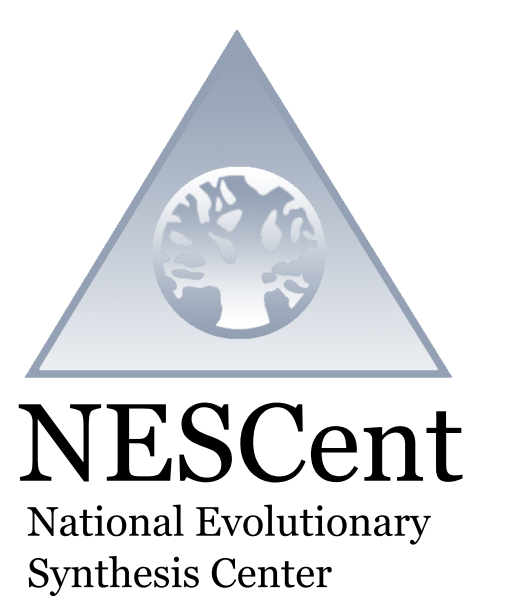


GMOD tools for evolutionary biology

> **NESCent's 4th (!) hackathon** [Nov. 8-12, 2010]

Organizers: Robert Buels, Scott Cain, Dave Clements, Hilmar Lapp, Sheldon McKay, Nicole Washington

Participants: Filip Balejko, Jim Balhoff, Yuri Bendana, Terry Casstevens, Delphine Dauga, Chris Fields, Maren Friesen, Vladimir Gapeyev, Richard Hayes, Yözen Hernández, Andrew Hill, Ben Hitz, Greg Jordan, Jonathan Leto, Eric Lyons, Olaf Mueller, Sander Pick, Sergei Kosakovsky Pond, Weigang Qiu, Seth Redmond, Mitch Skinner, Jason Stajich, Andrew Tritt, Hong Yi, Matt Yoder



> Summary and motivation

NESCent sponsors hackathons to work on open-source biological software projects, as part of its mission to foster cross-disciplinary evolution research. NESCent sponsored the GMOD Tools for Evolutionary Biology Hackathon (**GMODEvoHack**) in November 2010.

This event brought together **31 participants for 5 intense work days** to fill some critical gaps in the capabilities of the GMOD tool set that limited their usefulness in evolutionary research.

Attendees were a mix of bioinformaticists, software developers, and informatics-savvy biologists.

Early in the planning phase, the organizers defined three broad work objectives:

> **GMODEvoHack** work objectives

1. **Comparative genomics visualization**
2. **Visualization of phylogenetic data and trees**
3. **Population diversity and phenotypes support**

These objectives were kept **intentionally very broad**, in keeping with the unique hackathon event format.

