Synteny based networks for sRNA homolog prediction

Fkz 031L0106B

Steffen C. Lott, Dominik Rabsch, Jens Georg, Wolfgang R. Hess
Genetics & Experimental Bioinformatics, Faculty of Biology, University of Freiburg

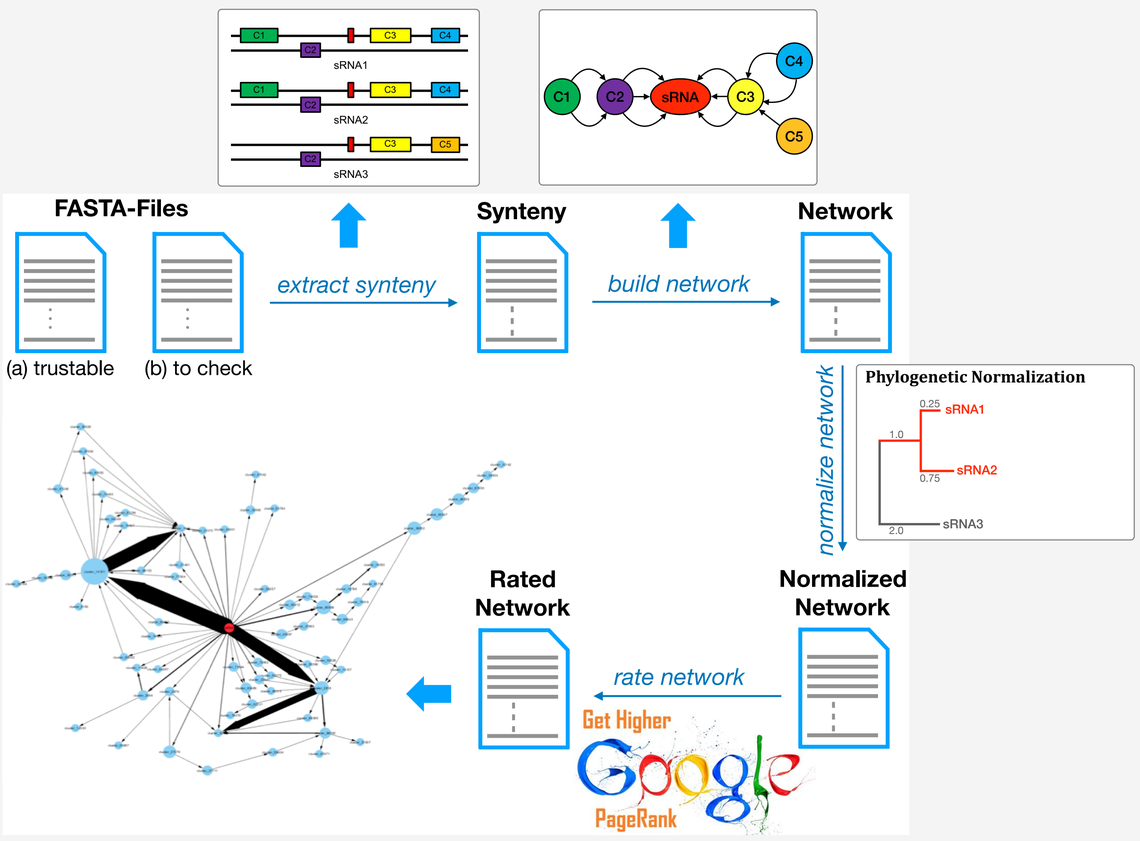
Short description of the project

Within the context of the de.NBI partner project “Structured Analysis and Integration of RNA-Seq experiments (de.STAIR)” we focus on the systemic relevance of regulatory RNAs. One class of regulatory RNAs are the trans-acting small RNAs (sRNAs), post-transcriptional regulators of gene expression. The numbers and types of known sRNAs have rapidly been increasing during recent years due to the extensive application of transcriptomics analyses, also in non-model organisms. The functional and evolutionary characterization of sRNAs requires the identification of homologs. However, finding homologs is frequently challenging due to their heterogeneity, shortness, and partly little sequence conservation. We have established GLASSgo as a standard tool for finding sRNA homologs from a single sequence. Although sequences predicted by GLASSgo are highly reliable, homologs with very low sequence conservation might be missed if the default settings of the workflow are used. To overcome this limitation, we take advantage of analyzing meta-data (synteny) for re-evaluating rejected GLASSgo hits.
Here we present a synteny network build from trustable GLASSgo hits to test sequences below the adjusted threshold of a minimum pairwise similarity of 52%. Therefore, we made use of Google‘s PageRank algorithm to rate the importance of individual homologs/nodes in the network and to assess significance criteria to score the candidate sRNA homologs based on their synteny.

