

RBC-LE

RNA Bioinformatics Center University of Leipzig

Fkz 031A538B

Jörg Fallmann, Thomas Gatter, Peter F. Stadler
Bioinformatics, University of Leipzig, Germany

Short description of the project

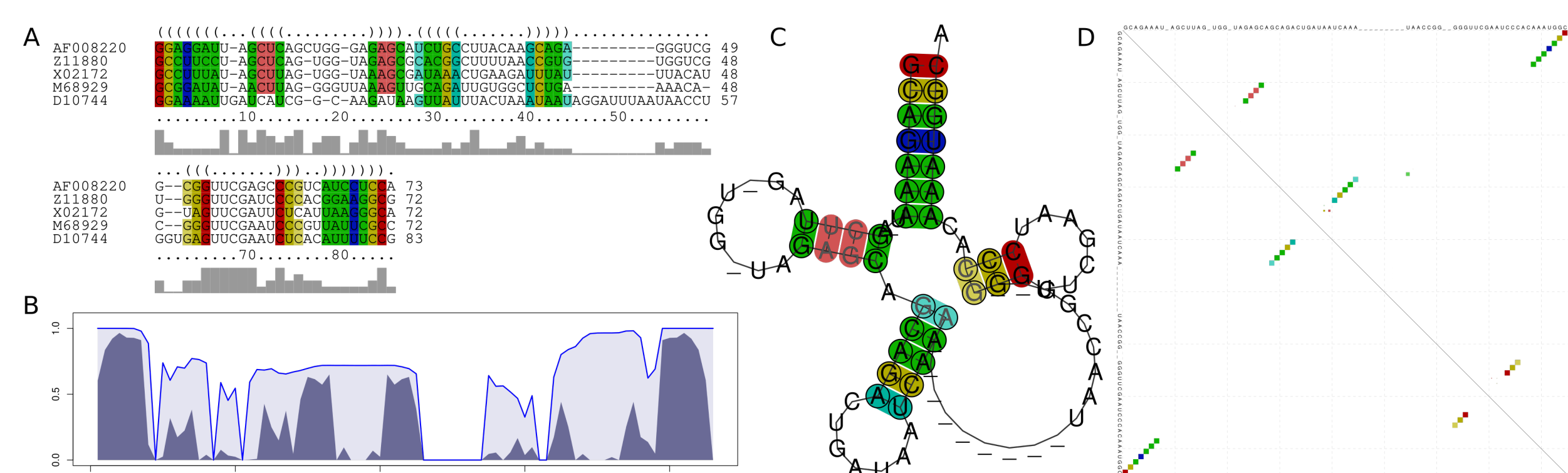
The RNA Bioinformatics Center (RBC) covers all aspects of RNA bioinformatics in the German Network of Bioinformatics Infrastructure (de.NBI), which is establishing a comprehensive bioinformatics infrastructure in Germany.

The Leipzig center focuses on the development and maintenance of tools for RNA (coding and non-coding) structure prediction, conservation analysis, motif detection, target prediction and annotation. These tools are made available to users as stand alone versions as well as an integrated part of the RBC galaxy workbench. Leipzig is offering services and training for the analysis of non-coding RNAs specializing on advanced secondary structure analysis. This also includes RNA-Seq and Galaxy related training and workshops.

de.NBI services

Vienna RNA package – RNA secondary structure prediction and analysis (implementing many folding variants: free energy minimization, partition function folding, local folding, folding of alignments, and subopt folding)

LocARNA package – Multiple comparative analysis of RNAs (incl. LocARNA for RNA alignment and folding, LocARNA-P for predicting reliabilities, SPARSE for even faster alignment, and ExpaRNA-P for matching and folding)



Vienna RNA and LocARNA package: Example results from selected tools. A) multiple structure-based alignment of five tRNAs by LocARNA. B) sequence and structure reliability plot of the alignment (STAR plot) of Subfigure A by LocARNA-P. C) Optimal structure of the alignment by RNAalifold and D) probability dot plot of the alignment by partition function folding of RNAalifold.