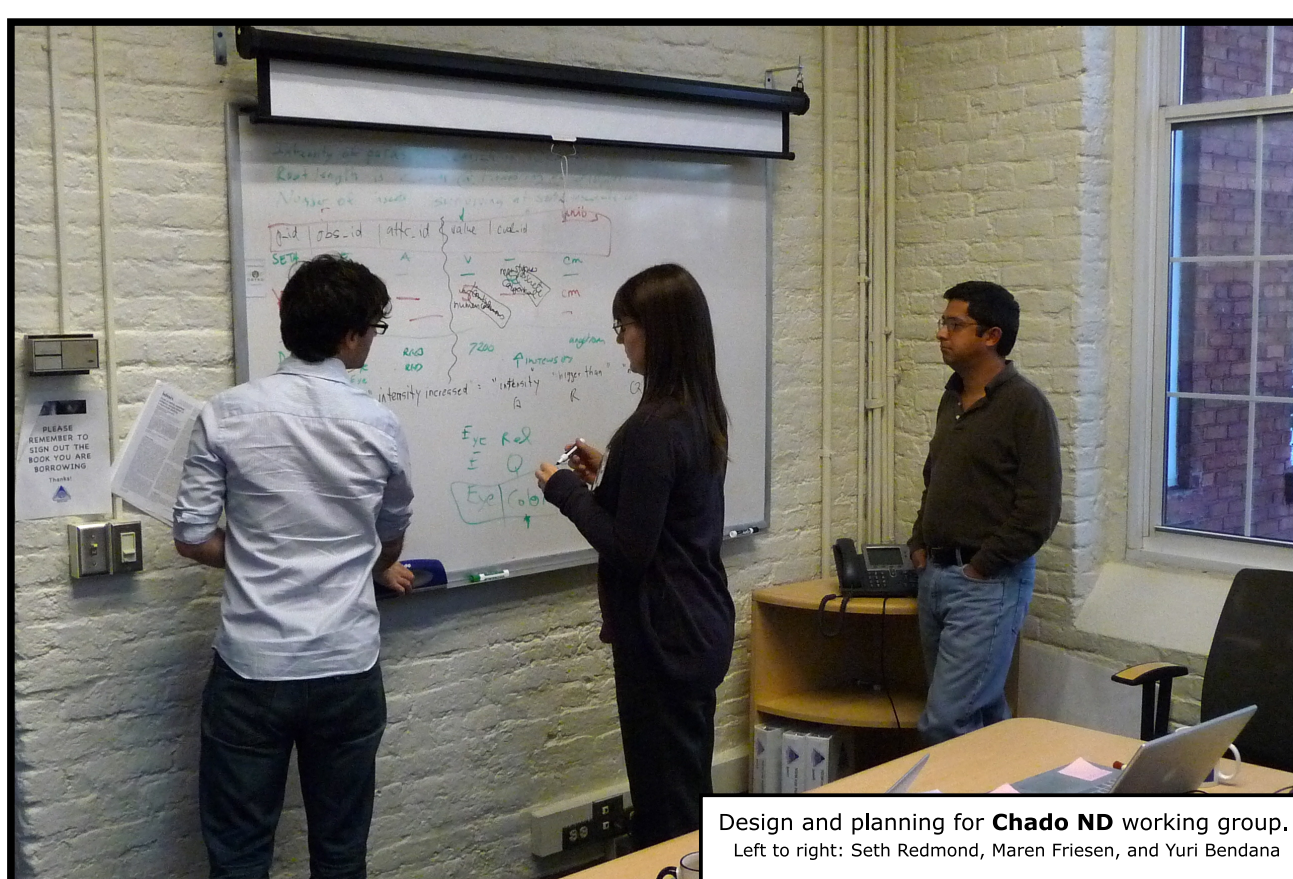
> The **hackathon** format

Before

- * Organizing committee makes work objectives, secures funding and venue, and solicits participant applications, and chooses participants.
- * Committee organizes one or more "pre-hackathon planning" conference calls among the participants to "jump start" their thinking on the event and begin the process of participants setting more concrete work goals.

During

- * On the first day, organizers lead a collaborative self-organization session using the **Open Space Model** to identify concrete implementation goals and form focused (maximum 5 people) working groups for each.



Design and planning for **Chado ND** working group. Left to right: Seth Redmond, Maren Friesen, and Yuri Bendana

- * Groups plan and work *ad libitum* until the last day of the hackathon.
- * In the last hours of the event, each working group presents its outcomes: running code, documentation, and plans for the future.

After

- * Hackathon email lists are kept operational. Organizers actively solicit followup reports from group leaders.



In the thick of the action. This event was hosted at **NESCent**.

> **Emergent working groups**

The 31 participants self-organized into 8 groups using the Open Space Model.

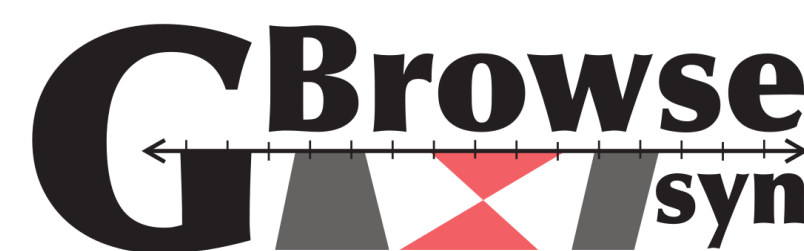
http://www.nescent.org/wg_gmodevohackathon/Open_Space_Model

GMatchbox

Establish a common database backend and JSON-based service API for comparative genomics data.

