

TEA – TRON Expression Atlas

A Web-based Tool for Visualizing and Interpreting Next Generation Sequencing Expression Data

Marius Byl, Thomas Bukur, Jelle Scholtalbers, Ludmila Schemarow,
Martin Löwer, John C. Castle, Ugur Sahin

*TRON – Translational Oncology at the University Medical Center of the Johannes
Gutenberg University in Mainz gGmbH*

marius.byl@tron-mainz.de; john-castle@tron-mainz.de

The next generation sequencing (NGS) technology has significantly increased the dimensions of data output of biotechnological laboratories. Especially with high throughput sequencing, where the sequencing process has been parallelized, large data amounts are generated in short periods of time. Direct interpretation of the resultant terabytes of data is impossible in such raw data formats like FASTQ. Therefore, the use of various software tools is essential for the analysis of this kind of data. The bioinformatics group of TRON has implemented pipelines to efficiently process RNA-Seq sequencing data and output consistent gene expression values in terms of RPKM unit (*Reads Per Kilobase of exon model per Million mapped reads* [Mortazavi *et al.*, 2008]).

Part of this pipeline is the TRON Expression Atlas (TEA). This web-based tool enables visual processing of expression values of transcripts within a range of samples, allowing, for example, tissue comparisons. The raw data of the Expression Atlas is stored in a database which was developed for this project and connects expression values to additional information including sample type and sample processing.

The TRON Expression Atlas offers both, the generation of interactive diagrams (which can be a very powerful tool to visualize gene expression patterns) and also a filtering function. It is possible to filter all records of the database or only a smaller selection of genes on a basis of specified parameters and then receive a list of genes matching a set of criteria. This can be used to find tissue specific genes which then can be used for further research.

The amount of data which represents the basis of the tool is extendable by importing new NGS expression data. The software package in the current version is a fully functional platform which is available for the scientific staff at TRON and is furthermore under active development.

[Mortazavi *et al.*, 2008] Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by rna-seq. *Nat Methods*, 5(7):621-628.