Aurora Galaxy Tools: Using R Markdown as a framework to build interactive Galaxy tool output reports

Ming Chen¹, Bradford Condon², Abdullah Almsaeed², Connor Wytko³, Stephen Ficklin³, Margaret Staton²

¹ University of Tennessee, Knoxville, Tennessee (USA). Email: ming.chen0919@gmail.com
² University of Tennessee, Knoxville, Tennessee (USA)
³ Washington State University, Pullman, Washington (USA)

Source code: https://github.com/statonlab/aurora-galaxy-tools
License: MIT license

Galaxy enables computational analysis in an intuitive web-based interface without programming [1], and its success relies on the thousands of Galaxy tools contributed by numerous developers from the community. The high diversity of bioinformatic tools increases the scope of Galaxy to support many different types of projects. Unfortunately, it also results in many tool output formats, which often lack user-friendly reports or summaries to guide interpretation and exploration of the results. We propose to use R Markdown [3] as a framework for Galaxy tool development as a straightforward way to construct custom reports describing tool output while wrapping any command line tool. Tools developed under the R Markdown framework generate HTML-based analysis reports which can include both the analysis code and visualizations of results (Figure 1). These reports can range from simple and quick descriptions to sophisticated, interactive visualizations. Reports can be viewed directly within the main Galaxy interface, downloaded by the user to view or share, and ported to other websites, such as Tripal databases, that call Galaxy workflows remotely. We have wrapped an initial set of bioinformatic software packages, the Aurora Galaxy tools, to demonstrate this approach and provide analysis reports for some of the most common next generation sequencing software tools. This presentation will demonstrate the outputs from these tools and also describe the process of developing within the R Markdown framework.

Figure 1. Example outputs from two Aurora Galaxy tools: aurora_fastqc (A1 & A2) and aurora_star_site (B1 & B2). aurora_fastqc is a wrapper of the short reads evaluation tool FastQC and aurora_star_site is a wrapper of the RNA-Seq mapping tool STAR. The output report from aurora_fastqc is a single HTML file which has links to the original outputs from FastQC (A1) and use interactive plots to visualize summarised results. The output report from aurora_star_site is a static website which provides analysis code to show how the mapping process is done (B1) and mapping statistics from the samtools flagstat tool.