GBrowse_syn2

Refactor GBrowse_syn to fit more naturally into the GBrowse 2.x-series codebase.



JBrowse_syn

Extend JBrowse with comparative genomics features.

BioPerl

Improving tree handling in BioPerl, especially large trees and large numbers of small trees.

PhyloBox

Extend PhyloBox with more capabilities for interacting with other JavaScript components, also added a Perl server-side API (WWW::PhyloBox) for embedding in web sites.

PhyloBox <-> JBrowse Integration

Work with JBrowse_syn, PhyloBox and GMatchbox groups to enable integration of these three technologies.

Natural Diversity and Phenotypes in Chado (Chado ND)

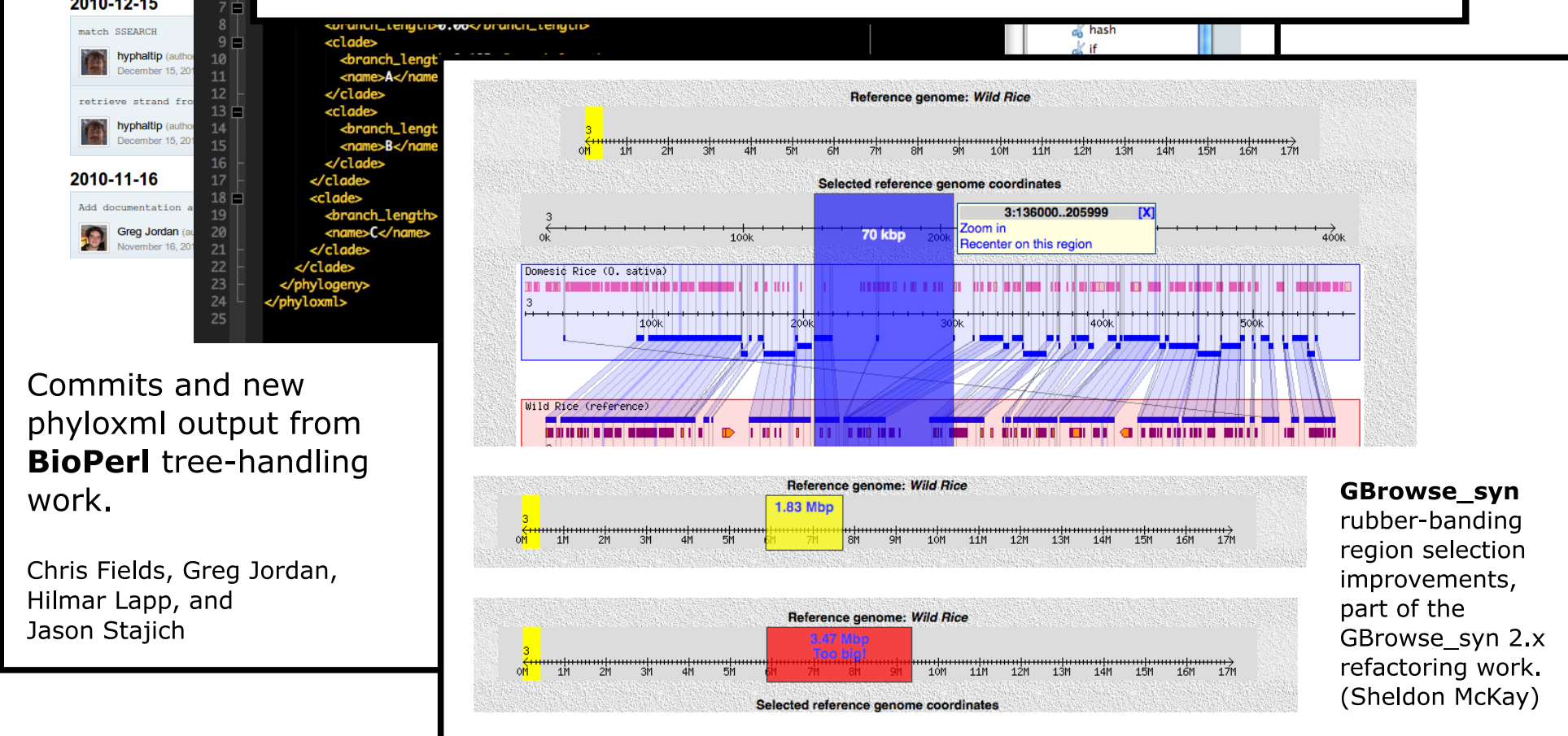
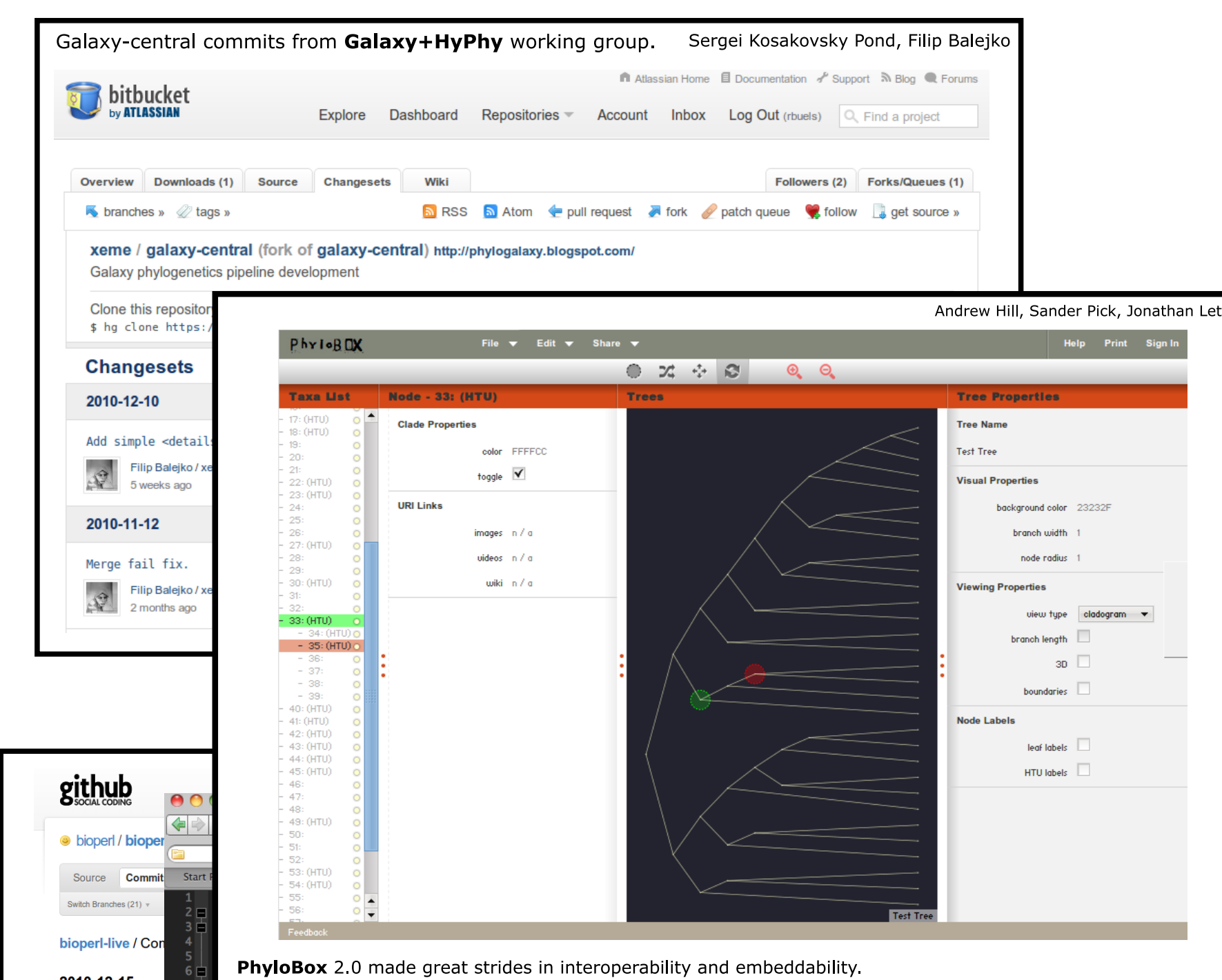
Develop a prototype Rails application as a web interface to the new Chado Natural Diversity (ND) module. Also, polish and further document Chado's Phenotype module.

