Building the Internet of Tools With CyVerse: The Cloud Integrated RNA-sequencing Platform
fRNAkenseq

Allen Hubbard
PhD Student, Bioinformatics Systems Biology

ABSTRACT:
Scalable bioinformatics job execution that exploits multiple platforms is a challenge of the sequencing era. As processing algorithms become faster and average size of datasets increases, the data burdens encountered by a laboratory are daunting. A strategy to counteract these challenges is to develop a modern generation of tools capable of seamless data sharing and communication by leveraging remote, distributed computing environments. fRNAkenseq is a powered-by-CyVerse informatics platform that accomplishes these goals for RNA-seq analysis. fRNAkenseq leverages APIs across multiple bioinformatic resources to execute RNA-seq analyses in a scalable fashion and to exploit cloud integration and cross-talk with other tools. The backend of fRNAkenseq is fully integrated with the Agave API’s developed by the Texas Advanced Computing Center (TACC) for computational scalability. The architectural design of fRNAkenseq provides a blueprint for the accelerated development of server-based and cloud-integrated informatics tools using CyVerse resources to keep pace with the explosion in sequencing data.