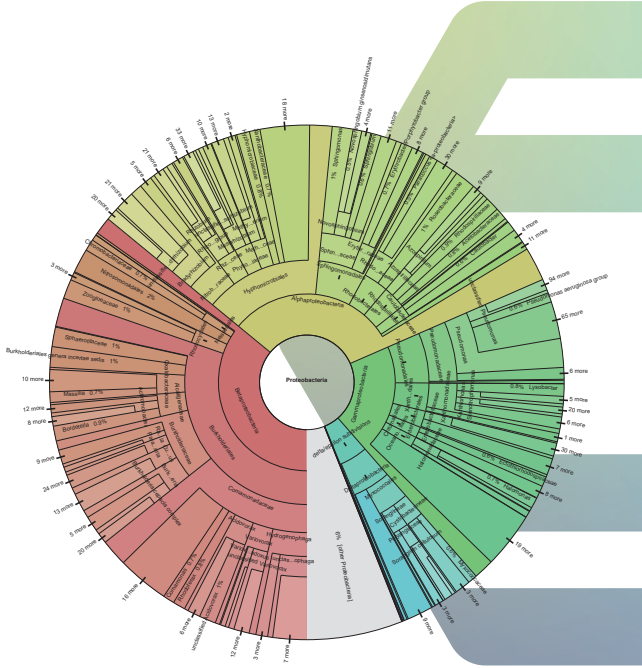

OmicsBox
Bioinformatics Made Easy
biobam
BIOINFORMATICS SOLUTIONS

OmicsBox Overview

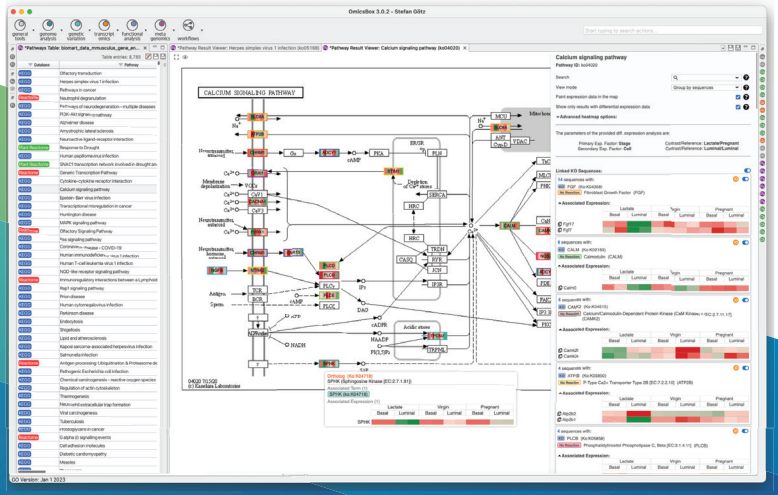


OmicsBox is a leading bioinformatics solution that offers end-to-end NGS data analysis of genomes, transcriptomes, and metagenomes.

The application is designed to be user-friendly, efficient, and with a powerful set of tools to extract biological insights from omic data.

OmicsBox is used by top private and public research institutions worldwide and allows researchers to easily process large and complex data sets, and streamline their analysis process.

The software is structured in different modules, each with a specific set of tools and functions designed to perform different types of analysis, such as de-novo genome assemblies, genetic variation analysis, differential expression analysis, and taxonomic classifications of microbiome data, including the functional interpretation and rich visualizations of results. The functional analysis module, which includes the popular Blast2GO annotation methodology makes OmicsBox particularly suited for non-model organism research. This is demonstrated by over 15k scientific research citations. OmicsBox works out of the box on any standard PC or laptop with Windows, Linux, or Mac.



OmicsBox Modules

The leading bioinformatics platform for the analysis of omics data.

