Institut für Informatik
Praktische Informatik und Bioinformatik
Prof. Dr. Ralf Zimmer
Dr. Gergely Csaba
Dr. Florian Erhard

Wintersemester 2015/2016

GoBi-Conference: Call for Papers, Presentations, Posters

Conference date: Fri, 15.04.2016, 10:00
Conference venue: Munich, Amalienstr. 17

The GoBi'16 conference addresses fundamental questions about the interpretation of RNA-seq experiments, in particular a data set from the ENCODE ([4]) consortium comparing two tissues (liver, melanocyte of skins), and a recent time series on H1 cell differentation from Kurian et. al. ([5]).

The submitted manuscripts, presentations and posters should cover the following topics (but are not limited to):

- Differences in basic properties of mapped read data depending of the read-mapping method used.
- Differential expression analysis: influence of mapping and DE-methods, interpretation of differences between methods.
- Differential splicing between tissues and upon cell differentiation: novel methods, in-depth analysis of examples for differential splicing.
- Expressed intergenic regions: finding, characterization (splicing, differential behaviour, potential co-regulations)
- Gene set enrichments: on differential expressed genes, on diff. spliced genes, on sets of genes with (anti-) correlated behaviour over time course experiments.

Hints and possible extensions:

You find the needed mappings (bamfiles) under:
`/home/proj/biosoft/praktikum/genprakt/assignments/data/bams`

- **Differential splicing**: one simple option to use the available toolkit for differential splicing analysis is to extract read counts for specific isoforms / splice junctions / intron retentions, apply methods for differential expression on the extracted counts and compare the fold changes between different isoforms / splice junctions / intron retentions of the same gene
- **Expressed intergenic regions**: one can compile new "transcripts" using covered regions of the genome from "intergenic" reads. By extracting read counts of these (replicate consistent) transcripts one can analyse their differential behaviour.
- **Differential expression analysis without replicates**: for the time course dataset there are no replicates provided. While the methods used in the toolbox provide workarounds for such cases, one can also try to apply the local-foldchange method to derive fold-changes along with their credibility intervals and extend the consistency analysis of DE-methods.
- **Gene set / network enrichment analysis**: There are several ways to extend the currently used gene get analysis package: one can check for active transcription factors by performing the enrichment analysis on known target gene sets, try to condensate the findings from GO enrichment analysis by clustering the enriched gene sets [6, DAVID], integrating the GO-DAG structure into the enrichment analysis [1, 3], taking into account overlaps between KEGG pathways [7], or integrating regulatory network information into the enrichment analysis [8, 2].

Submission checkpoints abd acceptance:

Prior to the submission every planned submission has to pass three submission-checkpoints, each about at one third of the work planned for the submission. To pass a submission-checkpoint you have to make an appointment with the organizers and inform them about the progress and plans. In addition, the submitting groups are encouraged to contact the organizers any time for discussion. All submissions passing the checkpoints will be accepted for presentation at the conference.

Literatur

[1] A. Alexa, J. Rahnenführer, and T. Lengauer. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13):1600–7, Jul 2006.

[2] L. Geistlinger, G. Csaba, R. Küffner, N. Mulder, and R. Zimmer. From sets to graphs: towards a realistic enrichment analysis of transcriptomic systems. *Bioinformatics*, 27(13):i366–73, Jul 2011.

[3] S. Grossmann, S. Bauer, P.N. Robinson, and M. Vingron. Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics*, 23(22):3024–31, Nov 2007.

[4] M. Kellis, B. Wold, M. P. Snyder, B. E. Bernstein, A. Kundaje, G. K. Marinov, L. D. Ward, E. Birney, G. E. Crawford, J. Dekker, I. Dunham, L. L. Elnitski, P. J. Farnham, E. A. Feingold, M. Gerstein, M. C. Giddings, D. M. Gilbert, T. R. Gingeras, E. D. Green, R. Guigo, T. Hubbard, J. Kent, J. D. Lieb, R. M. Myers, M. J. Pazin, B. Ren, J. A. Stamatoyannopoulos, Z. Weng, K. P. White, and R. C. Hardison. Defining functional DNA elements in the human genome. *Proc. Natl. Acad. Sci. U.S.A.*, 111(17):6131–6138, Apr 2014.

[5] L. Kurian, A. Aguirre, I. Sancho-Martinez, C. Benner, T. Hishida, T. B. Nguyen, P. Reddy, E. Nivet, M. N. Krause, D. A. Nelles, C. Rodriguez Esteban, J. M. Campistol, G. W. Yeo, and J. C. Izpisua Belmonte. Identification of novel long noncoding RNAs underlying vertebrate cardiovascular development. *Circulation*, 131(14):1278–1290, Apr 2015.

[6] B. T. Sherman, d. a. W. Huang, Q. Tan, Y. Guo, S. Bour, D. Liu, R. Stephens, M. W. Baseler, H. C. Lane, and R. A. Lempicki. DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. *BMC Bioinformatics*, 8:426, 2007.

[7] A.L. Tarca, S. Draghici, G. Bhatti, and R. Romero. Down-weighting overlapping genes improves gene set analysis. *BMC Bioinformatics*, 13:136, Jun 2012.

[8] A.L. Tarca, S. Draghici, P. Khatri, S.S. Hassan, P. Mittal, J.S. Kim, C.J. Kim, J.P. Kusanovic, and R. Romero. A novel signaling pathway impact analysis. *Bioinformatics*, 25(1):75–82, Jan 2009.