Organization Name

Princess Margaret Cancer Centre/University of Toronto

Organization Website

https://hoffmanlab.org/

Organization Type

Research Lab

Product

From a Galaxy web interface, users can select publicly available genomic datasets, or upload their own and then run Segway to identify recurring patterns across the entire genome to create a non-overlapping segmentation.

Existing Product

We would like to build a tool for the web-based scientific workflow manager Galaxy (https://galaxyproject.org/) that allows people to easily use our command-line automated genome annotation software, Segway (https://segway.hoffmanlab.org/).

Both Segway and Galaxy are both developed but integrating them together is not. Segway is largely written in Python (https://bitbucket.org/hoffmanlab/segway). Galaxy's tool architecture involves creation of XML configuration files (https://galaxyproject.org/admin/tools/add-tool-tutorial/.

Users/Stakeholders

This project would aid genome researchers who are less technically proficient to create and visualize genome annotations.

Key functionalities & features

A protocol posted on bioRxiv (https://doi.org/10.1101/080382) shows the command-line steps that would need to be implemented as a web tool for a minimum viable product. We can supply an updated draft of this protocol that is somewhat simpler.

This project will have intermediate datasets (such as genomic dataset storage) and outputs (probabilistic parameters, segmentations) used in further analysis and visualization within Galaxy. We would like to limit the project in scope. However, keen students may take interest in possible CI integration and regression testing if any internal software issue arises.

Value of the project to the organization

This project would be beneficial to our organization and the larger scientific community since it will permit researchers without a technical background to perform a very technical analysis of any set genomic data they wish. There are a significant number of researchers and clinicians who may only specialize in performing biological experiments and forgo or delegate the technical analysis and this product would help bridge that gap in knowledge.

Additional Notes

The project aims to be published to the "app store" of Galaxy where it would be publicly hosted. The code and any necessary additional products will be stored and revisioned our public repository (https://bitbucket.org/hoffmanlab/segway). We plan to maintain the completed project with our existing staff.