

Reliable annotation services

Biomax Informatics AG provides comprehensive manual annotation of entire genomes, including those from eukaryotic organisms. This highly specialized service is carried out by a team of more than 20 biologists with doctorate degrees. The Biomax annotation team has annotated several genomes for clients worldwide and represents one of the most skilled and experienced teams available.

High-quality results

Biomax offers the following levels of annotation:

- Automatic
- Basic manual
- Intermediate manual
- Advanced manual

The Pedant-Pro™ Sequence Analysis Suite provides the automatic annotation results, which become a framework for further annotation. Our annotators provide expert manual verification and annotation to improve the quality of the automatically generated data. The annotation results are delivered as a relational database supporting MySQL™ and Oracle® database management systems.

High-quality annotation allows researchers to perform the following tasks:

- Efficiently extract further biologically significant information
- Perform functional and structural analyses more effectively
- Optimize the design of experiments to answer questions about molecular and genetic aspects
- Narrow down options for drug targets or regions that should be further characterized and studied

Biomax and Pedant-Pro are registered trademarks and FunCat is a trademark of Biomax Informatics AG in Germany and other countries. MySQL is a trademark of MySQL AB. Oracle is a registered trademark of Oracle Corporation. Registered names, trademarks, etc., used in this document, even when not specifically marked as such, are not to be considered unprotected by law. PPR0309BEGAN

Automatic annotation

The automatic level of annotation involves gene prediction using the Pedant-Pro Sequence Analysis Suite. Open reading frames (ORFs) for prokaryotes are predicted by the ORPHEUS algorithm, while eukaryotic gene models are predicted by FGENESH and other algorithms. Both algorithms self-adapt by optimizing the parameter set for each organism, a process called training.

Automatic analysis and annotation predicts ORFs, tRNAs, rRNAs and other genomic features, and automatically annotates these features by the following:

- Searches for similarity using up-to-date protein and DNA databases
- Classification of function
- Identification of protein families
- Identification of protein domains including PROSITE, COGs and sequence blocks
- Identification of Enzyme Commission (EC) numbers
- Identification of Protein Information Resource (PIR) keywords
- Prediction of structure including SCOP domains, known three-dimensional structure and low-complexity regions

Basic manual annotation

Basic manual annotation improves the automatically annotated data by removing ORFs that overlap with genetic elements. The Biomax annotation team will perform the following tasks:

- Identification of internal and overlapping ORFs or predicted genes
- Analysis of intergenic regions to identify missed ORFs or genes
- Verification of exon borders and start and stop codons of each ORF or predicted gene

**Contact Biomax for more information
about genome annotation**



Intermediate manual annotation

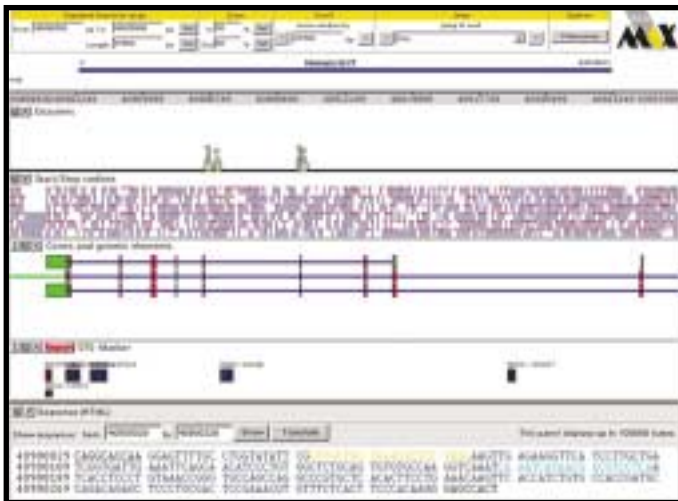
For the intermediate level of manual annotation, Biomax scientists provide the following additional information:

- Title assignment
- Functional classification using the FunCat™ Functional Catalog developed by MIPS* and Biomax
- ORF classification according to homology and including the following descriptions:
 - Known protein
 - Strong similarity to known protein
 - Weak similarity to known protein
 - Similarity to unknown ORF
 - Unknown ORF
 - Questionable ORF

Advanced manual annotation

Advanced manual annotation includes the following additional steps:

- Verification of all predicted features
- Annotation of EC numbers which were not automatically predicted
- Identification of literature references
- Identification of additional information about specific ORFs (for example, catalytic activity, function and pathways)



The Pedant-Pro DNA Viewer allows you to visualize the annotation results

FunCat Functional Catalog

Biomax intermediate and advanced levels of manual annotation use the FunCat Functional Catalog developed at MIPS and Biomax. The catalog provides access to a broad range of biological information through a hierarchically organized, controlled vocabulary. The FunCat system facilitates the annotation process:

- Groups biochemical functions and processes, expression data, and protein localization information according to the involved biological processes
- Provides intuitive navigation due to its hierarchical organization
- Allows biological information to be described in an organism-independent context
- Enables efficient data integration, manual and automatic information retrieval, and multidimensional annotation
- Provides consistency checks for data integrity and completeness
- Smooth connections between data sources
- Allows information to be integrated into and transferred between different bioinformatics applications

Annotation expertise

Many members of the annotation team gained experience at MIPS before joining Biomax. At MIPS, Biomax scientists worked on projects including the first comprehensive annotation of a eukaryotic genome (*Saccharomyces cerevisiae*) and the first complete annotation of the plant genome (*Arabidopsis thaliana*). *Nature* published the results of the *S. cerevisiae* genome in May 1997 and the *A. thaliana* genome in December 2000.

*The former Munich Information Center for Protein Sequences (MIPS) is now the Institute for Bioinformatics (IBI).

Selected references

Mewes HW, Albermann K, Bahr M, Frishman D, Gleissner A, Hani J, Heumann K, Kleine K, Maierl A, Oliver SG, Pfeiffer F and Zollner A (1997) Overview of the yeast genome. *Nature* 387:7–65

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815