Temperature dependant RNA secondary structure regulates gene expression in *Trypanosoma brucei*.

T. S. Andrews, A. Steif, S. Christiano, J. W. J. Anderson, J. Hein, R. Lyngsø and S. Kelly‡
Department of Statistics, University of Oxford, 1 South Parks Road, OX1 3TG, United Kingdom
‡ steven.kelly@plants.ox.ac.uk

Motivation

The regulatory potential of RNA secondary structure has been recognized since the 1980s. RNA secondary structure is highly sensitive to temperature. *Trypanosoma brucei*, like many organisms, experiences temperature change through its life-cycle. However, the potential for organisms to utilize RNA secondary structure change in response to temperature change to mediate life-cycle gene expression change has not been explored.

More specific background

Trypanosoma brucei, the species responsible for human African sleeping sickness, is spread by the bite of the tsetse fly, causing over 300,000 deaths per year (WHO 1998). Large-scale gene expression changes occur in response to the transition from the mammalian bloodstream (BF), constant temperature of 37^{o}C, to the insect midgut (procyclic form, PF), temperature approximately 22-28^{o}C. Unlike most eukaryotes, trypanosomatids transcription proceeds through polycistronic units containing dozens of genes, restricting gene regulation at the level of transcription. RNA secondary structure, particularly in the 5′ and 3′ untraslated regions (UTRs), may affect protein binding, the susceptibility of the molecule to degradation by exonuleases, and the efficacy of ribosome binding. By examining the predicted structure of *T. brucei* UTRs at representative temperatures of its two hosts, we determined if temperature dependent RNA secondary structure could be responsible to regulating gene expression on a genome-wide scale.

References

[1] Kolev, N. G., Franklin, J. B., Carmi, S., Shi, H., Michaeli, S. and Tschudi, C. (2010). The transcriptome of the human pathogen Trypanosoma brucei at single-nucleotide resolution. *PLoS Pathog* **6**(9), e1001090.

[2] Kozak, M. (1986). Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. *PNAS* **83**(9), 2850-2854.

[3] Siegel, T. N., Hekstra, D. R., Wang, X., Dewell, S. and Cross, G. A. M. (2010). Genome-wide analysis of mRNA abundance in two life-cycle stages of Trypanosoma brucei and identification of splicing and polyadenylation sites. *Nucleic Acids Res.* **38**(15), 4946-4957.

[4] WHO (1998). Control and surveillance of African Trypanosomiasis, Technical report. *WHO Technical Report Service*.

Results

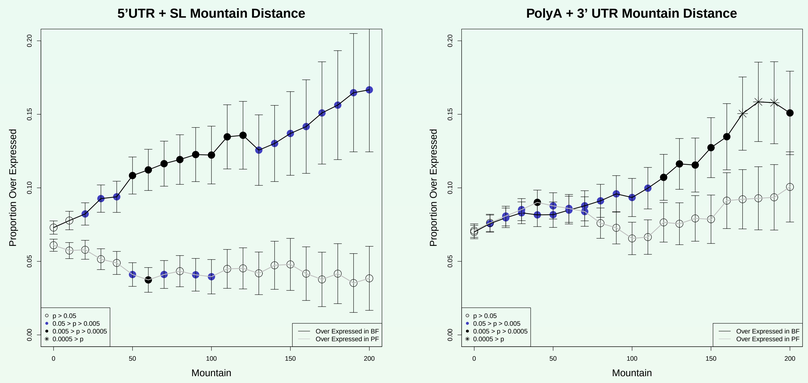

Figure 1: Genes with temperature dependent structural change are up-regulated in bloodstream form trypanosomes. A proportion test was used to test for over-representation of genes which are up-regulated in either blood or insect stage parasites. The test was performed at a range of mountain distance threshold values. The larger the mountain distance, the larger the change in structure between 37^{o}C and 28^{o}C. Error bars represent the standard error of the sample.