Progress report

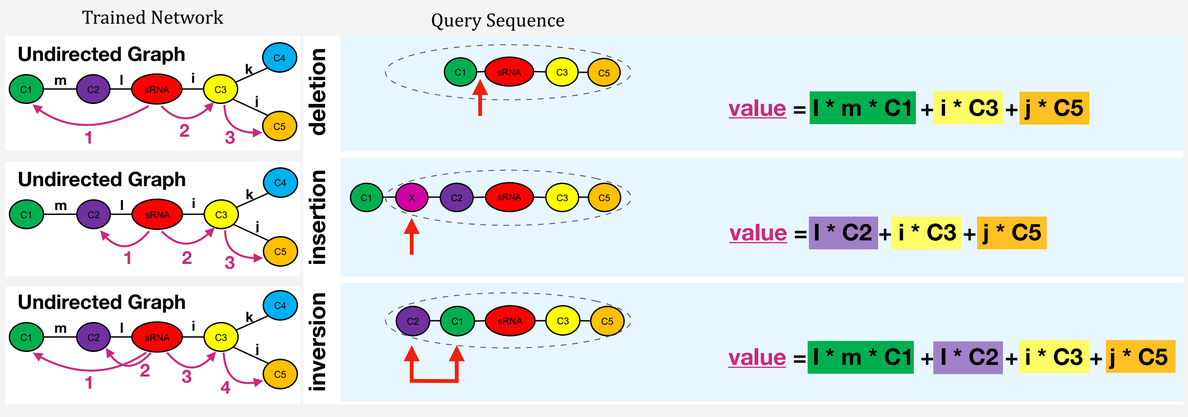
We developed a novel method to analyze meta-data to compute a similarity score called PRS (PageRank based Score). Therefore we utilize all reported GLASSgo hits to build a directed network and rank the nodes with Google's PageRank. Further, we established methods to consider insertions, deletions, and inversions, to calculate a similarity score for all questionable sequences and their synteny.
For evaluating the performance of the newly developed algorithm, we selected 40 sRNA-related RFAM families and constructed for each family a directed network. Afterwards, a single sequence per family was selected and used as input for the GLASSgo algorithm. The unfiltered GLASSgo results were used for extracting the associated synteny and calculating the PRS based on the corresponding pre-computed network.
Additionally, we worked on several improvements regarding GLASSgo, such as the easily accessible web-server, docker integration as well as on the full GALAXY support.

Project – Algorithm description



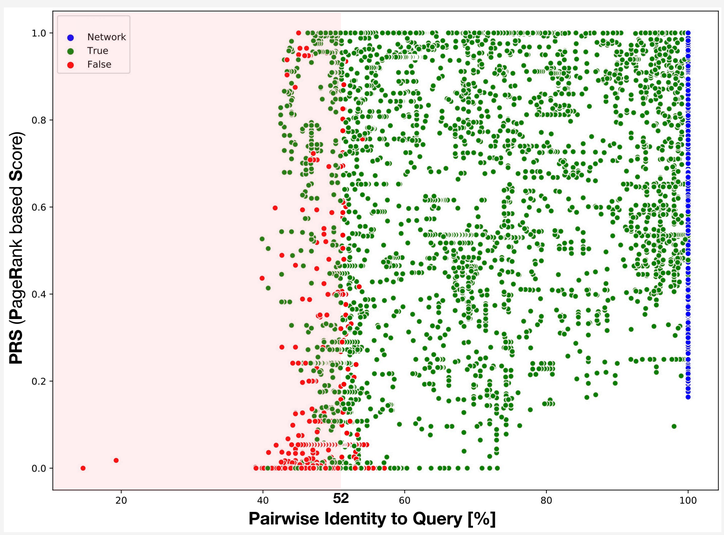
The algorithm takes trustable homologs in multi-FASTA-format as input and extracts for each entry (e.g., sRNA) the surrounding gene neighborhood. Given the obtained synteny, a directed network is built with the sRNA as the center. Phylogenetic tree normalization decreases database biases, such as organism overrepresentation. Finally, PageRank takes the normalized network as input and assigns a value of significance to each node.

