LibrAlign — A powerful Java GUI library for MSA and attached raw and meta data

Why developing LibrAlign?
Several applications currently developed in our group and by cooperating colleagues deal with multiple sequence alignments (MSA) or associated raw and meta data (e.g. trace files, positions of repeats or inversions,…), and allow the user to view and edit them in a graphical user interface (GUI). Instead of implementing independent solutions for these different tasks, we decided to create a common library that contains powerful and reusable GUI components and is easily extendable by others.

Features and design
- MultipleAlignmentsContainer in one of the main GUI components of LibrAlign and contains a set of AlignmentAreas
- An AlignmentArea can be used inside such a container or individually
- It contains the aligned sequences and allows to edit these (as well as a set of Database components displaying attached data)
- Sequence data can come from any data source by an implementation of the Interface SequenceDataProvider (as well as a set of Database components displaying attached data)
- LibrAlign contains several default implementations that provide compatibility to e.g. BioJava or allow a compressed storage of very large data sets.
- The GUI components allow the user to easily edit the alignment (as well as a set of Database components displaying attached data).
- The parent class DatabaseArea can also be extended by application developers to create new data areas (see @ and example in fig. 5).

Software using LibrAlign
Currently there are a number of software projects available or under development which make use of the LibrAlign API in their current or future versions.
- Taxonomic editor of the EDIT platform: The EDIT platform is currently extended to support sequence data for the Campusana portal of EDIT using LibrAlign (see fig. 2, http://cybertaxonomy.eu/taxeditor/)
- AlignmentComparator: Visualizes differences between alternative alignments of the same data set and allows annotating the comparison (as well as a set of Database components displaying attached data).
- PhyDE: Phylogenetic data editor: An alignment editor that provides features especially useful for phylogenetic purposes (as well as a set of Database components displaying attached data).
- HIIR-Finder: Locates microstructural mutations like tandem repeats possibly associated with hairpins and visualizes these in a special data area (as well as a set of Database components displaying attached data).

Compatibility and IO
- All components have a Swing and a SWT version, so they can be integrated into nearly all thinkable kinds of Java GUIs, including more complex projects based on the Eclipse Rich Client Platform (http://www.eclipse.org/home/)
- The toolkit independence is made possible by the TCL library of bioinfweb.commons (http://commons.bioinfweb.info/java/)
- LibrAlign is fully interoperable with the BioJava 3 API (Pliéz et al., 2012) and also contains bridges to BioJava 1.x legacy release.
- NeXML (Vos et al., 2012) is the main format of LibrAlign, since its flexible meta-data annotation allows to store data associated with the data areas. (The definition of according ontologies for the different data areas is currently ongoing.)
- Several additional alignment formats are supported for im- and export.

Acknowledgements
The funding of parts of the development of LibrAlign with grant MU 2875/3–5 to BKM by the DFG (German research foundation) is highly appreciated. BCS wants to thank the KNF (National Science Foundation) and IPAM for partly financing the presentation of this poster at the MSA Workshop at IPAM, UCLA 2015.

Furthermore the authors are very thankful to the developers of the other open source projects LibrAlign is build on: Biokatana, Apache commons, NeXML, especially to Roger Vos for his advice on the NeXML integration. A special thank you is also due to Joaquin Rischkebloch for providing demonstration data and to Ruth Hendriks for very helpful comments on this poster.

Citations

