

A couple of UI prototypes*

Jonathan Crabtree

jcrabtree@som.umaryland.edu

Institute for Genome Sciences (IGS)
University of Maryland, Baltimore

* Dramatization: actual presentation may have been significantly less well-organized

1. New! Dojo tree widget, now with 100% more bar graphs.

- Took standard Dojo tree widget
- Added graphs of node-associated quantities
- Use case: Review counts of 454 sequence reads classified at different taxa in a taxonomic tree, using data from metagenomic survey sequencing classified using the RDP's Naïve Bayesian classifier and TOBA-based taxonomy.

Classification result file: endo_prelim-0.9-data.txt

show 2 levels

Help

3 0% |
 61 1% |
 1 0% |
 726 21% |
 479 13% |
 16 0% |
 12 0% |
 2 0% |
 1010 29% |
 0 0% |
 506 14% |
 2 0% |
 61 1% |
 0 0% |
 166 4% |
 178 5% |
 2 0% |
 22 0% |
 160 4% |
 11 0% |
 1 0% |
 0 0% |
 11 0% |

- endo_prelim, score_threshold=0.9 [3|3430]
 - Bacteria [61|3416]
 - + Acidobacteria [0|1]
 - + Actinobacteria [0|726]
 - + Bacteroidetes [137|479]
 - + Cyanobacteria [0|16]
 - + Deinococcus-Thermus [0|12]
 - Firmicutes [2|1520]
 - + "Bacilli" [5|1010]
 - "Clostridia" [0|506]
 - + Clostridiales [60|506]
 - + "Erysipelotrichi" [0|2]
 - + Fusobacteria [0|61]
 - Proteobacteria [0|528]
 - + Alphaproteobacteria [3|166]
 - + Betaproteobacteria [0|178]
 - + Deltaproteobacteria [0|2]
 - + Epsilonproteobacteria [0|22]
 - + Gammaproteobacteria [1|160]
 - + Spirochaetes [0|11]
 - + Tenericutes [0|1]
 - Unclassified [0|11]
 - too_short [11|11]

Clostridiales [60|506] : 60 sequence(s) classified at this node (506 sequence(s) at this node or below)

seq_id	seqlen	percent_gc	revcomp	classification
YQ002	464	52	0	Root/Bacteria/Firmicutes (0.99)"/Clostridia" (0.99)/Clostridiales (0.99)
YQ008	476	51	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales
YQ009	478	51	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales
YQ012	476	51	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales
YQ022	478	51	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales
YQ024	479	52	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales
YQ029	478	51	0	Root/Bacteria/Firmicutes"/Clostridia"/Clostridiales

Tree/graph widget features

- Synchronized scrolling of graph and tree panes
- Graphs dynamically update as tree nodes are expanded or collapsed.
- Clicking on a tree node loads a list of all the sequence reads classified at that node.
- Graph colors differentiate between: 1. sequences classified at a node and 2. sequences classified at OR below a node.

