Jalview and friends ..
Sequence analysis for all

Jim Procter, Jalview Coordinator
j.procter@dundee.ac.uk | foreveremain
http://libreplanet.org/wiki/User:Riffraff
Analysis involves lots of different kinds of data:

Sequences, alignments, trees, structures, functional assays, literature.

Data, analysis, evidence and results must be properly recorded.

Fig. 1 Structural overview of mDNMT1(650–1602)–DNA 19-nucleotide oligomer complex with bound AdoHcy.

J Song et al. Science 2011;331:1036-1040
Bioinformatics data is not fun to read.....
What is Jalview?

- A java alignment viewer
  
  **Java alignment viewer**

  .. but not just for viewing..

- Java?
  - Platform independence
  - Standalone or web based tool

- Open source since ‘98
  - Coordinated in Dundee
  - Core development funded by ...
A brief history of Jalview

**Jalview Version 1**

- **1997**
  - Michele Clamp
    - *Interim Director of Research Computing, Harvard.*
  - James Cuff
    - *Cycle Computing, USA.*

**Jalview Version 2**

- **2005**
  - Andrew Waterhouse
    - *U. Basel.*
  - Jim Procter
  - Peter Troshin
  - VAMSAS eScience project

**2004**

- Jalview 1 published.
  - Steve Searle
    - *Now at Sanger, UK*
  - David Martin

**2009**

- Jalview 2 Paper
  - New BBSRC Funding 2009-2014
Jalview Desktop monthly usage from 2007-2012

Desktop launches per month:
- 2008: 144,868 downloads, 23,924 unique IPs
- 2009: 181,118 downloads, 32,537 unique IPs
- 2010: 191,321 downloads, 38,483 unique IPs
- 2011: 233,320 downloads, 48,563 unique IPs
- 2012: 275,539 downloads, 55,805 unique IPs

Unique IPs per month:
- 2008: 23,924
- 2009: 32,537
- 2010: 38,483
- 2011: 48,563
- 2012: 55,805
Alignment on Web

Jalview

Inkscape/Illustrator

LaTeX/LibreOffice/Word... Etc...

- Key qualities for a workbench
  - Undo and Redo!
  - Archival of data, results and all display settings
  - Filtering & Multiple views
  - Easy access to databases and programs
Distributed Annotation System

Analysis

Pairwise alignment

Consensus Conservation & Clustering

Shading

Interactive Editing

Alignment

Visualisation

Pairwise alignment

Sequences

Features

Annotation

Trees

Trees/PCA

Alignments

Structures

PCA

Clickability

HTML

Line Art

Images
Jalview and friends ..

Sequence analysis for all
Jalview and friends ..
Sequence analysis for all

VARNA
Visual Analysis of RNA

http://varna.lri.fr/

http://jmol.sourceforge.net/
Requires

- ant
- icedtea-netx-common
- java-headers (>= 0.1.22)
- javahelp2
- jmol-applet
- libapache-mime4j-java
- libaxis-java
- libcastor-core-java
- libcastor-xml-java
- libcommons-discovery-java
- libcommons-logging-java
- libhttpclient-java
- libhttpcore-java
- libhttpmime-java
- libjaba-client-java
- libjmol-java (>= 12.1.42)
- libjswingreader-java
- liblog4j1.2-java
- libmiglayout-java
- libvamsas-client-java
- openjdk-6-jdk
- openjdk-6-jre | cacao-oj6-jre | java6-runtime

Packaged for Debian
by Vincent Fourmand

With help from all
the folk in DebianMed:

... particularly..

Steffen
Moller
JDAS is a Java DAS client library which simplifies DAS queries and maps Java classes to DAS XML representations. JDAS includes JAVA classes generated from DAS RNG schemas. These Java classes are used through the JAXB API to store and retrieve data in memory without the need to implement a specific set of XML loading and saving routines. JAXB provides two main features: the ability to marshal Java objects into XML and the inverse, to unmarshal XML into Java objects. JDAS additionally implements adaptor classes which are used to refactor generated classes with a complex structure. Finally JDAS implements client classes to facilitate querying DAS information.

Developed by Rafael Jiminez, Jonathan Warren and other Java developers in DAS Community
Bob Hanson awarded Blue Obelisk award for Jmol in 2006

Image taken at the Panton Arms, in Cambridge UK
http://blogs.ch.cam.ac.uk/pmr/2006/09/14/blue-obelisk-award-bob-hanson-of-jmol/
VARNA is developed by Yann Ponty CNRS Fellow at LIX in France
Lauren Lui, UC Santa Cruz.
http://jalview-rnasupport.blogspot.com/

Purine/pyrimidine colourscheme

Rfam alignment fetcher

WUSS annotation parser (from RALEE)

NESCent

SUMMER OF CODE // 2010
code.google.com/soc
Jan Engelhardt (Uni. Leipzig)

VARNA: Visualization Applet for RNA
A Java lightweight component and applet for drawing the RNA secondary structure

H.sapiens.11/1-1 U A C C U G U U C A G U U G U U A C A - G U A
H.sapiens.10/1-9 U A U C U G U U C A G U U G U U A C A - G U A
P.abelii.3 Selection H.sapiens.10 Structure
P.abelii.2 Hide Sequences
P.abelii.1

E.coli.11/1-10 H.A.G
C u G u U C a g u u G u a C a . G u a
C U G U U C A G u U G U A C A G U A

Google SUMMER OF CODE 2011

NESCent
10% of Summer of Code orgs in 2012 create & maintain scientific software
Science Code Manifesto

Manifesto  Discussion  Endorse  Resources  About

Software is a cornerstone of science. Without software, twenty-first century science would be impossible. Without better software, science cannot progress.

But the culture and institutions of science have not yet adjusted to this reality. We need to reform them to address this challenge, by adopting these five principles:

Code
All source code written specifically to process data for a published paper must be available to the reviewers and readers of the paper.

Copyright
The copyright ownership and license of any released source code must be clearly stated.

Citation
Researchers who use or adapt science source code in their research must credit the code's creators in resulting publications.

Credit
Software contributions must be included in systems of scientific assessment, credit, and recognition.

Curation
Source code must remain available, linked to related materials, for the useful lifetime of the publication.

iPython is a perfect example!

Panton Arms, again…

Further Reading
- Climate Code Foundation
- Panton Principles
- Open Knowledge Definition

CCF
The Jalview 5 Year Plan

1 Oct 2009 to 30th Sep 2014

More Flexible

More Capable

Large Datasets

Sustainable

Maintainable

Extensible

Community

GUI refactor & plugins

Software Engineering

Analysis services
The Jalview 5 Year Plan

More Flexible

More Capable

Community Outreach

- Website redesign
- Training
  - Users and Developers
- Internationalisation
- True Open Source Development
  - Issue tracker and open repository

Software
Engineering

GUI refactor & plugins

Frontend

Large Datasets

Community
Users and Developers

Analysis services

30th Sep 2014

1 Oct 2009
Thank you.