Genome Analysis

Quality Control And Assessment
Genome Mapping and Assembly
Repeat Masking and Gene Finding
Genome Curation Tools

Transcriptomics

De-Novo Assembly and Alignment
Differential Expression Analysis
Single-Cell RNA-Seq
Long Read Transcriptomics

Metagenomics

Metagenomic Assembly
Taxonomic Classification
Functional Annotation and Analysis
Comparative Analysis

Genetic Variation

Fast Variant Calling and Filtering
Supports GBS and WGS data
Variant Annotation
Genome Wide Association Studies

Functional Analysis

Basic Sequence Analysis
Blast2GO Functional Annotation
GSEA and Fisher's Enrichment Analysis
Pathway Analysis with KEGG

General Features

Manage Projects and Files
Access to Cloud Computing
Design, Run and Save Workflow
Visualize Data, Genome Browser

Efficient

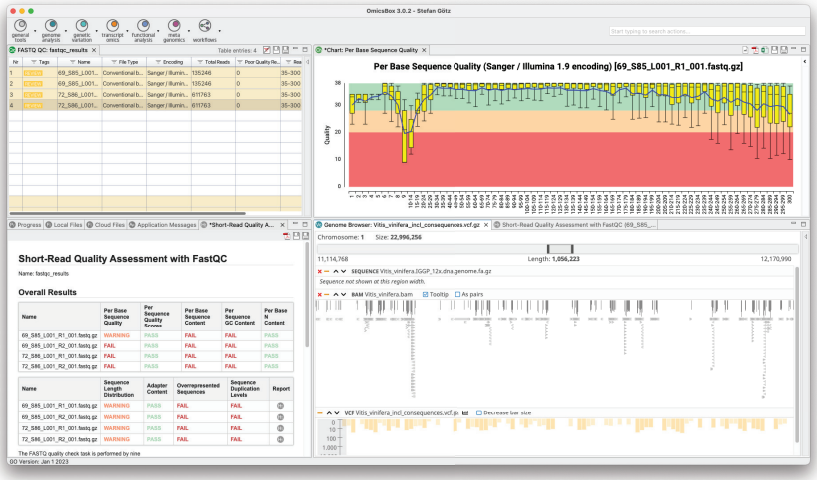
Powerful

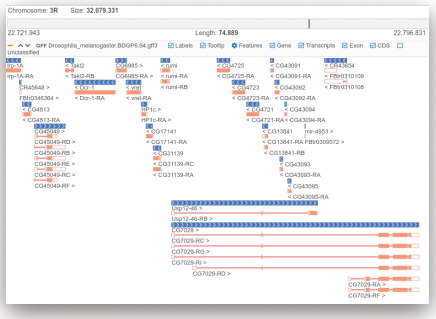
Versatile

Genome Analysis

Convert raw DNA-Seq reads into a structurally annotated and curated draft genome.

- Quality Control And Assessment
- De-Novo Assembly
- Alignment and Polishing
- Long Reads Genome Analysis
- Repeat Masking and Gene Finding





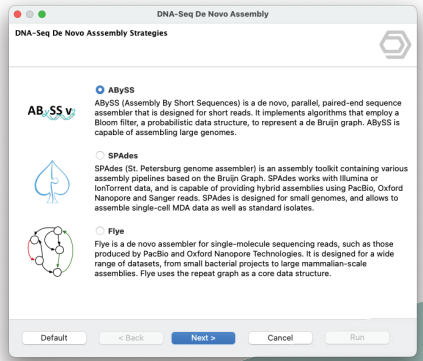
Complete Genome Analysis for Short and Long Reads

The Genome Analysis module allows for the characterization and analysis of newly sequenced genomes, from raw reads to gene structures in an efficient and user-friendly way. Quality control of the data can be performed using FastQC and Trimmomatic, allowing for the filtering of reads and removal of low-quality bases.

The assembly feature allows for the reconstruction of whole genome sequences without a reference genome or any specific hardware requirements. The module offers the ability to assemble sequencing data from both **short and long-read technologies** using popular algorithms such as **ABySS, SPAdes, and Flye.** Flye can also be used in combination with **Pilon** for sequence polishing.

