**APOLLO:**
Improving collaborative genome curation

Monica Munoz-Torres | @monimunozto
Nathan Dunn, Deepak Unni, Colin Diesh, Eric Yao, Christine Elsik, Ian Holmes, Suzanne E. Lewis
Berkeley Institute for Integrative Studies | School of Computer Science and EE | UC Berkeley

Outline

- I will show you how we are empowering curators with the latest improvements in Apollo, and discuss how these changes can assist in advancing your research.

**Anatomy of a genome sequencing project**

**We must care about curation**

The gene set of an organism informs a variety of studies:
- Characterization: Gene number, GC%, TE, repeats.
- Functional assignments.
- Molecular evolution, sequence conservation.
- Gene families.
- Metabolic pathways.
- What makes an organism what it is? What makes a bee a “bee”? 

**Genome Curation**

**APOLLO: versatile genome annotation editing**

- Apollo is a web-based genome annotation editor, integrated with JBrowse
- Supports real time collaboration & generates analysis-ready data
Adding a gene model

Adding an exon supported by experimental data

- RNAseq reads show evidence in support of a transcribed product that was not predicted.
- Add exon by dragging up one of the RNAseq reads.

Editing functionality

Adjusting exon boundaries supported by experimental data

- RNAseq reads show evidence in support of a transcribed product that was not predicted.
- Add exon by dragging up one of the RNAseq reads.
Improvements

Architecture and Interface improvements can accommodate a variety of genome projects and support the needs of our growing research community.

ARCHITECTURE: Supports extension and integration
1. Web-based client
2. Annotation-editing engine
3. Server-side data service

Web-based Client

GWT / AngularJS / Bootstrap on front-end provide rich application behavior.

REST / Websockets to communicate with server for flexibility, speed, and service reuse.

Annotation Engine

GRAILS controllers (a J2EE servlet) route requests to the appropriate JBrowse data directory for each organism

Server-side Data Service

Single Data Store PostgreSQL, MySQL, MongoDB, ElasticSearch

Update: Web Services

A suite of secure web services drives the interface, making it easier to integrate with other software and create customized interfaces.
Web Services

Update: Support multiple organisms
Minimizes server resources and allows optional public access.

Interface Updates

Annotator Panel
Interface Updates

Annotator Panel

Interface Updates

Improved user and group level permissions for editing and viewing annotations and supporting participation by larger teams of collaborators.
Interface Updates

Extensible Administrator options to produce customized reports.

Update: Transforming coordinates

Bringing exons closer together to facilitate annotation of gene models with long introns.

Transforming coordinates

Assembly artifacts may cause gene models to be split across two or more scaffolds. To facilitate annotation, Apollo allows the generation of an artificial space where the annotation can be completed.

Apollo Development

BBOP

- Scott Lewis, Principal Investigator
- Nathan Dunn
- Eric Yao
- Deepak Unni

Christine Dike's Lab, University of Missouri
Thank you!

- Berkeley Bioinformatics Open-source Projects (BBOP), Berkeley Lab: Apollo and Gene Ontology teams. Suzanna E. Lewis (PI).
- § Christine G. Elsik (PI), University of Missouri.
- * Ian Holmes (PI), University of California Berkeley.
- Arthropod genomics community: i5K Steering Committee.
- Stephen Ficklin, GenSAS, Washington State University.
- Apollo is supported by NIH grants 5R01GM080203 from NIGMS, and 5R01HG004483 from NHGRI. Also supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

For your attention, thank you!

Start using Apollo at http://GenomeArchitect.org