DEPARTMENT OF PHYSICS

DEPARTMENT PROFILE: https://physics.osu.edu/people/shatoff.1

The long range effect of SNPs on RNA-protein binding

Elan Shatoff, Ralf Bundschuh

INTRODUCTION

RNA-protein binding is mediated by RNA secondary structure, since many proteins cannot bind to base-paired RNA. Changes in secondary structure can have long range effects, if one binding event changes the secondary structure in such a way that it is easier for another binding event to proceed. In this study we investigate the effect of single nucleotide polymorphisms (SNPs) on RNA-protein binding.

SNPs:

- SNPs are variations of a single nucleotide in a genome that appear in a certain fraction of the population
- Missense SNPs in coding regions are usually associated with disease
- SNPs are also abundant in non-coding regions and the effect of these SNPs is less well understood

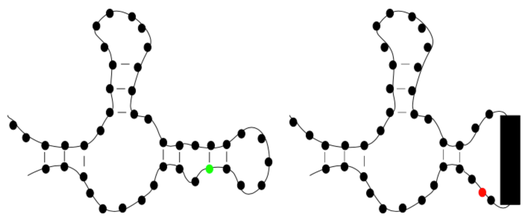
Cooperative RNA-protein binding:

- When proteins bind to RNA they affect the RNA secondary structure
- Secondary structure changes may make it easier or harder for something else to interact with the RNA [1]
- Changes in the sequence of an RNA, such as a SNP, will also change the secondary structure of RNA
- We hypothesize SNPs can effect protein binding even if the SNP is not directly in the protein binding footprint
- We hypothesize that SNPs that negatively effect protein binding through changes in RNA secondary structure may be selected against by evolution

AIM

In this study we aim to investigate the secondary structure mediated effect of single nucleotide polymorphisms on RNA-protein binding, in particular non-coding SNPs.

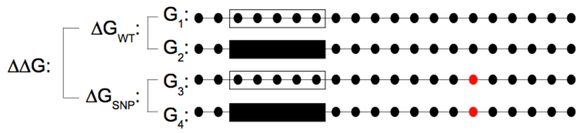
Interactions between SNPs and protein binding



Quantitatively measuring the effect of SNPs on protein binding

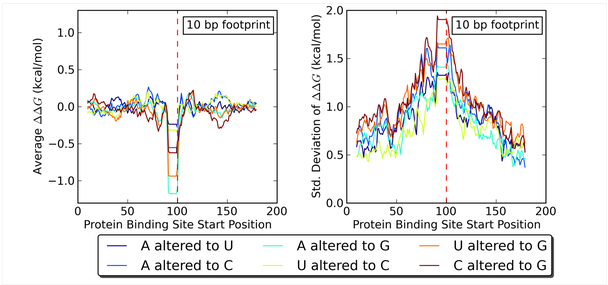
In this project we began by showing that SNPs have an effect on protein binding by measuring the effect on random sequences.

- If a SNP is present the RNA will have different free energies, and different free energy changes for a protein to bind ΔG_{WT} and ΔG_{SNP}
- The effect of SNPs on protein binding is quantified by $\Delta\Delta G$, the difference in the differences between the energy states
- Free energies are derived from the Vienna Package [2], using hard constraints to simulate bound proteins
- If the presence of the SNP is cooperative with the protein binding, then it will be easier (or harder) for the protein to bind when the SNP is present than when it is absent



We show that SNPs can have a significant effect on the binding of proteins to random sequences, with a standard deviation in $\Delta\Delta G$ on the order of 1 kcal/mol even tens of nucleotides away from the protein footprint.