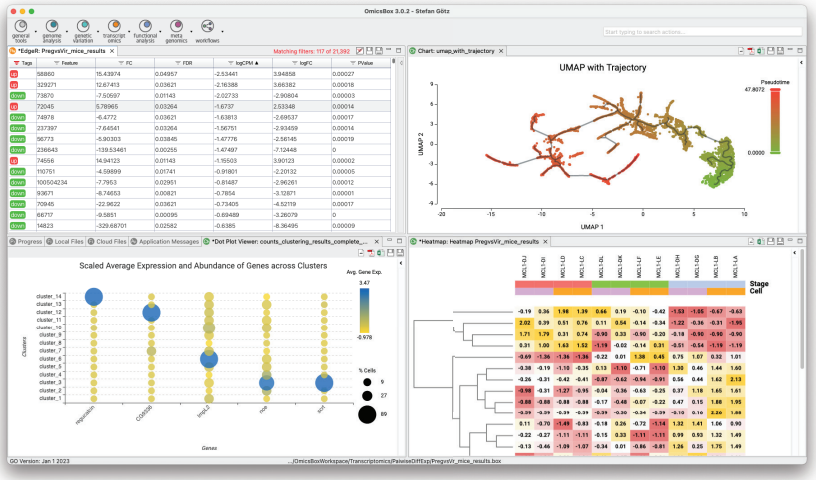
The module also provides efficient execution of **BWA and Bowtie2** for read alignments. Additional tools for repeat identification and masking prior to gene prediction are also available.

Finally, OmicsBox allows visualization of your annotations in the form of tracks that combine genome sequences (.fasta) with alignments (.bam), intron-exon structure (.gff), and variant data (.vcf).



Convert your RNA-Seq Samples into Biological Insights.

Transcriptomics

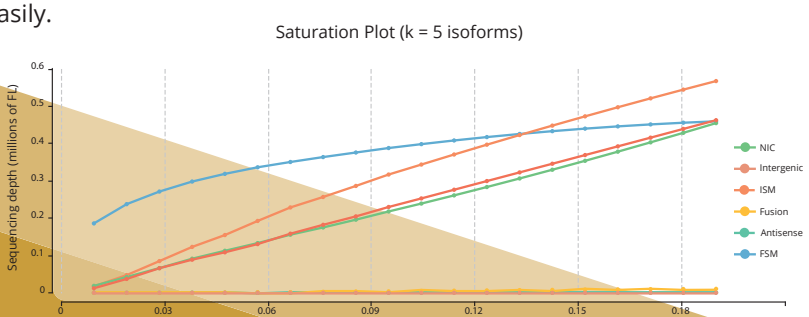


- De-Novo Assembly ✓
- RNA-Seq Alignment ✓
- Differential Expression Analysis ✓
- Enrichment Analysis ✓
- Single-Cell RNA-Seq ✓

Quality Control, Assembly, Quantification and Differential Expression

The Transcriptomics module allows for the processing of RNA-Seq data from raw reads down to the functional analysis in a flexible and intuitive way. Once quality control has been applied using FastQC and Trimmomatic, RNA-seq data can be aligned to a reference genome using **STAR or BWA,** or assembled without a reference genome using **Trinity** to obtain a de-novo transcriptome. Additional tools for predicting coding regions and assessing coding potentials aid in refining the datasets.

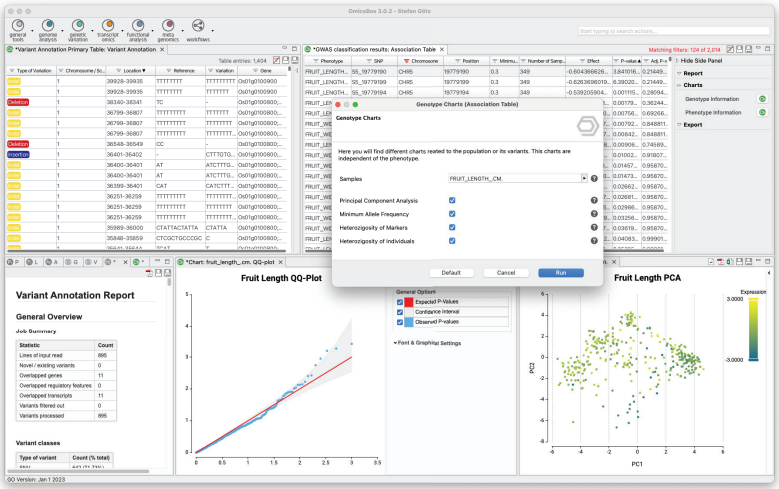
The module also allows quantification of expression at the gene or transcript level - with or without a reference genome, using **HTSeq or RSEM.** Different statistical charts provide information about the assembly and quantification processes, as well as a quality assessment of the results. Differentially expressed genes can be detected between experimental conditions or over time using statistical packages like NOISeq, edgeR, or maSigPro. Interactive heatmaps help to interpret results intuitively. The combination of IsoSeq3 and SQANTI3 allows for the identification and characterization of long-read-sequenced transcripts. The available **Single-Cell RNA-Seq** tools help identify groups of cells and gain insight into cell lineage trajectories in pseudotime. Enrichment analysis is directly integrated to identify over- and underrepresented biological functions easily.



