wasabi

a web-based platform for evolutionary sequence analysis supporting reproducible research and sharing

Andres Veidenberg
PhD student, University of Helsinki, Finland
NETTAB 2016
wasabi, and:

1) Data provenance
2) Visualization and publishing
3) Plugin system

Details and nice video: http://wasabiapp.org
Portable HTML5 webapp + python server module

http://wasabiapp.org:8000

Background jobs

Dataset library
Tree + sequence alignment

Evolutionary events

Annotations & graphs

Ancestral sequences
DATA PROVENANCE
<table>
<thead>
<tr>
<th>Name</th>
<th>Created Date</th>
<th>Collection ID</th>
<th>Analysis ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wasabi use cases</td>
<td>07.08.15 at 12:09</td>
<td>usecases</td>
<td></td>
</tr>
<tr>
<td>Tiger genome</td>
<td>23.06.15 at 15:03</td>
<td>tiger</td>
<td></td>
</tr>
<tr>
<td>EGLN1 genetree (end result: DNA)</td>
<td>08.07.15 at 14:32</td>
<td>38VMbS</td>
<td></td>
</tr>
<tr>
<td>PAGAN placement of big cats</td>
<td>08.07.15 at 15:53</td>
<td>oTIVt</td>
<td></td>
</tr>
<tr>
<td>Tree pruning</td>
<td>08.07.15 at 16:10</td>
<td>u6MkXH</td>
<td></td>
</tr>
<tr>
<td>Relocated subtree of the big cats</td>
<td>14.08.15 at 17:55</td>
<td>NRWwhe</td>
<td></td>
</tr>
<tr>
<td>Gaps collapsed; codons</td>
<td>17.08.15 at 11:38</td>
<td>Xbfz7l</td>
<td></td>
</tr>
</tbody>
</table>
wasabiwebapp@gmail.com

Shared data from Wasabi
To:  Andres Veidenberg

To view the dataset that was shared with you, go to http://wasabiapp.org:8000?id=tiger

Wasabi is a web application for phylogenetic sequence analysis. You can learn more about Wasabi from http://wasabiapp.org

This message was sent to the e-mail address registered in Wasabi app.

You can edit your Wasabi account at http://wasabiapp.org:8000/G0CEXI
PUBLISHING
Wasabi: An Integrated Platform for Evolutionary Sequence Analysis and Data Visualization

Andres Veidenberg,¹ Alan Medlar,¹ and Ari Löytynoja*¹
¹Institute of Biotechnology, University of Helsinki, Helsinki, Finland
*Corresponding author: E-mail: ari.loytynoja@helsinki.fi.
Associate editor: Naoki Takebayashi

Abstract

Wasabi is an open source, web-based environment for evolutionary sequence analysis. Wasabi visualizes sequence data together with a phylogenetic tree within a modern, user-friendly interface: The interface hides extraneous options, supports context sensitive menus, drag-and-drop editing, and displays additional information, such as ancestral sequences, associated with specific tree nodes. The Wasabi environment supports reproducibility by automatically storing intermediate analysis steps and includes built-in functions to share data between users and publish analysis results. For computational analysis, Wasabi supports PRANK and PAGAN for phylogeny-aware alignment and alignment extension, along with drag-and-drop import of local files, Wasabi can access remote LYNX and EPO alignments directly from Ensembl. To demonstrate a few case studies, these case studies can be browsed within Wasabi at http://wasabiapp.org:8000 id=usecases. You can start using it at http://wasabiapp.org:8000.

Introduction

In evolutionary sequence analysis, phylogenetic trees and sequence alignments are intrinsically linked: Sequence alignments define character homologies upon which phylogenetic inference is based and trees are integral in correcting for hierarchical dependencies among sequence data. In multiple sequence alignment, this connection was noticed early (Sankoff 1975) and was central in the first progressive alignment algorithm (Hogeweg and Hesper 1984). Although popular alignment programs use a tree to guide the alignment procedure, they consider the tree a nuisance parameter. The

Wasabi is inspired by our earlier tool, webPRANK (Löytynoja and Goldman 2010), a web interface to the PRANK alignment program. webPRANK was limited to the analysis of only small data sets and to a single analysis method. Wasabi, however, was developed using modern web technologies to scale to much larger data sets, while maintaining responsiveness. Wasabi additionally contains features for integrating heterogeneous data sets, file management and an intuitive, modern web interface (see fig. 1 for overview) for navigating and selecting appropriate analysis options. Additionally, due to the increasing complexity of these workflows the curation of PWAs became cumbersome and an integration of data and analysis was needed in a single tool. Thus, the goal was to create a platform that would allow efficient and reproducible analysis and data visualization.
Hypoxia-related genes are soft targets in adaptation to new environments