- An increase in mountain distance in 5′ UTRs is accompanied by a significant increase in the number of genes exhibiting two-fold or greater up-regulation in blood stage parasites (left).
- Up-regulation of genes in insect stage parasites is not associated with a change in RNA structure.
- An increase in mountain distance in 3′ UTRs is accompanied by a significant (but smaller) increase in the number of genes exhibiting two-fold or greater up-regulation in blood stage parasites (right).

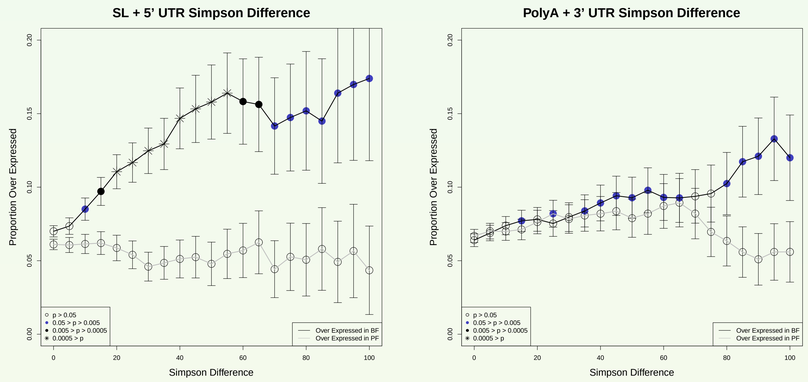

Figure 2: Genes with temperature dependent structural stability are up-regulated in bloodstream form trypanosomes. Methods and symbols are the same as Figure 1. Positive Simpson difference indicates a decrease in stability at 37^{o}C compared to 28^{o}C.

- An increase in Simpson difference in 5′ UTRs is accompanied by a highly significant increase in the number of genes exhibiting two-fold or greater up-regulation in blood stage parasites (left).
- Up-regulation of genes in insect stage parasites is not associated with a change in RNA stability.
- An increase in Simpson difference in 3′ UTRs is accompanied by a low significance increase (much smaller) in the number of genes exhibiting two-fold or greater up-regulation in blood stage parasites (right).

Conclusion

Overall, our results show structural instability of 5′ UTRs at 37^{o}C is responsible for up-regulation of genes in BF trypanosomes. Instability at the 5′ end of mRNA has been suggested to affect translation efficiency in other organisms (Kozak 1986). Our data suggests the length and sequence of the 5′ UTR of BF proteins have been selected to exploit this phenomenon for regulation. Instability of the 5′ UTR could increase flexibility of the mRNA and enhance recognition and binding of the ribosome. Ribosome binding increases translation of the mRNA but may also protect the mRNA from degradation. Structural change in 3′ UTR is also involved in expression regulation, perhaps by the gain/loss of particular structures influencing mRNA degredation by the exosome.

Data & Methods

Sequence Data: 6,546 5′ UTRs and 4,729 3′ UTRs, generated by Kolev et al. (2010) using RNASeq, were extracted from the *T. brucei* genome available from TriTrypDB.
Expression Data: mRNA levels for 11,412 genes in BF and PF were obtained from Siegel et al. (2010), of which 5,199 had known 5′ UTR sequences and 3,814 had known 3′ UTR sequences.
RNA Structure Prediction was computed using the UNAFold package using default parameters at 37^{o}C and 28^{o}C. Boltzmann probabilities were computed using the partition function in UNAFold.
RNA Structural Comparision employed ten distinct measures: number of hairpins, number of stems, fractions of nucleotides involved in base pairing, number of suboptimal structures, Simpson difference, free energy, mountain distance, tree edit distance, symetric set distance, and Hausdorff distance.

Acknowledgements

This work was carried out as part of the Oxford Summer School in Computational Biology, 2011, in conjunction with the Department of Plant Sciences, and with support from the Department of Zoology. Funding was provided by EPSRC, BBSRC, the EU grant COGANGS, the Departments of Statistics and Plant Sciences of Oxford Univesity, and Oxford Centre for Integrative Systems Biology. We thank Dr Steven Kelly for providing computational resources.