RNAz – genome-wide de-novo prediction of non-coding RNAs

CARNA – Alignment and folding of RNA ensembles and pseudoknots

snoSTRIP – homology search and analysis of small nucleolar RNAs

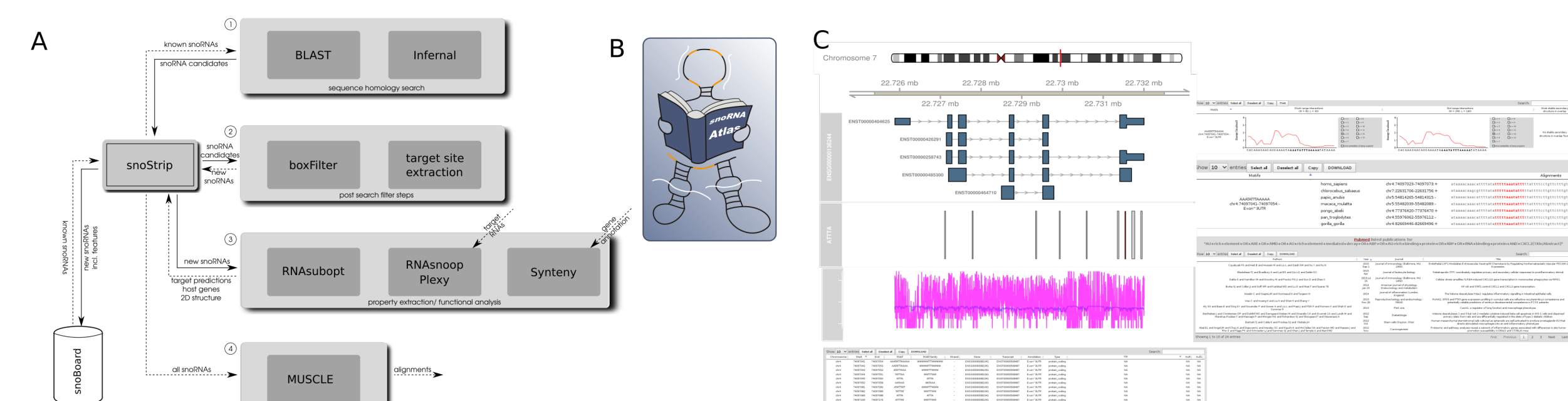
PLEXY – predicting interactions of C/D-box snoRNAs with targets

RNAsoop - efficient target prediction for H/ACA snoRNAs

MITOS – annotation of metazoan mitochondrial genomes

tRNAdb – a comprehensive database of tRNAs

AREsite2 - a database for investigation of AU/GU/U-rich elements



snoStrip, the snoRNA atlas and AREsite2. A) snoStrip pipeline for identification of snoRNAs in arbitrary sets of genomes B) snoStrip Atlas, a database of human snoRNAs C) AREsite2 output for IL6 with annotated AU-rich elements

Blockbuster – detection of read blocks in NGS mapping data

snakes - Highly configurable multi conditional NGS workflows

Ryütō - Transcriptome reconstruction based on network-flow

MIRfix - Automatic curation of miRNA annotation

KPIs: > 50.000 User/Downloads per year
> 600 citations > 200 user requests

Progress report

Based on user demands and developments, new software releases were published and integrated into the Conda package management system for easy integration into the Galaxy framework and to ease installation for end users.

Amongst others, tools of the ViennaRNA package are under constant development and the integration and adaption of new versions into the RBC infrastructure is part of our mission.

Leipzig is also working on a collection of workflows regarding RNA related experiments and analysis of short and long reads, with focus on quality control. An RBC collection of How-Tos and tutorials is available via the Galaxy Training Network.

In tight RBC-wide cooperation we finished work on version 2.0 of the RNA-workbench, a collection of tools, pipelines and workflows. It is an independent instance of Galaxy available via the european Galaxy server (<https://rna.usegalaxy.eu>).

Updated versions of several tools were released or are under development: Galaxy-RNA-workbench 2, MIRfix 2.0, Vienna RNA package 2.4.14, Mitos 2, snakes2.0.1

Packages for Galaxy and Conda were created and updated in an continuous effort. Together with members of the group of Jan Gorodkin (ELIXIR) and the other RBC centers, we continued work on the ELIXIR registry, updated and added tools in various workshops and meetings.

de.NBI Training and education

Leipzig hosted and participated in several bioinformatic training courses. General topics were an introduction into the command line environment, hands-on training with transcriptomic and genomic data sets, as well as the usage of the Galaxy environment. Depending on the needs of participants, specialized topics like predicting RNA secondary structure were covered as well.

List of courses 2019:

- Various hands-on NGS analysis training sessions with small groups
- Training cooperation with industry collaborator ecSEQ:
 - January: NGS Data Analysis: A Practical Introduction
 - February: RNA-Seq Data Analysis Workshop
 - March NGS Data Analysis: Variation Analysis
 - November: DNA Methylation Data Analysis

General information on the project

3 Researchers of the RBC Leipzig are (partially) paid from de.NBI, 1x1FTE, 2x0.5 FTE. Administrative staff and contributions from other lab members are financed from household or project money.

Publications

Automatic curation of large comparative animal MicroRNA datasets. Ali M Yazbeck, Peter F Stadler, Kifah Tout, Jörg Fallmann (2019). Bioinformatics

The RNA workbench 2.0: next generation RNA data analysis. Jörg Fallmann, Pavankumar Videm, Andrea Bagnacani, Bérénice Batut, Maria A Doyle, Tomas Klingstrom, Florian Eggenhofer, Peter F Stadler, Rolf Backofen, Björn Grüning (2019). Nucleic Acids Research

LOTTE-seq (Long hairpin oligonucleotide based tRNA high-throughput sequencing): Specific selection of tRNAs with 3'-CCA end for high-throughput sequencing. Lieselotte Erber, Anne Hoffmann, Jörg Fallmann, Heike Betat, Peter F. Stadler, Mario Mörl (2019). RNA Biology

Ryütō: network-flow based transcriptome reconstruction. Gatter, T., Stadler, P.F. BMC Bioinformatics 20, 190 (2019)