Effect of SNPs on proteins binding random sequences

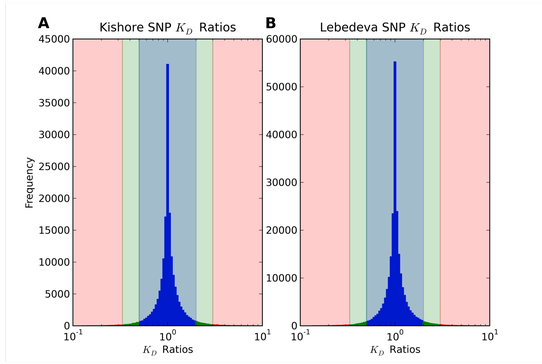


Effect of SNPs on HuR binding

We measure the effect of actual SNPs on actual protein binding sites in humans.

- We map known SNP locations and PAR-CLIP data for HuR (ELAVL1) binding sites to RNA transcripts [3][4]
- Using a modified version of the Vienna Package [5], we are able to determine the K_d of proteins binding to real RNA sequences
- We are able to measure the change in K_d at HuR binding sites with and without a nearby SNP, and take a ratio of the two to see the strength of the effect

Effect of SNPs on Kd of HuR

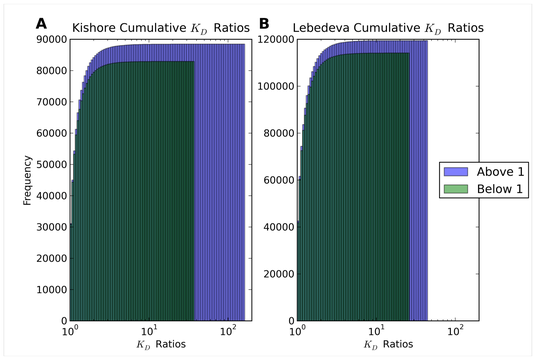


- We find that SNPs can have a many fold effect on the K_d of HuR to binding sites on specific RNA transcripts
- SNPs are visibly depreciated on HuR binding motifs
- We find that the distance distributions for SNPs with different effect strengths are very similar

SNPs may be under selection

- We have shown that SNPs have the ability to affect RNA-protein interactions, but we want to know if this effect has functional relevance
- In the null hypothesis where SNPs do not play a functional role, we would expect SNPs raise and lower protein affinities at equal rates, but we in fact see a significant asymmetry ($p<10E-28$)
- We find that SNPs that make binding easier are over-represented, and have larger effect sizes, indicating that SNPs that make binding harder may be under negative selection

Cumulative distributions of SNP Kds



CONCLUSIONS

We have shown that SNPs can act at a distance, through RNA secondary structure, on RNA-protein binding interactions. Due to this effect, SNPs can have a many-fold effect on the K_d of HuR binding to RNA transcripts.

FUTURE PLANS

In the future we plan to further asses the biological relevance of this effect both on a genome wide level and at HuR sites that are most affected.

BIBLIOGRAPHY

1 Yi-Hsuan Lin, Ralf Bundschuh; RNA structure generates natural cooperativity between single-stranded RNA binding proteins targeting 5′ and 3′UTRs, *Nucleic Acids Research*, Volume 43, Issue 2, 30 January 2015, Pages 1160–1169, https://doi.org/10.1093/nar/gku1320

2 Hofacker I.L., Fontana W., Stadler P.F., Bonhoeffer L.S., Tacker M., Schuster P.. Fast folding and comparison of RNA secondary structures, Monatsh. Chem. , 1994, vol. 125 (pg. 167-188)

3 Kishore S, et al. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat. Methods. 2011;8:559–564. doi: 10.1038/nmeth.1608.

4 S. Lebedeva, M. Jens, K. Theil, B. Schwanhausser, M. Selbach, M. Landthaler, N. Rajewsky Transcriptome-wide analysis of regulatory interactions of the RNA-binding protein HuR Mol. Cell, 43 (2011), pp. 340-352

5 Robert A. Forties, Ralf Bundschuh; Modeling the interplay of single-stranded binding proteins and nucleic acid secondary structure, *Bioinformatics*, Volume 26, Issue 1, 1 January 2010, Pages 61-67 https://doi.org/10.1093/bioinformatics/btp627

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. DMR-1719316.

THE OHIO STATE UNIVERSITY

shatoff.1@osu.edu