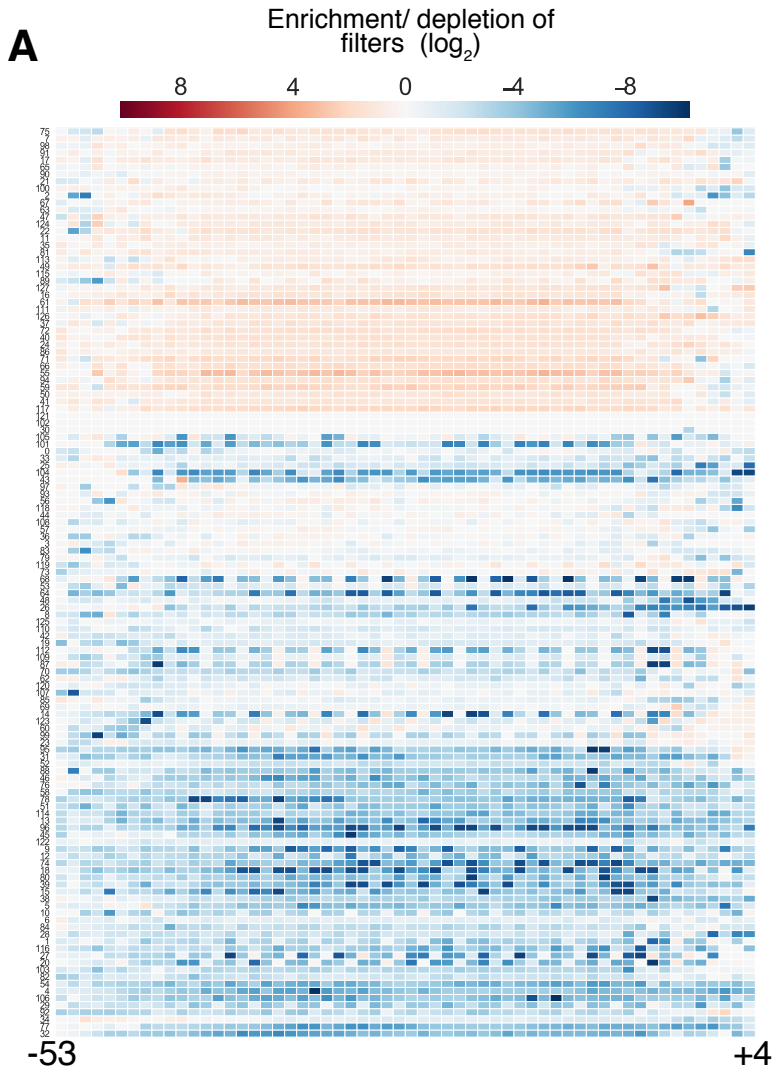
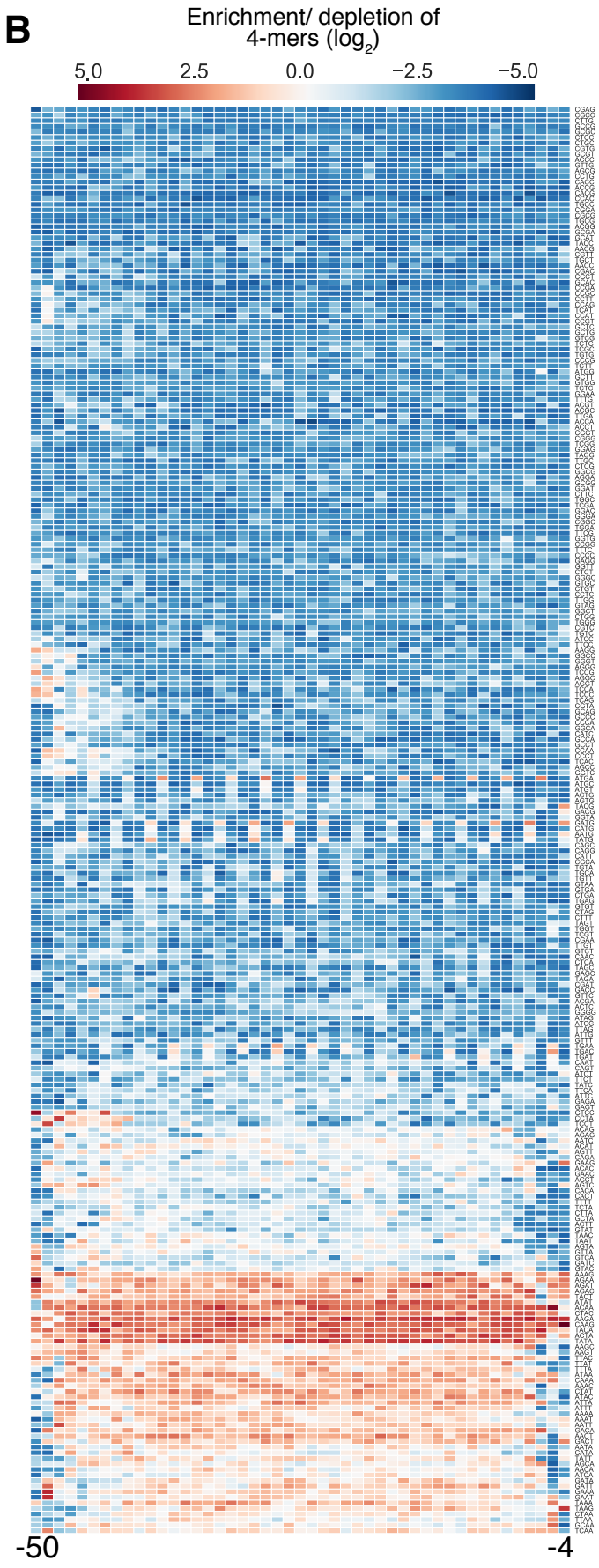

A
Enrichment/ depletion of filters (log2)
8
4
0
−4
−8
-53
+4


B
Enrichment/ depletion of 4-mers (log2)
5.0
2.5
0.0
−2.5
−5.0
-50
-4

Supplemental Figure 10. Position-specific enrichment and depletion of particular motifs and sequences across the 5,000 5’ UTRs optimized via the model-guided in-silico evolutions. In both plots, the x-axis indicates the position across the 5’UTR, while the y-axis indicates the 128 filters in the first convolutional layer of the CNN, or all possible 4-mer sequences. (A) The relative change in the position-specific filter scores between the beginning and end of the evolutions. The 5,000 initial 5’ UTR sequences were compared to the sequences from the point in the evolutions where no additional increase in expression was predicted (a potentially different step for each 5’ UTR being optimized). The average filter score was calculated at every position across each set of 5,000 sequences. The log2 difference of the averages from the last step and the first step are presented. Positive values (in blue) indicate filters that are more highly activated at the indicated 5’ UTR position after the evolutions, while the opposite is true for negative values (red). (B) The enrichment of every possible 4-mer. The sum of the number of times a given 4-mer sequence appeared at each position in the 5’ UTR was calculated for the first and 40th steps of the evolutions. Enrichment was calculated by taking the log2 difference of these two values for each 4-mer sequence at each 5’ UTR position. Positive values indicate 4-mers that increased in number at the indicated positions by the 40th step, while negative values indicate the opposite.