

Visualization of RNA-Seq results with CummeRbund



Requirements

Before diving into this slide deck, we recommend you to have a look at:

- [Galaxy introduction](#)
- [Quality control](#)



? Questions

- How are RNA-Seq results stored?
- Why are visualization techniques needed?
- How to select our desired subjects for differential gene expression analysis?



Objectives

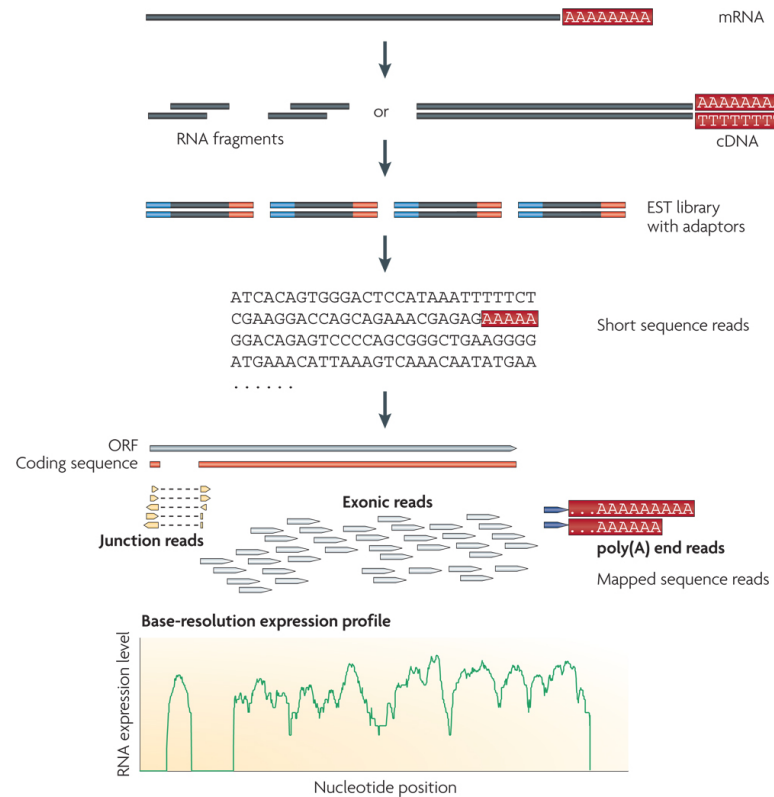
- Manage RNA-Seq results
- Extract the desired subject for differential gene expression analysis
- Visualize information



Why visualization?



Where is my data coming from?



Wang et al, Nat Rev Genet, 2009



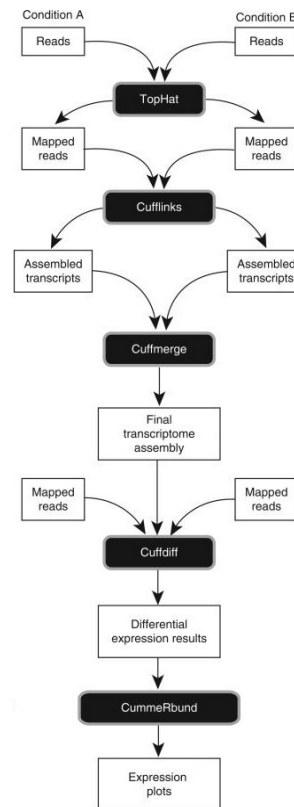
Bioinformatic tools for RNA-Seq analysis

Once the RNA-Seq pipeline is implemented, we still need to handle and analyse all data that is generated. This requires:

- computer science skills to be handled
- mathematical knowledge to be interpreted



Bioinformatic tools for RNA-Seq analysis



Trapnell et al, Nat Protoc, 2012



Bioinformatic tools for RNA-Seq analysis

The last step in our RNA-Seq analysis is CuffDiff. Its output comprises multiple files containing the results of the differential expression analysis.

- Gene expression levels are reported as *tab-separated* values: a simple tabular output that can be viewed with any spreadsheet application. Such files contain statistics, gene-related, and transcript-related attributes

```
» ls
bias_params.info      genes.fpk_tracking    run.info
cds.count_tracking   genes.read_group_tracking  splicing.diff
cds.diff             isoform_exp.diff      tss_group_exp.diff
cds_exp.diff         isoforms.count_tracking  tss_groups.count_tracking
cds.fpk_tracking     isoforms.fpk_tracking   tss_groups.fpk_tracking
cds.read_group_tracking isoforms.read_group_tracking tss_groups.read_group_tracking
gene_exp.diff        promoters.diff         var_model.info
genes.count_tracking read_groups.info
```

- Another way to collect all these data is to organize it within a dedicated database for later consultation. CuffDiff can be instructed to do so

Generate SQLite

Yes No

Generate a SQLite database for use with cummeRbund.



Bioinformatic tools for RNA-Seq analysis

Whatever storage strategy you opted for, i.e. multiple tab-separated-value files or a SQLite database, all data is still retained within text format.

We need to have a bird's-eye view of that data, and *make sense* of it



Visualization



CummeRbund

CummeRbund is an R package for visualizing the results of a CuffDiff output.