Genetic Variation

Identify and analyze genetic variations within a population or a species.

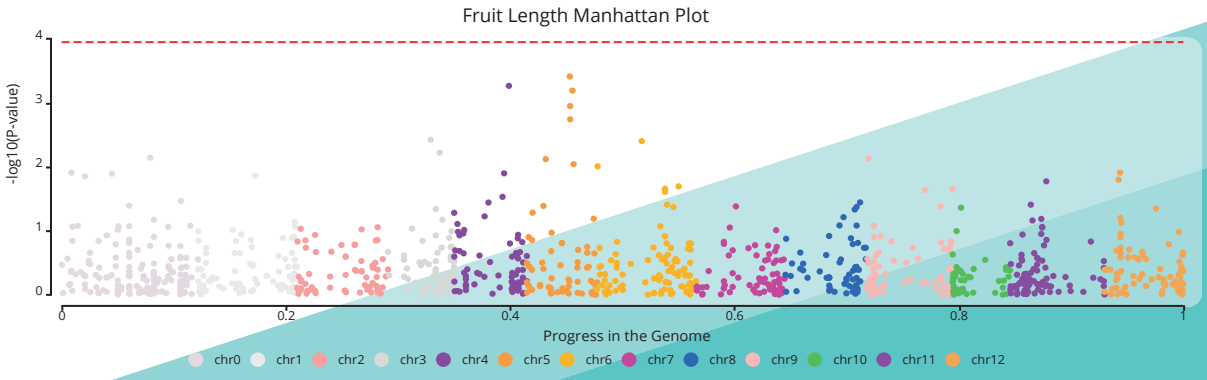
- Fast Variant Calling
- Variant Filtering
- Model & Non-Model Variant Annotation
- Guided Genome Wide Association Studies
- Supports GBS and WGS data



Cutting-Edge Genetic Variation Analysis

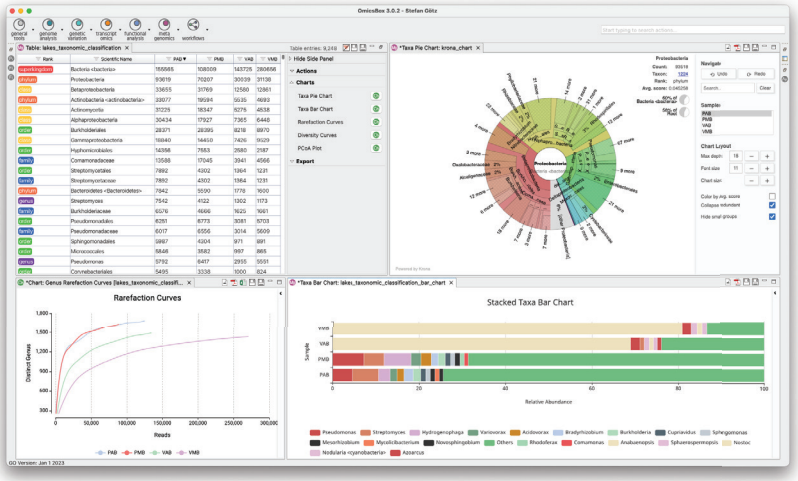
The Genetic Variation Module allows for performing variant calling, filtering, and annotation, as well as associating genetic variations with a particular trait or disease via genome-wide association studies.

