

Statistical methods for miRNA motif finding

Scott Norton

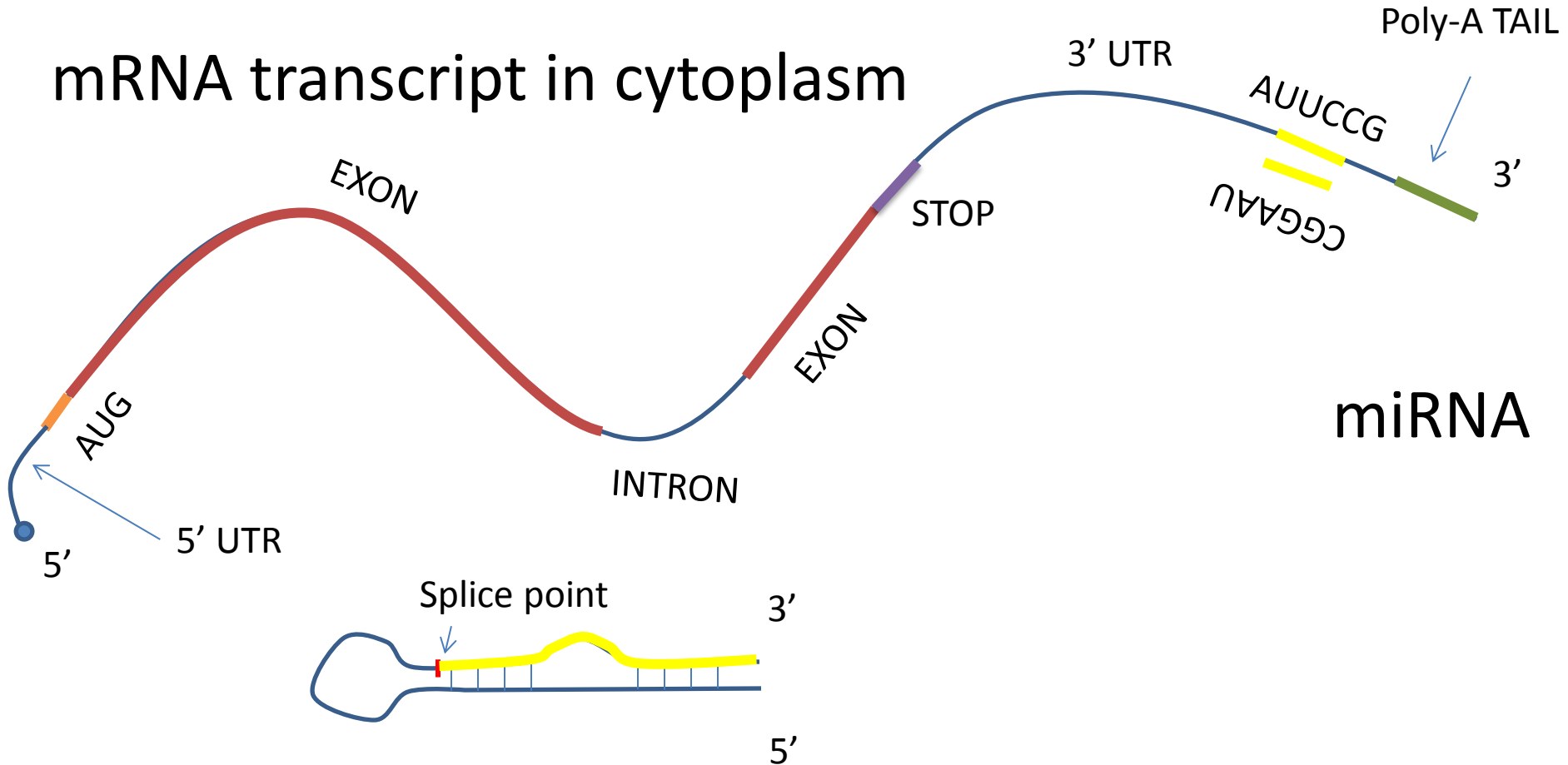
Mentors: Dr. Kevin Chen, Liyang Diao

miRNA

- First discovered in 1993 by Dr. Ambros in *Caenorhabditis elegans*
- Class of small non-coding RNA
- ~22 nt (nucleotides) in length
- Regulates gene expression by binding to mRNA transcript inside cytoplasm
- Acts on 3' untranslated region (3' UTR) of target mRNA
- Interested in motif (recognition sequence) (~6-8 nt)

