Supporting Text 1

Robustness Analysis of Shuffled MicroRNA Sequences:

An alternative approach to verifying that the increased neutrality observed in miRNAs is not a byproduct of a bias in the base composition of the native stem loops (compared with the sequences generated by the inverse folding algorithm) employs shuffled miRNA sequences. Each miRNA precursor sequence in the analysis is replaced with a shuffled version. Shuffling is performed according to the dinucleotide shuffle algorithm [1], which perfectly preserves both mononucleotide and dinucleotide frequencies. The robustness of these shuffled sequences is analyzed in the exact same method used for the non-shuffled miRNA (i.e., each shuffled sequence is folded and assigned a reference set comprising of sequences that fold into the same minimum free energy structure). If robustness was indeed the product of some base composition bias, such shuffled sequence (which maintain any existing bias) should demonstrate the same significant robustness.

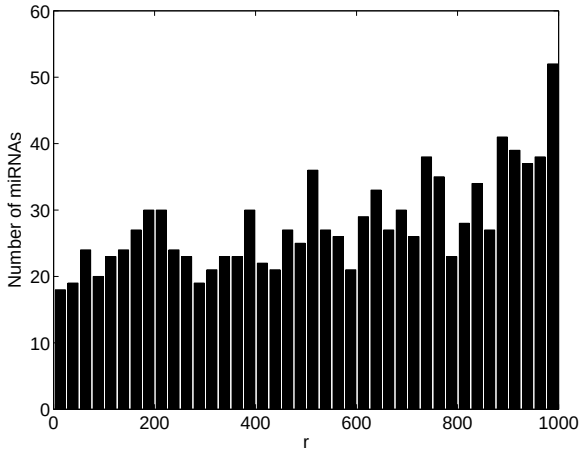

Figure 1: The distribution of the rank, r, for structural robustness of shuffled miRNA sequences. High ranks correspond to significantly robust genes. Only 3.3% of the shuffled miRNAs are significantly robust as opposed to 16.5% of the original miRNA sequences, indicating that the observed robustness of miRNAs cannot be attributed to base composition bias.

Comparing the neutrality of the shuffled miRNA stem loop sequences with the neutrality of inversely folded sequences that form the same structure as the shuffled stem loops demonstrates that the increased neutrality found in the original unshuffled sequences vanishes (Figure 1). This experiment further corroborates that the observed robustness is not the product of miRNAs' base composition

bias.

References

[1] C. Workman, A. Krogh, *Nucleic Acids Res* **27**, 4816 (1999).