Tree/graph widget cons

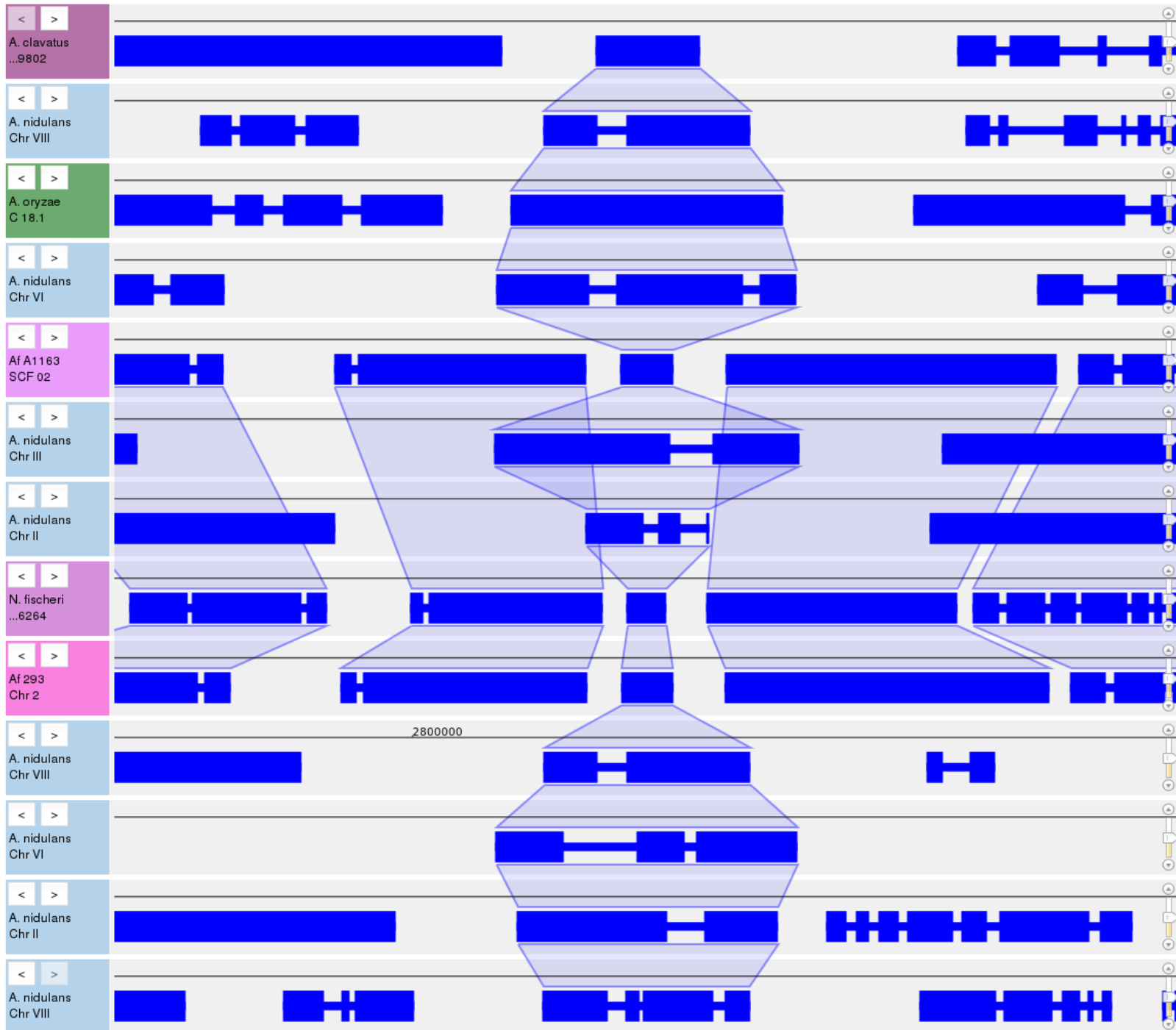
- Not under active development
- No connection with chado tools (currently reads data in simple JSON format)
- Just an idea/prototype for the dynamic display of tree-associated abundance data in the context of a JavaScript framework.

2. An AJAX-Style Synteny Viewer

- Goal: produce a proof of concept of an interactive version of the “Sybil” synteny viewer (see <http://sybil.sf.net>)
- Under (extremely) sporadic development for several years now.
- Uses Nested Containment Lists like JBrowse, but only on the server side.
- Client uses Dojo, server is custom C app that uses libmicrohttpd, Cairo, NCLists.