Project – Evaluation: PRS (PageRank based Score) calculation



The PageRank based Score (PRS) takes only a predefined number of genes into account (dashed line) and can handle deletions, insertions as well as inversions.

Project – Results evalutated with Infernal



For demonstrating the power of the newly developed PageRank based score, we analyzed unfiltered GLASSgo results coming from 40 different sRNA families. INFERNAL was used to evaluate the unfiltered GLASSgo results, and the color indicates green=True Positives, red=False Positives, as well as blue=Network. The network was built using the corresponding RFAM-families.

de.NBI Training and education

09 - 11/10/19 – University of Freiburg
de.NBI / de.STAIR: Galaxy for linking bisulfite sequencing with RNA sequencing

06 - 09/03/19 – University of Rostock
de.NBI / de.STAIR: Galaxy for linking bisulfite sequencing with RNA sequencing

de.NBI services

Galaxy atoms & workflow generator plugin	github.com/destairdenbi
GLASSgo	10.3389/fgene.2018.00124
- web server	rna.informatik.uni-freiburg.de/GLASSgo/
- docker container	hub.docker.com/r/lotts/glassgo_acc_version
- GALAXY tool	https://toolshed.g2.bx.psu.edu

General information on the project

No. of staff paid from de.NBI grant (FTE) : Steffen C. Lott

Other staff involved: Prof. Wolfgang R. Hess, Jens Georg, Dominik Rabsch

Publications

1. Barshishat S., Elgrably-Weiss M., Edelstein J., Georg J., Govindarajan S., Haviv M., Wright P.R., Hess W.R., Altuvia S. (2018) OxyS small RNA induces cell cycle arrest to allow DNA damage repair. EMBO J. 37, 413-426, DOI: 10.15252/embj.201797651.
2. Georg J, Lalaouna D, Hou S, Lott SC, Caldelari I, Marzi S, et al. The power of cooperation: Experimental and computational approaches in the functional characterization of bacterial sRNAs. Molecular Microbiology. 2019. doi:10.1111/mmi.14420
3. Hagemann M, Gärtner K, Scharnagl M, Bolay P, Lott SC, Fuss J, et al. Identification of the DNA methyltransferases establishing the methylome of the cyanobacterium Synechocystis sp. PCC 6803. DNA Research. 2018. pp. 343–352. doi:10.1093/dnares/dsy006
4. Lott SC, Schäfer RA, Mann M, Backofen R, Hess WR, Voß B, et al. GLASSgo – Automated and Reliable Detection of sRNA Homologs From a Single Input Sequence. Frontiers in Genetics. 2018. doi:10.3389/fgene.2018.00124
5. Lott S.C., Wolfien M., Riege K., Bagnacani A., Wolkenhauer O., Hoffmann S., Hess W.R. (2017) Customized workflow development and data modularization concepts for RNA-Sequencing and metatranscriptome experiments. J. Biotech. 261, 85-96. DOI: 10.1016/j.jbiotec.2017.06.1203.
6. Raden M, Ali SM, Alkhnbashi OS, Busch A, Costa F, Davis JA, …, Steffen C. Lott, … et al. Freiburg RNA tools: a central online resource for RNA-focused research and teaching. Nucleic Acids Research. 2018. pp. W25–W29. doi:10.1093/nar/gky329
7. Weise S.C., Arumugam G., Heidrich S., Villarreal A., Nebel N., Videm P., Dumit V.I., Sananbenesi F., Reimann V., Craske M., Schilling O., Hess W.R., Fischer A., Backofen R., Vogel T. (2018) FOXG1 regulates PRKAR2B transcriptionally and posttranscriptionally via miR200 in the adult hippocampus. Molecular Neurobiology, DOI: 10.1007/s12035-018-1444-7.

GEFÖRDERT VOM Bundesministerium für Bildung und Forschung
Projektträger Jülich Forschungszentrum Jülich
www.denbi.de