The module offers two different analysis strategies for variant calling and filtering based on the popular tools, **BCFtools and FreeBayes.** Many options are available to adjust the analysis to specific methodologies (e.g GBS vs WGS) or requirements (level of ploidy, genome coverage, presence of repetitive regions, etc.). The resulting VCF file can then be annotated using the Variant Effect Predictor from Ensembl. This combination of tools has been shown to outperform alternative pipelines in several recent review studies (e.g. Nature Sci. Rep. 12, 11331 (2022)). Finally, the module also allows for the execution of **guided genome-wide association studies (**GWAS) to identify genetic variations associated with a particular trait.



Metagenomics

Allows different types of microbiome analysis like taxonomic classification and assembly with functional annotation.



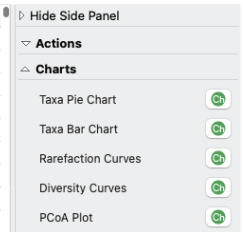
- Quality Control ✓
- Metagenomic Assembly ✓
- Taxonomic Classification ✓
- Functional Annotation and Analysis ✓
- Comparative Analysis ✓

Complete WGS and 16S Microbiome Data Analysis

The Metagenomics Module enables the execution of different types of microbiome data analysis, including the assembly, annotation, and classification of metagenomic data. It allows to combine and integrate all necessary steps for a **complete microbiome data analysis** in a flexible and intuitive way and offers custom pipelines for individual analysis strategies.

The module offers **Kraken** to identify bacteria, archaea, fungi, protozoa, and viruses down to strain levels. Rich visualizations help to gain insights more easily. With the tools MetaSPAdes and MEGAHIT it is possible to assemble large datasets easily and fast in the cloud. Additionally, the use of FragGeneScan for plain reads and Prodigal for assembled data allows to identify and extract possible genes and proteins. Finally, functional annotation with various databases is available and various statistical tests can identify the **differential abundance** of taxonomies and biological functions between samples.

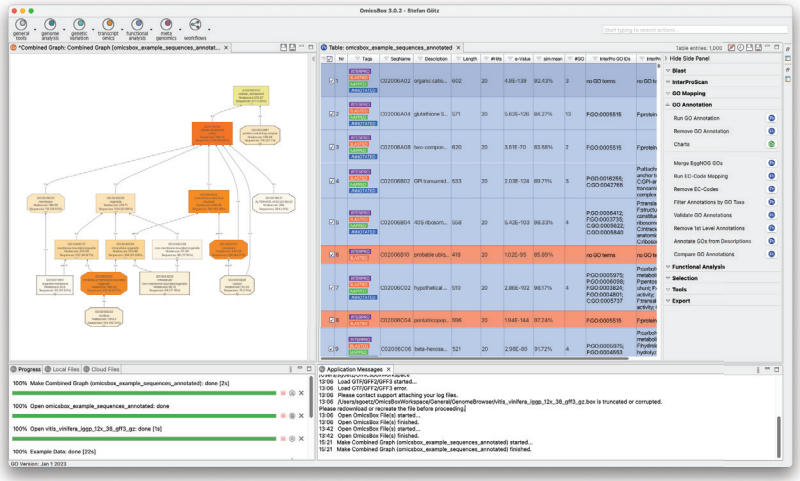
Rank	Scientific Name	PAB ▼	PMB	VAB	VMB
superkingdom	Bacteria <bacteria>	155565	108009	143725	280656
phylum	Proteobacteria	93619	70207	30039	31138
class	Betaproteobacteria	33655	31769	12580	12861
phylum	Actinobacteria <actinobacteria>	33077	19594	5535	4693
class	Actinomycetia	31225	18347	5275	4538
class	Alphaproteobacteria	30434	17927	7365	6448
order	Burkholderiales	28371	28395	8218	8970
class	Gammaproteobacteria	18840	14450	7426	9529
order	Hyphomicrobiales	14356	7553	2580	2187



Functional Analysis

Functionally annotate and analyze any sequence dataset from scratch.

