ASSESSING NEW TOOLS AND BEST PRACTICES FOR RNA SEQ DATA ANALYSIS AND VISUALIZATION WITH IPLANT CYBERINFRASTRUCTURE

Kapeel M Chougule[1,4], Andrew Olson[1], Jurandir Vieira De Magalhaes[2], Peter Van Buren[1,4], Liya Wang[1,4], Doreen Ware[1,3,4]

[1]Cold Spring Harbor Laboratory, Ware Laboratory, Cold Spring Harbor, NY, [2]Embrapa, Parque Estação Biológica, Brasília, Brazil, [3]USDA ARS NEA Plant, Soil & Nutrition Laboratory Research Unit, Ithaca, NY, [4]iPlant Collaborative, Bio5 Institute, Univerity of Arizona, Tucson, AZ

Common analysis steps in all RNA-seq experiments include quality control, read alignment, assigning reads to genes or transcripts, and estimating gene or transcript abundance. Choice of analysis tool and experimental strategy are critical in ensuring high quality data for analysis and visualization of the results. In this study we compare and assess new tools available for spliced alignment(STAR & HISAT), assembly(StringTie) and quantification(Ballgown,eXpress,Sailfish,Kallisto and Salmon) for RNA seq data using cyberinfrastructure(CI) services developed by iPlant collaborative project.

The CI unites high-end computing, large scale data storage and networking with user-interfaces that simplify creation, integration and sharing of software tools and workflows. Here we compare the established Tuxedo protocol for RNA seq analysis against the new tools and evaluate the performance of the workflows on the CI. Most of these new tools were benchmarked on human RNA seq datasets, we want to test the efficacy and robustness of these tools on the CI using RNA seq data from plants.The alignment and assembly tools were integrated in Discovery Environment platform that executes jobs in a Docker container while the quantification tools were integrated as a virtual image in Atmosphere cloud computing service within the CI.The results were visualized on a Biodalliance genome web browser for assessing quality of the annotation.

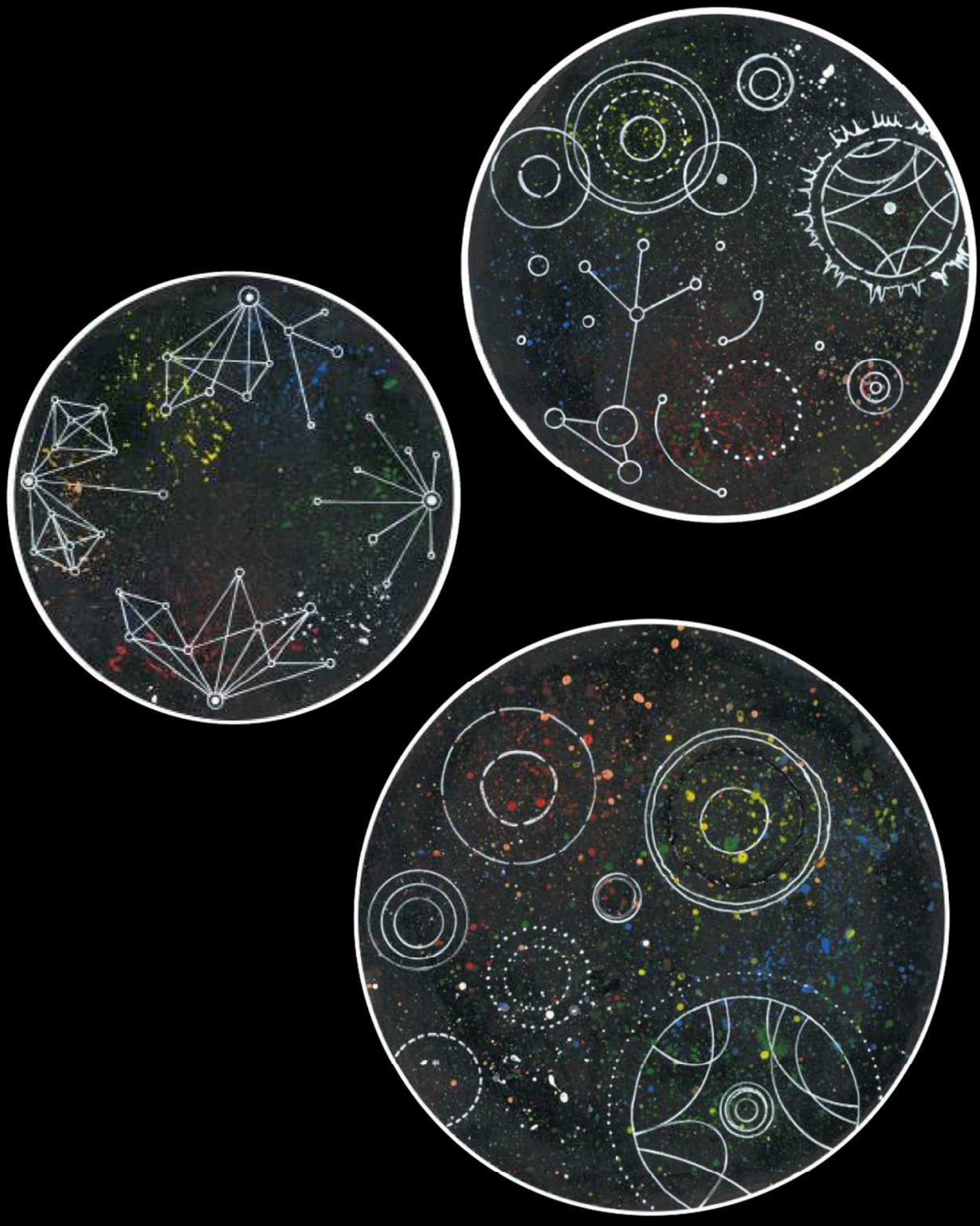
The iPlant Collaborative is funded by a grant from the National Science Foundation (#DBI-0735191)

← Back

Abstracts of papers presented
at the 2015 meeting on

GENOME INFORMATICS

October 28–October 31, 2015



CSH Cold Spring Harbor Laboratory 1890 125 2015

Abstracts of papers presented
at the 2015 meeting on

GENOME INFORMATICS

October 28–October 31, 2015

Arranged by

Janet Kelso, *Max Planck Institute for Evolutionary Anthropology, Germany*
Daniel MacArthur, *Massachusetts General Hospital*
Michael Schatz, *Cold Spring Harbor Laboratory*