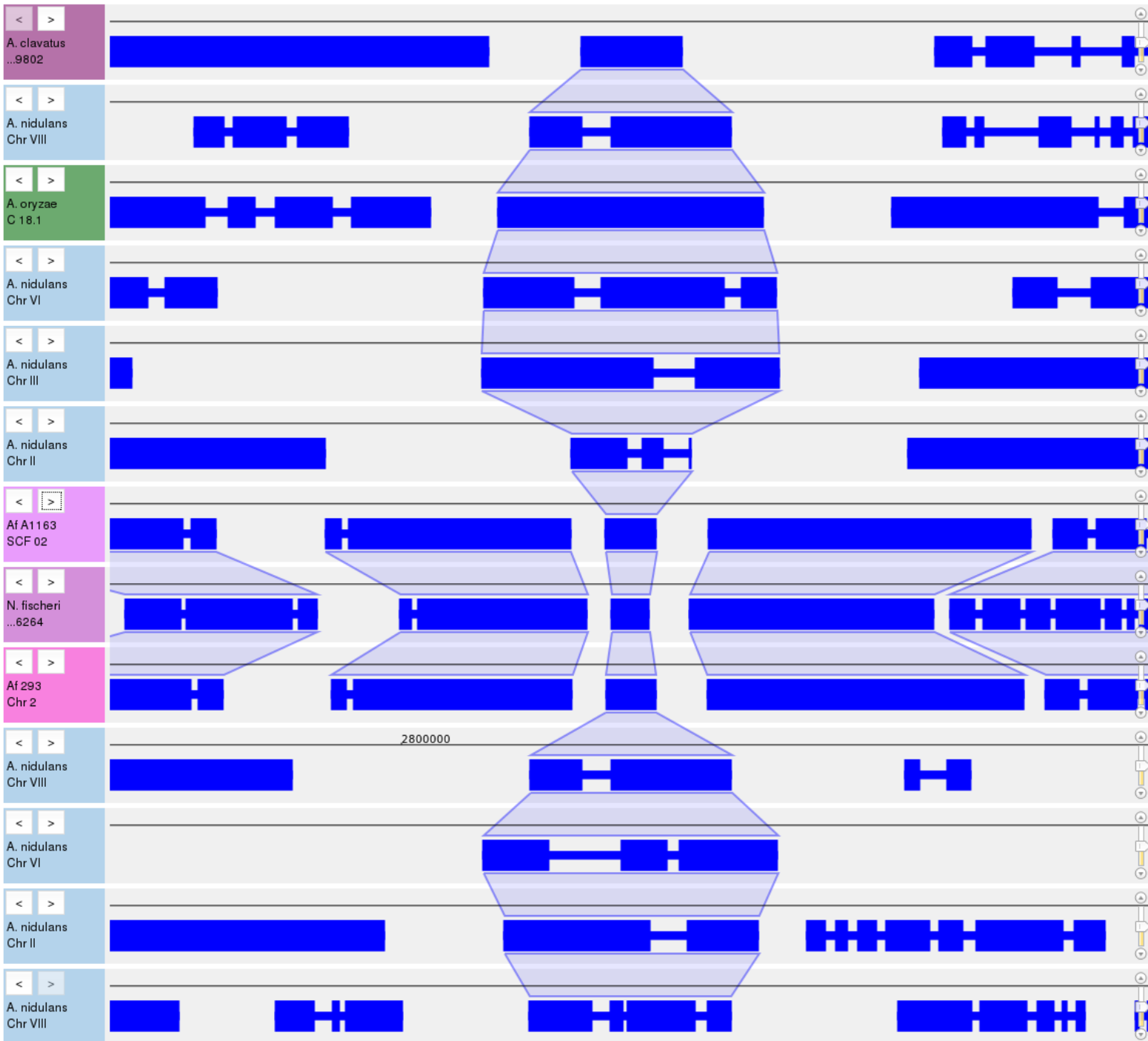
Synteny viewer example

- 13 sequence/6 genome example reading from a 10 genome PostgreSQL chado comparative database
- A protein clustering analysis is used for the comparative data.
- The current prototype displays single protein clusters in their genomic context.
- Sequences can be scrolled by dragging and the protein cluster display updates in real time.
- Color coding indicates source genome.



