Chipster is a user-friendly analysis software for next generation sequencing data with interactive genome viewer.

Massimiliano Gentile, Aleksi Kallio, Taavi Hupponen, Petri Klemelä and Eija Korpelainen
CSC - IT Center for Science, Espoo, Finland

Chipster is a Java-based client-server system with a rich collection of up-to-date analysis tools for high-throughput and next generation sequencing (NGS) data, interactive visualizations and workflow functionality. Version 2.0 includes functionality for analysis of ChIP-seq, RNA-seq and miRNA-seq data, and its integrated genome browser allows interactive and smooth visualization of sequence reads and analysis results in their genomic context.

Chipster is a generic platform, which has now been extended for NGS applications and supports integration of different data types.

Powerful preprocessing tools allow users to perform quality control with the FASTX toolkit and align reads using Bowtie or BWA. Integration of the SAMtools package enables format conversion from SAM to BAM, as well as sorting and indexing. In addition, Chipster features capabilities to find, remove, fuse and combine overlapping genomic regions.

The ChIP-seq analysis tools enable users to detect peaks with MACS, filter them based on p-value, number of reads etc., and scan them for common sequence motifs to be matched against the JASPAR database. Identification of nearby genes, filtering with regards to peak distance, location and various genomic features is also possible. Interpretation of results is aided by incorporating biologically relevant information through pathway analysis.

Data from both single and multi-factor RNA-seq and miRNA-seq experiments can be normalized and analyzed for differential expression through integration of the edgeR Bioconductor package. Identification of predicted miRNA gene targets is provided via a number of databases and gene sets can be tested for overrepresentation of biologically relevant classifications with pathway analysis tools.

The ability to view mapped sequence reads and results in their genomic context greatly benefits the data analysis and interpretation of results. Smooth zooming and navigation makes it easier to keep the sense of location and context, and the ability to zoom in to nucleotide level enables a more detailed inspection. Large datasets are not loaded into memory, but instead sampling and hierarchical indexing is used for efficient generation of large scale data views.

Technical implementation

Chipster is a Java-based client-server system with messaging architecture. The analysis jobs are run on the server side, which enables them to benefit from the CPU and memory of central computing servers. Analysis packages and the associated databases are maintained and updated centrally, which makes it easy for the users and the bioinformatics core facilities supporting them.

SUMMARY

Chipster 2.0 is a software for the analysis of high-throughput and NGS data providing:

- Intuitive graphical user interface
- Integrated genome browser for viewing reads and results, from whole chromosome to nucleotide level
- Tools for QC, alignment, region matching and statistical analysis for ChIP-seq, RNA-seq and miRNA-seq data
- Ability to integrate different data types
- Possibility to reuse and share workflows
- Runs on Windows, Linux and Mac OS X
- Open source (GPL)

http://chipster.csc.fi
http://chipster.sourceforge.net

Contact: chipster@csc.fi