Galaxy + HyPhy

Improve Galaxy's ability to integrate interactive tools, using HyPhy as the prototype application.

> Outcomes: software

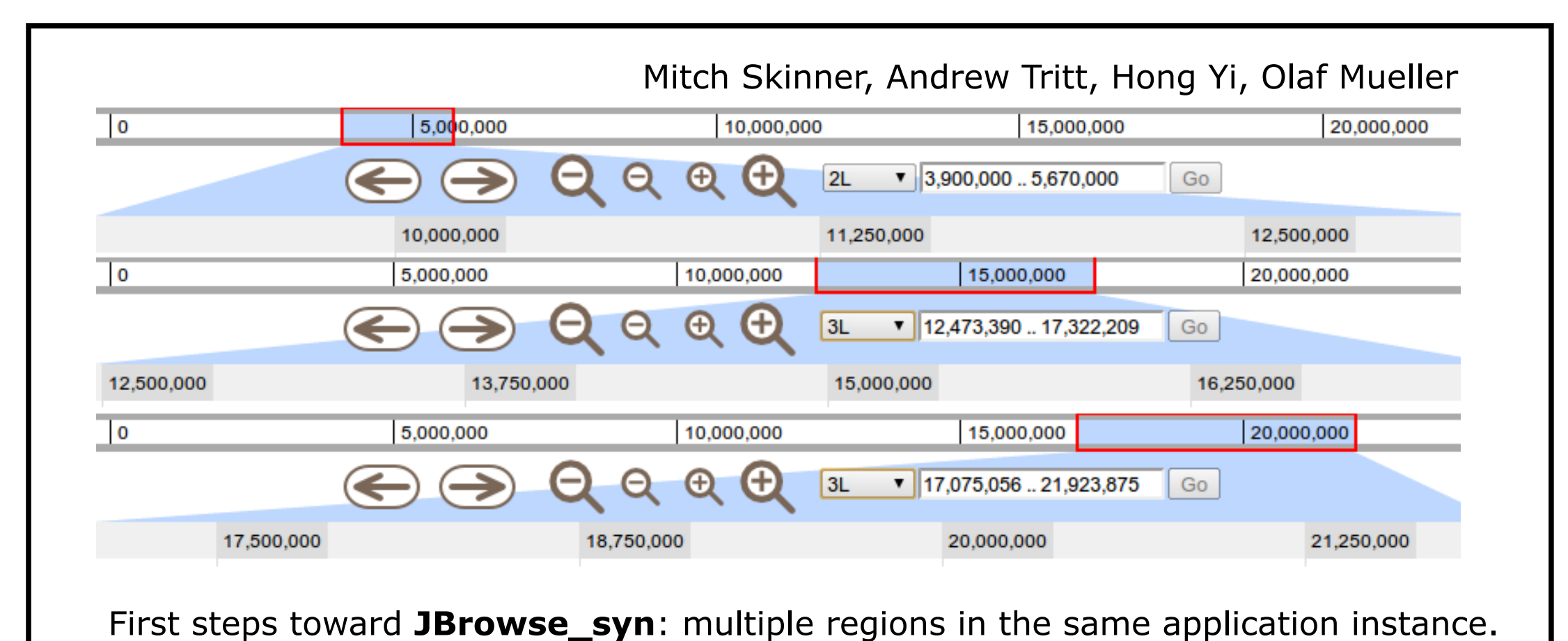
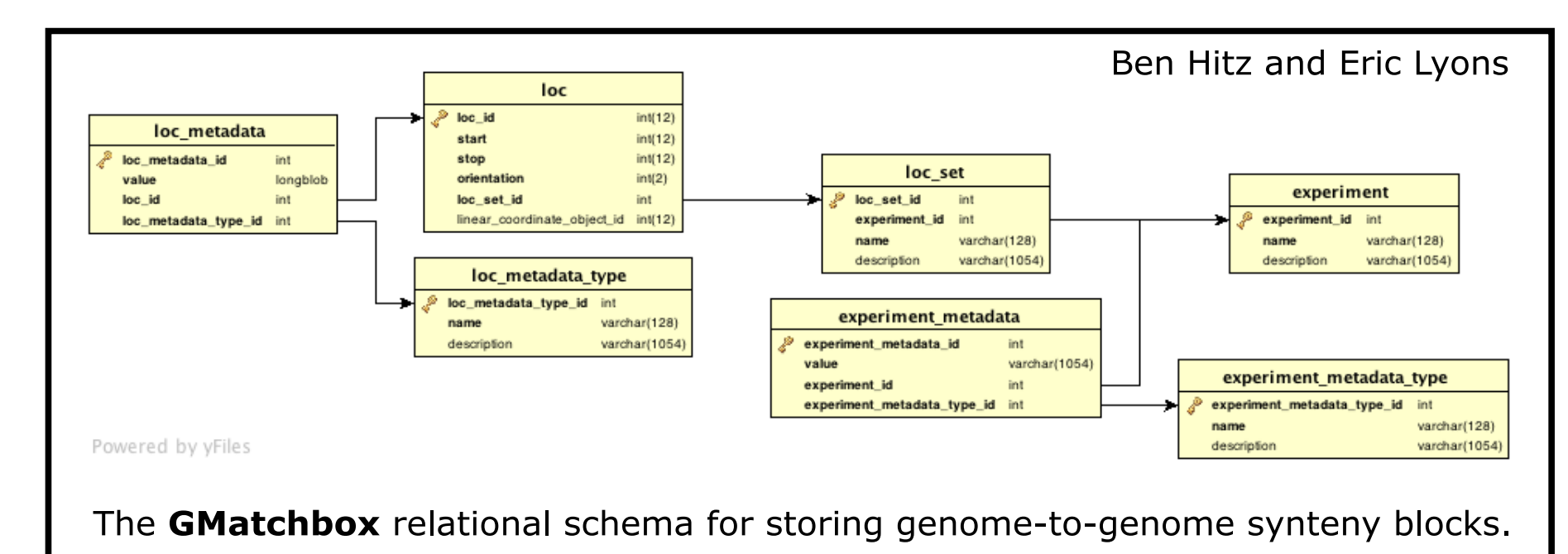
Below are selected screen shots and diagrams showcasing work done at the hackathon.



> Outcomes: community development

GMODEvoHack brought together a very focused group of people from many different communities who may otherwise never have been able to work together in person. As a result:

- * **At least four more people** are now familiar with the JBrowse code, and therefore potential contributors.
- * Non-programmers saw to it that **good documentation** was written.
- * The Chado Natural Diversity module working group solidified, and a **publication of their work is now pending**.
- * As a result of their work on GBrowse_syn during the hackathon, Robert Buels and Richard Hayes were named **GBrowse_syn core developers**.



> Acknowledgments



The US National Evolutionary Synthesis Center (NESCent, <http://nescent.org>) is a collaborative effort of Duke, NC State University and UNC Chapel Hill, and is supported by the National Science Foundation (EF-0423641).

GMOD is supported by a specific cooperative agreement from the USDA Agricultural Research Service, and by NIH grants co-funded from the National Human Genome Research Institute and the National Institute of General Medical Sciences.

We are indebted to the enthusiasm and energy of the hackathon participants that made the event a success.