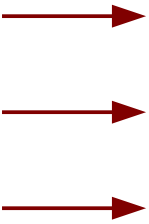
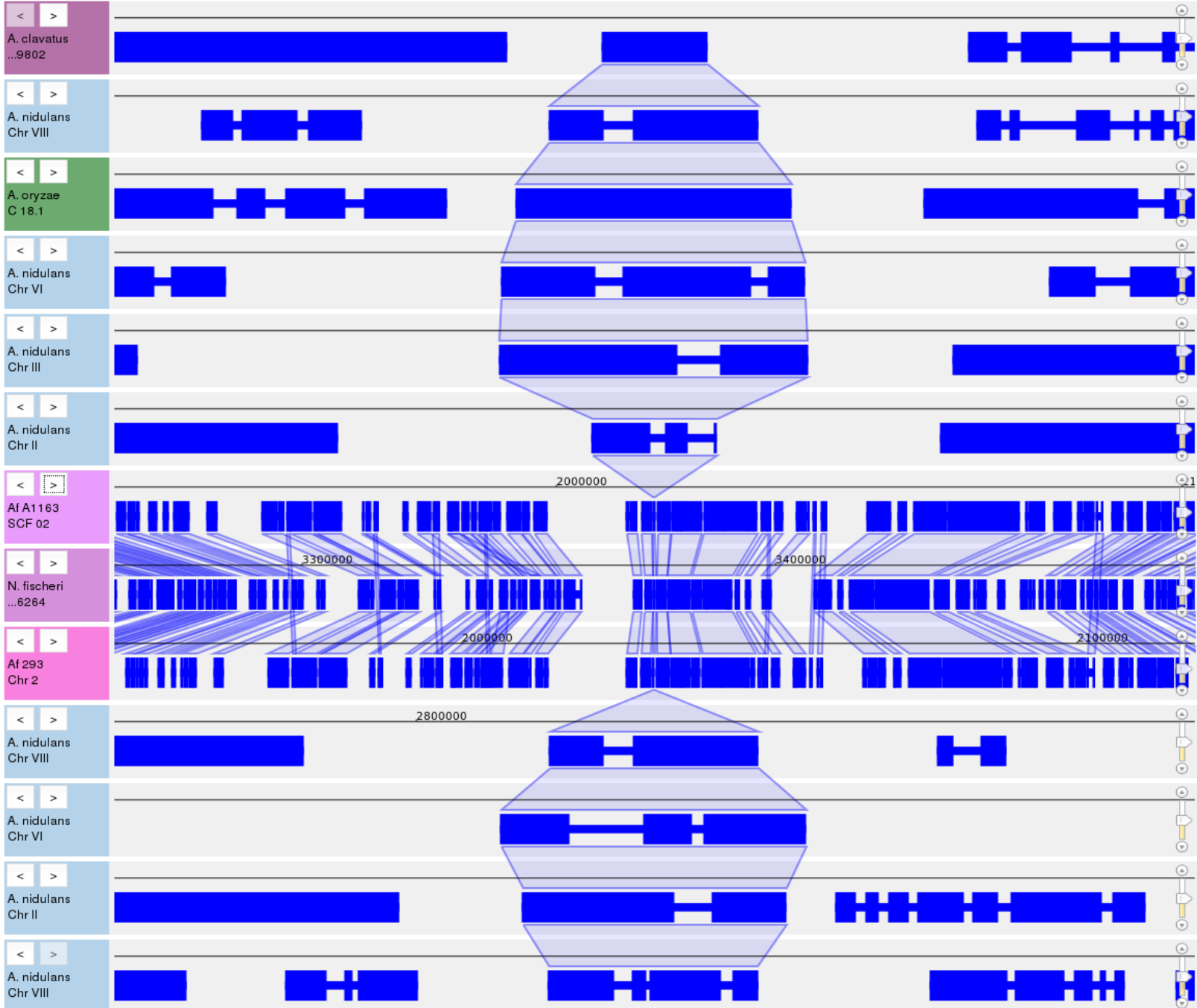
In the next slide...

- Click on the arrows to reorder the sequences, or pick them up and drag them directly.
- Here a sequence (the 5th from the top) is manually moved to be closer to those to which it is more similar (2 rows lower on the page.)



