



annoscoreTM
automated genome and functional annotation

Accelerate your next annotation projects

with our automated, scalable pipeline

high throughput

fully automated pipeline

scalable, flexible and adaptable



Benefit from our personalized service and support

At Computomics, we customize our gene finding pipeline to your species. We perform functional and structural de novo annotation using machine learning technologies. We make use of all available data, including existing annotations, RNAseq reads, ESTs, PacBio complete isoform sequences, and others.

Structural Annotation:

We use the latest advances in gene prediction and annotation technologies. Our machine learning-based algorithms deliver the most accurate gene models and will speed up your product pipeline with better annotations. In your gene finding project we collaboratively ascertain the project objectives to ensure that a quality structural annotation is developed. The results can then be utilized effectively by the client.

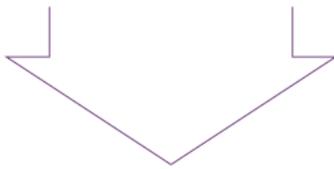
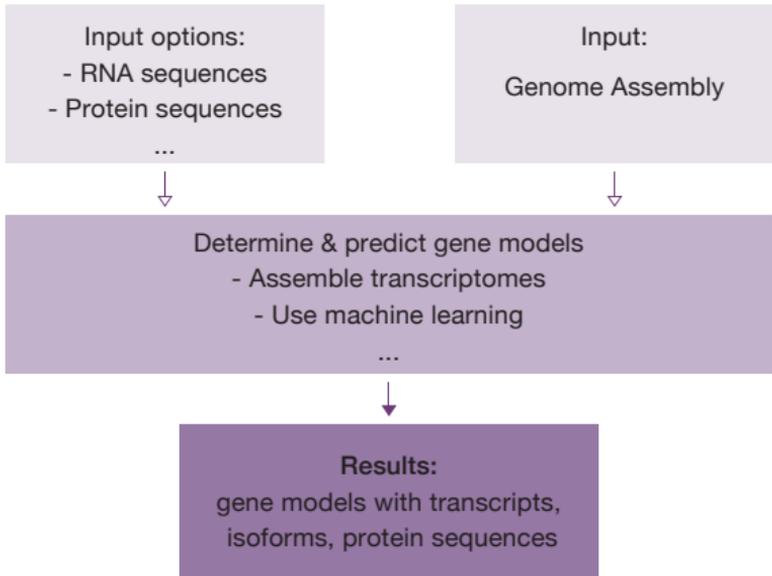
Functional Annotation:

AnnoScore is our fully automated, scalable and adaptable functional annotation solution. It includes mappings to metabolic pathways, gene ontologies and protein domain databases. With AnnoScore we have a ready to go solution that performs functional annotation and can be independently run.

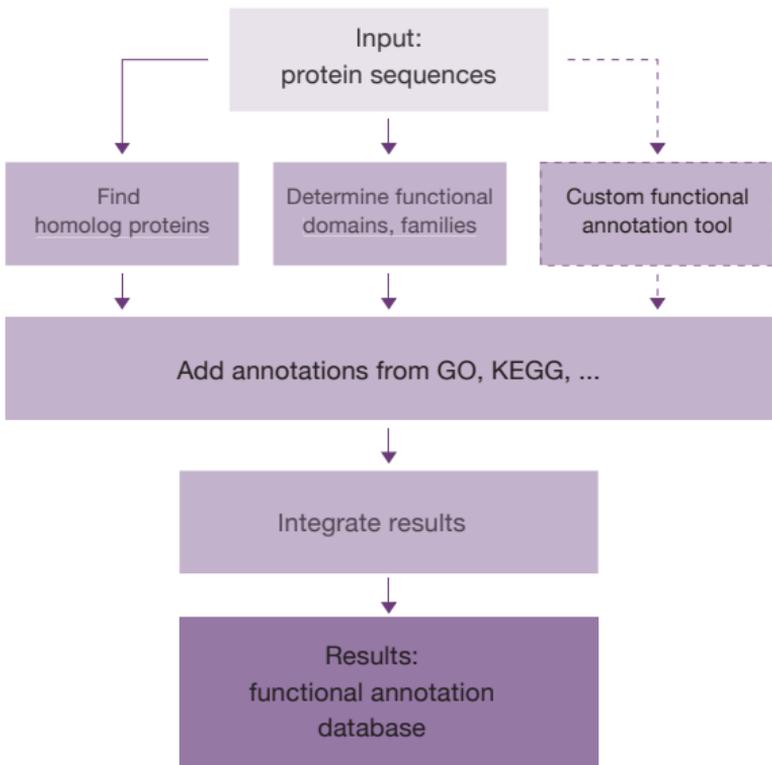


GENOME ANNOTATION WORKFLOW

Structural annotation / Gene finding (project specific)



AnnoScore: Functional annotation





Speed up your product pipeline with automated annotations

- ▶ *Fully automated processing of hundreds of thousands of protein annotations at once*
- ▶ *Easily customizable and scalable to customer needs*
- ▶ *Results are presented in an intuitive way and are easily accessible through database query or via a genome browser*
- ▶ *Using diamond blastp for alignment can speed up the process by a factor of thousands*
- ▶ *In-house developed reciprocal best hit allows to find not only one reciprocal best hit, but also similarly good reciprocal hits*
- ▶ *Runs on any server infrastructure and is high performance computing ready*
- ▶ *Results are stored in an integrated database of the client's choice*
- ▶ *Highest data security as all data stays in house*



Highly scalable and parallelizable functional annotation pipeline

- ▶ *Compatibility with common HPC environments, e.g. SGE, LSF, ...*
- ▶ *Maximum speedup in HPC/cloud setups:*
 - *Feasible parallelization on up to 40,000 CPUs*
 - *Typical runtime 5-7 hours on HPC cluster for 100,000 cDNAs*
- ▶ *Transfer annotations and cross-references from Gene Ontology (GO), KEGG, Reactome, Uniprot, NCBI Gene, Ensembl*
- ▶ *Find matches to InterPro databases: CDD, Coils, Gene3D, Hamap, MobiDBLite, PANTHER, Pfam, PIRSF, PRINTS, ProSitePatterns, ProSiteProfiles, SFLD, SMART, SUPERFAMILY, TIGRFAM*
- ▶ *Output format adaptable to your preferred type or database design, e.g. SQL, TSV*



Functional annotation of proteomes and DNA sequences

Challenge:

Our customer wanted to automatically submit thousands of annotation queries per day, collect the results in its own database and make them available via a genome browser. A commercial solution used before was technically out of date and didn't provide the necessary accuracy, flexibility and throughput.

Solution:

Using AnnoScore, our customer is now able to annotate a complete set of gene sequences from a genome in one run, using a multitude of annotation tools and query-relevant databases. AnnoScore uses parallelization extensively, which allows hundreds of thousands of annotation queries to be performed daily. The results are then stored within a database formatted to our client's needs.

Benefits:

With AnnoScore our customer is now able to process hundreds of thousands of protein annotations a day automatically which decreased the manual overhead to a minimum and sped up the process by a factor of hundred. The plethora of complex results is presented in an intuitive way and is accessible automated and in bulk via a database or for in-depth inspection through a genome browser.

Customer Quote

“AnnoScore solved all the issues we had with functional annotations. The automation and parallelization made the hassle of it finally easy and fast.”

Midrange Breeding Company



computomics
machine learning-based data analysis

Computomics GmbH

Phone: +49 7071 568 3995

info@computomics.com



@computomics



/computomics-gmbh



/computomics