Poster H-72

MAVIANT – a Multi-purpose Alignment VIewing and (SNP) Annotation Tool

Authors:
Panitz, F (*Dept. Genetics and Biotechnology, Danish Institute of Agricultural Sciences*)
Stengaard, H (*Dept. Genetics and Biotechnology, Danish Institute of Agricultural Sciences*)
Bendixen, C (*Dept. Genetics and Biotechnology, Danish Institute of Agricultural Sciences*)

Short Abstract: MAVIANT, a platform-independent Multi-purpose Alignment VIewing and (SNP) Annotation Tool, was designed for large-scale SNP mining projects. In addition to chromatogram and alignment views MAVIANT facilitates manual annotation, which is stored directly in an integrated database. The annotations are immediately accessible and can be easily shared with external collaborators.

Long Abstract:
The increasing availability of sequence trace data in public repositories makes it feasible to analyse sequences alignments and in-silico SNP predictions on the chromatogram level. Standard programs as Cap3, Phrap, PolyBayes and PolyPhred generate ace files that can locally be viewed using the UNIX-based Consed program. However, collaborating with (external) partners on large sets of contigs would require transferring trace/quality files or the complete analysis output, provided a UNIX based system for subsequent analysis and annotation is available. Our motivation was to provide easier, platform-independent access to trace data, contig information and options for annotation. We therefore devised MAVIANT, an application that provides contig and chromatogram views together with an integrated annotation tool in a platform-independent browser. The current version was designed with respect to cooperation on SNP mining projects using predictions based on contig clusters originating from a few hundred thousand sequences. MAVIANT output is generated using html, png and javascript files thus making it independent from the data source. The views can be accessed via a web-server or as stand-alone application on a local computer. By pregenerating the output files the application is optimized for fast loading speed. From a navigation panel chromatograms in the alignment views can be collapsed and expanded in order to provide overviews and quick navigation. As input chromatogram files (ab1, abd, ab1 trace files), base quality information (phd files) and clustering and alignment information (ace files) are used. Alignment overviews of sequences and features are available in three different modes: data (consensus and sequence overview), filter (overview of differences between consensus and sequences) and quality (phd-based quality overview). Alignments views are depicted as 50 base pair fragments in three different colouring schemes for easier visualisation: data (coloured base), background (coloured base background) and quality. Sequences can be annotated by BLAST hits in order to provide search options as gene id or description. Additional features like SNPs, repeats or other custom annotation can be added by using a feature file containing sequence identifier, annotation type (SNP, repeat, etc.), start and end position and a specific colour for each individual feature. Apart from being a mere viewing tool for chromatograms and alignments MAVIANT is designed to facilitate manual annotation, e.g. the assessment of predicted polymorphic sites. These individual evaluations are stored directly in an integrated flat file or MySQL database using a dynamic

website. The annotations are immediately accessible and can be easily shared with external collaborators. Additional features (e.g. the integration of the scf sequence format as well as other alignment formats like BLASTN, SSAHA, Dialign, ClustalW) are currently under development.