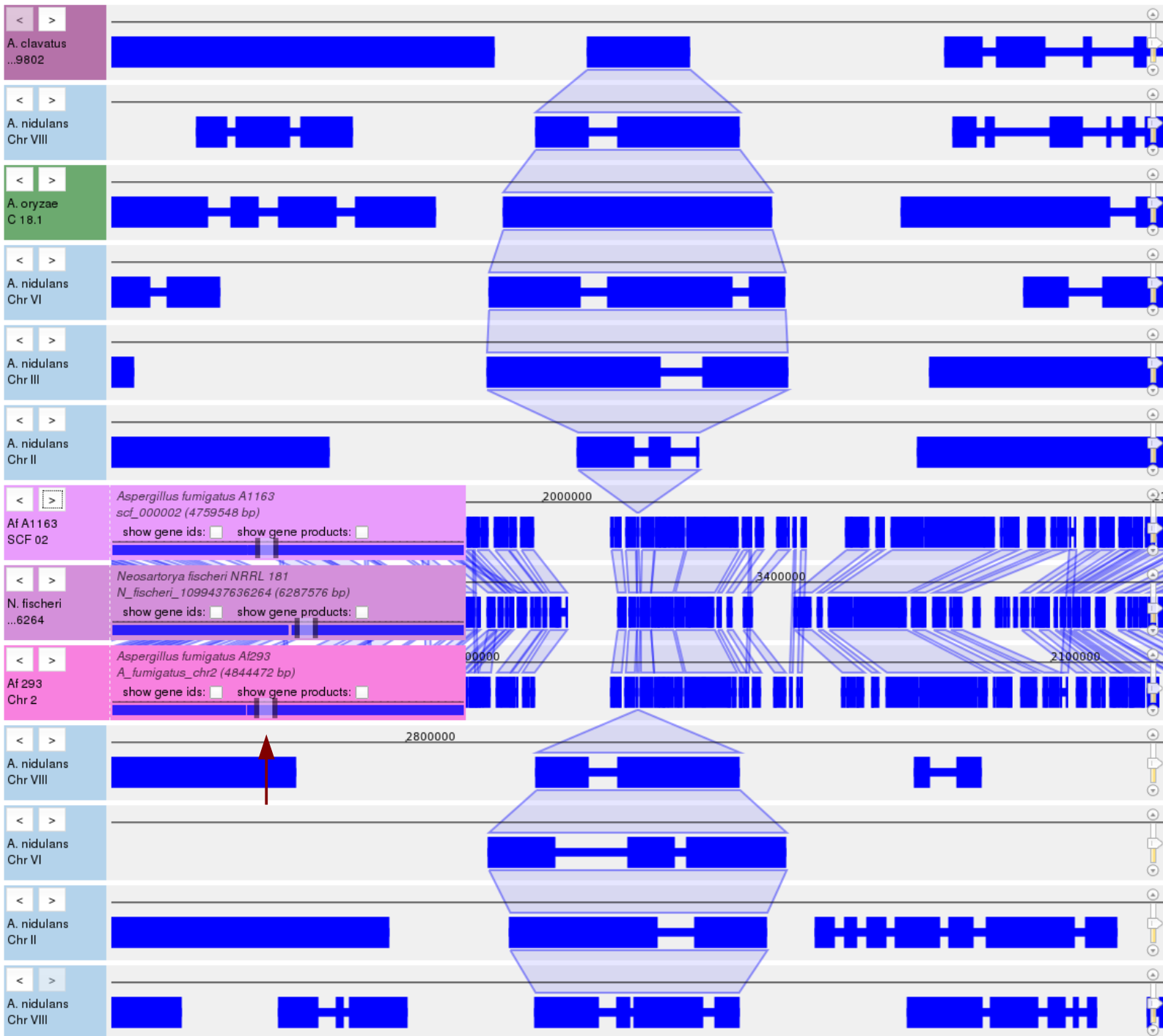
In the next slide...

- Zoom/change scale by using the mouse wheel or the zoom slider at the right edge of each sequence.
- Here we have zoomed out in 3 adjacent sequences.
- Scrolling and zooming is currently independently controlled in each sequence.



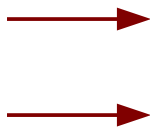
In the next slide...

- Double-click to open up control panel for each sequence.
- Control panel contains an overview of the entire sequence, which shows the currently-viewed position and can also be used as a scrollbar.



In the next slide...

- Click to display gene ids and/or gene product names.
- Gene ids and product names are retrieved asynchronously as needed.



Synteny viewer summary

- Still just a prototype/proof of concept, with a possible alpha release on the horizon if time permits.
- Similar in concept to proposed JBrowse_syn tool
- Current back-end is chado/PostgreSQL only
- No attempt to deal with 2nd gen data here: assumes assembled and annotated sequence data.
- All the heavy lifting is done on the server side, which could potentially be distributed on the cloud (although the current app runs fine—for a small number of users—with client (browser), server, and back-end chado db all on the same laptop.)