Structured Analysis and Integration of RNA-Seq experiments (de.STAIR)

Fkz 031L0106C

Andrea Bagnacani, Markus Wolfien, Olaf Wolkenhauer
Dept. of Systems Biology & Bioinformatics, University Rostock

Short description of the project

The Galaxy platform is being adopted by a growing body of Life Science communities. However, with its increasing number of tools, it becomes more difficult for users to make informed decisions about the selection of tools and implication of their parameterizations. For this reason, the Galaxy Training Network is organizing an extensive index of best-practice data analysis approaches. This material guides users through the construction of workflows for different types of analyses and experimental setups. However, the absence of a systematic overview of the alternative tools of a Galaxy instance does not train users on how to examine alternative best-practice algorithms and parameterizations.

We address this problem by dividing a workflow into its tasks, and provide, for each of them, alternative implementations in the form of interactive Galaxy tours. We call these entities atoms. At the same time, we couple each atom with a help section, to assist users in the selection of the most suitable tool and parameterization. Finally, we reorganize the training material as a series of user-selectable strategies for the customization of their workflows (see Fkz 031L0106D).

Progress report

Interactive comparison of the available alternative atoms

<table>
<thead>
<tr>
<th>Task</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation</td>
<td>Quality control</td>
<td>Data preprocessing</td>
<td>Quality re-check (optional)</td>
</tr>
<tr>
<td>FastQC</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Cutadapt</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Trinity</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Sample atoms to complete a task in the analysis

Workflow customization

Front-end dialogs for tool selection

Back-end workflow customization

General information on the project

Funded by de.NBI staff
Andrea Bagnacani (Scientist)
Prof. Olaf Wolkenhauer (PI)
Markus Wolfien (Scientist)
Dr. Martin Scharm (IT administrator)

Additional staff involved


d.e.NBI services

Galaxy Atoms & Workflow generator plugin
https://github.com/destairdenbi

TRAPLINE
10.1186/s12859-015-0873-9

TriplexRNA
10.1093/nar/gku465

de.NBI Training and education

Introduction to RNA-Seq data analysis with Galaxy
EASYM Berlin

Introduction to RNA-Seq data analysis with Galaxy
EASYM Ljubljana
29.03.2017 - 29.03.2017

RNA-Seq data analysis with Galaxy for clinical applications
GMDS Oldenburg
18.09.2017 - 18.09.2017

A primer for RNA-Seq processing, interpretation and visualization
Freiburg
04.10.2017 - 06.10.2017

Introduction to RNA-Seq data analysis with Galaxy
Kiel
07.03.2018 - 07.03.2018

A primer for RNA-Seq processing, interpretation and visualization
Jena
27.06.2018 - 29.06.2018

RNA-Seq data analysis with Galaxy for clinical applications
GMDS Osnabrück
04.09.2018 - 04.09.2018

Galaxy for linking Bisulfite sequencing with RNA sequencing
Rostock
06.03.2019 - 08.03.2019

RNA-Seq data analysis with Galaxy for clinical applications
GMDS Dortmund
08.09.2019 - 08.09.2019

Galaxy for linking Bisulfite sequencing with RNA sequencing
Freiburg
09.10.2019 - 11.10.2019

Publications


Batut et al. (2018). Community-driven data analysis training for biology. Cell Systems, 6(6), 752-758