John Doe¹ and Mary Major²

¹University of Cambridge, Harvard, USA; ²Trinity College, London, UK

Journal Communications 1, Article number: 1234 | doi:01.1001/jcomms1234
Received 02 December 2015 | Accepted 13 May 2016 | Published 02 June 2016

Abstract


Figure 1: Reanalysis of EGLN1 genetree in Wasabi.

Cho et al. (2013; see original study) detected a snow leopard-specific mutation in EGLN1 that is likely to be associated with adaptation to high altitude. We replicated the analysis with Wasabi and saw that the suggested position is not perfectly conserved across species and an identical Met39>Lys39 substitution has happened in alpaca, another species living in high altitudes. Browse the dataset in Wasabi pop-up.


In hendrerit nisi ipsum, nec iaculis dolor bibendum ac. Integer eget nibh vel massa tristique interdum id quis magna.
Gene: BRCA2  ENSG00000139618

Description
BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:11016]

Synonyms
XRCC11, FADC, FAD1, PNCA2, FAD, BRCC2, FANCD1, GLM3, BROVCA2, FANCD

Location
GRCh38:CM000675.2

About this gene
This gene has 7 transcripts (splice variants), 64 orthologues, is a member of 1 Ensembl protein family and is associated with 100 disease

Gene tree
GeneTree  ENSGT00390000003602

Number of genes  65
Number of speciation nodes  62
Number of duplication  2
Number of ambiguous  0
Number of gene split events  0

Ray-finned fishes: 10 homologs
Reptiles and birds: 7 homologs
Laurasiatherian mammals: 15 homologs
Old World monkeys: 3 homologs
BRCA2, Marmoset
BRCA2, Tarier
Wet nose lemurs: 2 homologs
BRCA2, Tree Shrew
Rodents and Rabbids: 8 homologs
Gene: BRCA2 ENSG00000139618

Description
BRCA2, DNA repair associated [Source:HGNC Symbol;Acc:HGNC:11016]

Synonyms
XRCC11, FACD, FAD1, PNCA2, FAD, BRCC2, FANCD1, GLM3, BROVCA2, FANCD

Location
Chromosome 13: 32.315.474-32.400.266 forward strand.
GRCh38:CM000675.2

About this gene
This gene has 7 transcripts (splice variants), 64 orthologues, is a member of 1 Ensembl protein family and is annotated by 63 gene names.

Gene tree
GeneTree ENSGT00390000003602

- Number of genes: 65
- Number of speciation nodes: 62
- Number of duplication: 2
- Number of ambiguous: 0
- Number of gene split events: 0

Ray-finned fishes: 10 homologs
Reptiles and birds: 7 homologs
Laurasiatherian mammals: 15 homologs
Old World monkeys: 3 homologs
BRCA2, Marmoset
BRCA2, Tarsier
PLUGINS
Executable

JSON
Executable

JSON

Wasabi interface
THANKS!

http://wasabiapp.org

Ari

andres.veidenberg@helsinki.fi

Wasabi: An Integrated Platform for Evolutionary Sequence Analysis and Data Visualization

Andres Veidenberg,1 Alan Medlar,1 and Ari Löytynoja*,1
1Institute of Biotechnology, University of Helsinki, Helsinki, Finland
*Corresponding author: E-mail: ari.loytynoja@helsinki.fi.
Associate editor: Naoki Takebayashi

Abstract

Wasabi is an open source, web-based environment for evolutionary sequence analysis. Wasabi visualizes sequence data together with a phylogenetic tree within a modern, user-friendly interface. The interface hides extraneous options, supports context sensitive menus, drag-and-drop editing, and displays additional information, such as ancestral sequences, associated with specific tree nodes. The Wasabi environment supports reproducibility by automatically storing intermediate analysis steps and includes built-in functions to share data between users and publish analysis results. For example, the paper provides a brief summary of the data used in the analysis and describes the methods used to obtain the results. The paper also presents a detailed description of the software and its capabilities, as well as examples of how it can be used in different research settings. The software is available for download from the website mentioned in the abstract.

Molecular Biology and Evolution

www.mbe.oxfordjournals.org