

Quantitative Modeling of Transcription

STREAM : A Practical Workbench

Denis C. Bauer
d.bauer@imb.uq.edu.au

Timothy L. Bailey
t.bailey@imb.uq.edu.au

Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld. 4072 Australia

Understanding the transcriptional regulation of a gene in detail is a crucial step towards uncovering and ultimately utilizing the regulatory grammar of the genome. Here, we present **STREAM** [1], the first publicly available framework for modeling, visualizing, and predicting the regulation of the transcription rate of a target gene. Given the concentrations of a set of transcription factors, the transcription factor binding sites in a regulatory

DNA region, and the transcription rate of the target gene, **STREAM** will optimize its parameters to generate a model that best fits the input data. This trained model can then be used to (a) validate that the given set of TFs is able to regulate the target gene and (b) to predict the transcription rate under different conditions (e.g. different tissues, knockout/additional TFs or mutated/missing TFBSs).

<http://bioinformatics.org.au/stream>

Top 10 applications of STREAM:



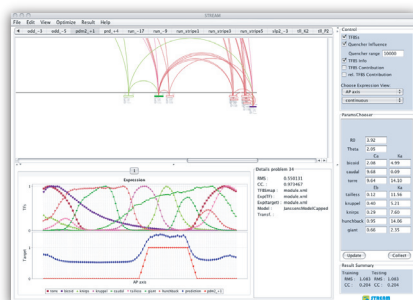
STREAM

STREAM, a Java-implemented framework to calculate and visualize transcriptional regulation using thermodynamic modeling approaches. STREAM offers several optimization methods including gradient descent and simulated annealing.

STREAM allows the user to customize the thermodynamic model by choosing to include additional features, such as co-operativity between TFs or distance dependent repression.

STREAM can optimize one model to the data of multiple CRMs, which increases the confidence in the produced model.

STREAM can be executed using a graphical user interface (GUI) as well as via the command-line.



The Graphical User Interface (GUI) of STREAM.

Analysis / Validation

- Identify the set of regulatory TFs of a target gene
- Identify co-regulated genes
- Identify orthologous genes [2].

Prediction

- Predict the transcriptional output for missing/mutated binding sites to explore the robustness of the system or identify the effect of a SNP.
- Predict the output for different TF protein concentrations, e.g. different tissues, situations or in TF-knockout mutants.

Research

- Identify a particular CRM-position by exhaustively scanning the DNA surrounding the gene for a region that causes STREAM to predict the expected activation pattern [2].
- Cluster groups of genes, which are regulated by the same set of TFs with the same activator-repressor configuration.
- Construct a dynamic regulatory network by identifying the TFs regulating a target gene as the ones for which the STREAM prediction is closest to the observation.
- Construct a library of thermodynamic models trained for a particular target gene or groups of target genes (like Pfam).
- The STREAM framework can be used to examine the properties of other thermodynamic models (e.g. Segal et al. [3]).

[1] Bauer, D. C., Bailey, T. L. STREAM: Static Thermodynamic REgulatory Model of transcription. *Bioinformatics*. 2008 Nov 1;24(21):2544-5

[2] Bauer, D. C., Bailey, T. L. Studying the functional conservation of cis-regulatory modules and their transcriptional output. *BMC Bioinformatics*, 2008, 9, 220

[3] Segal, E., Raveh-Sadka, T., Schroeder, M., Unnerstall, U., Gaul, U. Predicting expression patterns from regulatory sequence in *Drosophila* segmentation. *Nature*, 2008, Jan