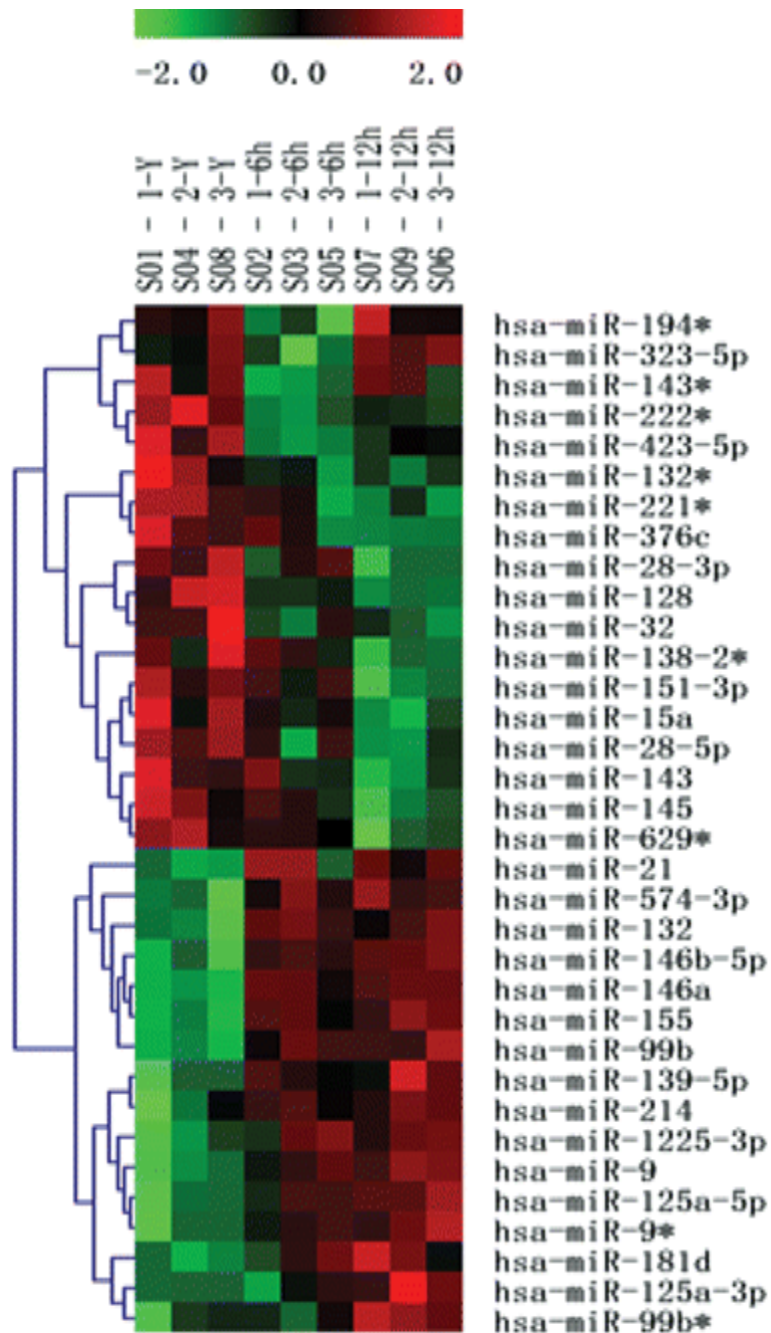
miRNA

mRNA transcript in cytoplasm



Existing method: miREDUCE

- Standard method for identifying miRNA motifs
- Compares gene expression microarrays looking for decreased output (inhibition)
- Searches for miRNA motifs correlated to observed expression ratio and corresponding to specific mature miRNA (miR)



Dysregulation of miRNA expression in monocytes after oxLDL stimulation. Differentially expressed miRNAs ($P < 0.05$) were analysed by hierarchical clustering of the log 2 value of each miRNA microarray signal at 0, 6, and 12 h. Red: up-regulation; green: down-regulation; black: no change. The legend on the right displays the microRNA represented in the corresponding row. The bar code on the top represents the colour scale of the log 2 values. Each column represents the data from a given time course and includes three repetitions at the top of the heatmap.

Image source:

<http://cardiovascres.oxfordjournals.org/content/early/2009/05/17/cvr.cvp121/F1.medium.gif>

Objective

- Create a new statistical method for discovering new miRNA motifs from expression arrays, and mapping them to their target genes
- Desired qualities compared to existing methods such as miREDUCE
 - Replicate established results
 - Faster run time
 - Accurately identifies new motifs that the previous method could not

Basic idea

- Optimize fit between expression log ratio and miRNA motif association strength
 - Standard linear
 - LASSO
 - Ridge
 - Elastic net
 - Other models

References

- Lee RC, Feinbaum RL, Ambros V (December 1993). "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*". *Cell* **75** (5): 843–54.
- Bussemaker, H. J., Li, H., & Siggia, E. D. (2001). Regulatory element detection using correlation with expression. *Nature*, 167-171.
- Ben-Moshe, N. B., Avraham, R., Kedmi, M., Zeisel, A., Yitzhaky, A., Yarden, Y., & Domany, E. (2012, August 29). *Context Specific MicroRNA Analysis (CoSMic): Identification of Functional MicroRNAs and their mRNA targets - Supplementary Data*. Retrieved from Nucleic Acids Research: Oxford Journals: <http://nar.oxfordjournals.org/content/suppl/2012/08/29/gks841.DC1/nar-00762-n-2012-File009.pdf>
- Chen, Ting, Huang, Zhouqing, Wang, Liansheng, Wang, Yue, Wu, Feizhen, Meng, Shu, & Wang, Chanqian. "MicroRNA-125a-5p partly regulates the inflammatory response, lipid uptake, and ORP9 expression in oxLDL-stimulated monocyte/macrophages". *Cardiovascular Research*, Oxford Journals. 12 June 2009, Vol 83.1 pg 131-139, at <http://cardiovascres.oxfordjournals.org/content/83/1/131>