- Manages, integrates, and visualizes all data produced by CuffDiff
- Simplifies data exploration
- Provides a bird's-eye view of the expression analysis
- Helps creating publication-ready plots



CummeRbund

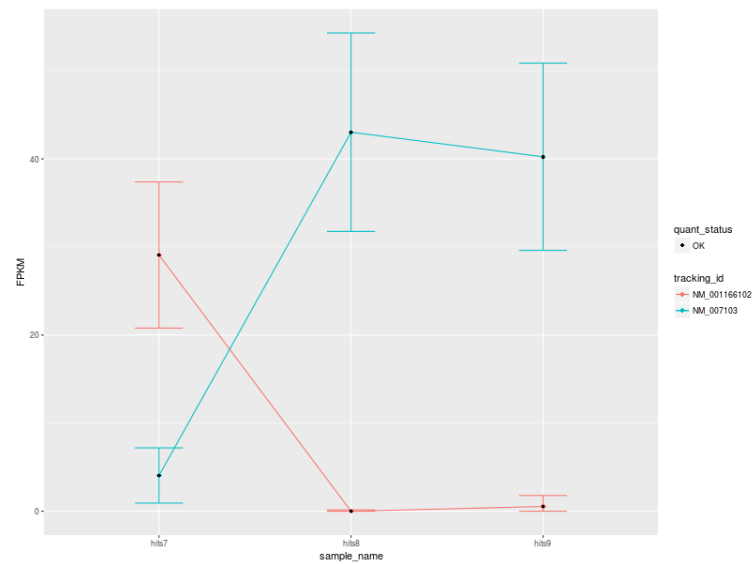
CummeRbund needs to be instructed on which data to be visualized:

- *Extract CuffDiff's* "Transcript differential expression testing" table
- *Filter* the table on the column storing the significance of a differentially expressed gene
- *Sort* all entries on the basis of most significant differentially expressed gene
- Identify the most significant differentially expressed gene



CummeRbund

Once the most significant differentially expressed gene has been identified, CummeRbund can generate publication-ready plots to highlight...

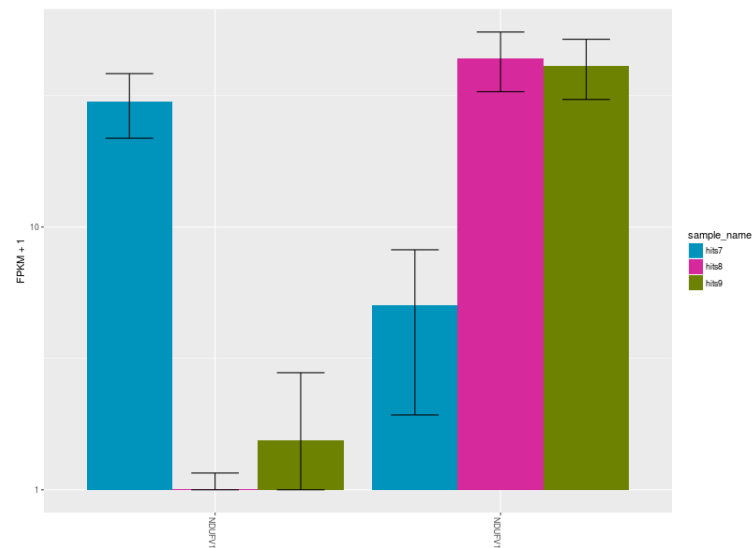


The expression of all isoforms of the single gene with replicate FPKMs



CummeRbund

Once the most significant differentially expressed gene has been identified, CummeRbund can generate publication-ready plots to highlight...

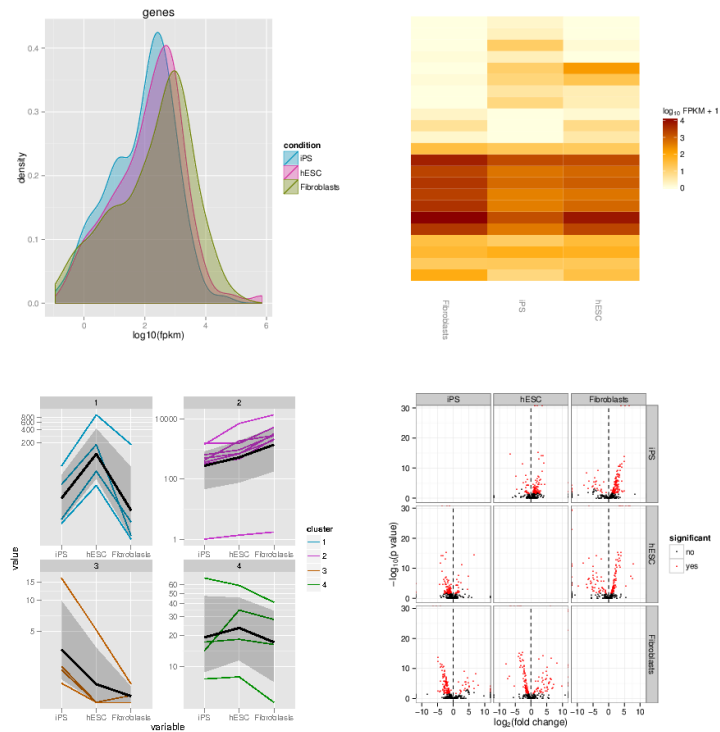


The expression bar-plot of all isoforms of a gene with replicate FPKMs



CummeRbund

...and many more



Have a look at [CummeRbund's tutorial](#) to overview all possibilities!



! Key points

- Extract informations from a SQLite CuffDiff database
- Filter and sort results to highlight differential expressed genes of interest
- Generate publication-ready visualizations for RNA-Seq analysis results



Thank you!

This material is the result of a collaborative work. Thanks the [Galaxy Training Network](#) and all the contributors (Andrea Bagnacani) !



Found a typo? Something is wrong in this tutorial?
Edit it on [GitHub](#)

