

NeoPipe is an open-source framework, distributed under the LGPL license. The source code is available at <http://sourceforge.net/projects/neopipealign/>

NeoPipe: A workflow for protein family analysis

Vincent Walter, Julie Thompson, Olivier Poch and Hoan Nguyen

Contact: nguyen@igbmc.fr

Integrated structural Biology department, (IGBMC), Illkirch, France

Integrative Genomics and Bioinformatic Laboratory-LBGI, Strasbourg, France

New advanced technologies including the next-generation DNA, and information technology have significantly improved our capacities of developing biological knowledge and changing our understanding of diseases, phenotype and genotype. In this post-genomic context, protein sequence analysis is a key issue to better understand the evolutionary, structural and functional aspects. NeoPipe is a tool of analyzing a protein family, which consists of 8 steps concerning the search for homologous sequences in multiple databases (protein, 3D structures,...) and functional and structural annotations of clustered multiple alignment of complete sequences (MACS). Those indicate the relationship between the protein subfamilies. The emphasis is to get a high quality alignment by performing refinement and corrections (evaluated by a quality score at each step) and giving a clustered and annotated alignment of potential subgroups. NeoPipe's application, APIs and REST Web Services are implemented in Java and supported on Linux. NeoPipe is open-source (under the LGPL license) and the source code is available on SourceForge at <http://sourceforge.net/projects/neopipealign/> using a GIT repository. NeoPipe website is developed in Java, JavaScript (jQuery), AJAX with all major browser supported. The website is available at <http://lbgi.fr/neopipe/>.