- Sequence Analysis with Blast, InterPro and EggNog
- Blast2GO Functional Annotation
- GSEA and Fisher's Enrichment Analysis
- Combined Pathway Analysis with KEGG
- Rich Gene Ontology Visualizations



High-Quality Functional Annotation and Enriched Pathway Analysis

The Functional Analysis module allows adding **biological context** to the different data analysis approaches in OmicsBox.

An initial step is the functional annotation of any given dataset. The popular **Blast2GO** methodology allows flexibility to assign the most reliable functional labels to novel sequence datasets. The annotation process takes into account source annotation quality and Gene Ontology hierarchies. Cloud-based Blast and InterPro allow for fast sequence alignment and domain searches against the reference datasets of your choice. Potential homologs and domains are linked to functional annotation from up-to-date and well-curated databases like UniProt and the Gene Ontology Consortia.

Once a dataset is functionally annotated, the module offers different **enrichment** approaches like the Fisher's Exact Test and GSEA to identify over and under-represented biological functions.

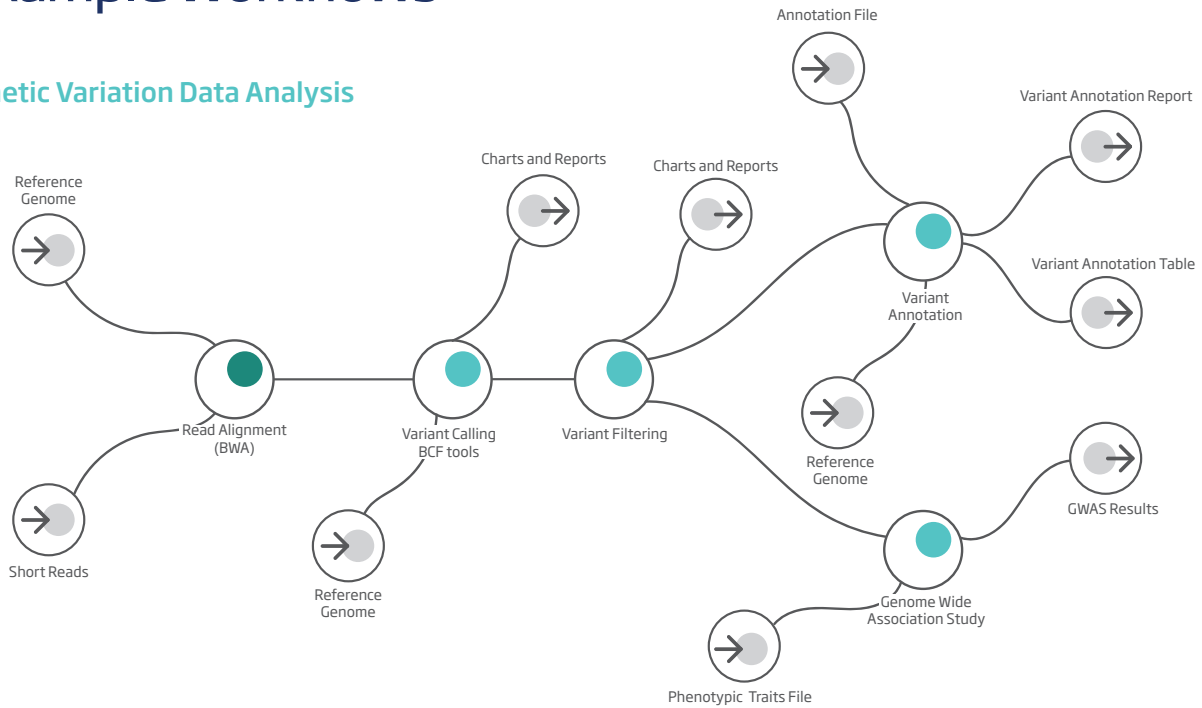
Finally, the **Combined Pathway Analysis** allows the identification of Reactome and **KEGG** pathways for any set of sequences. Combined with differential expression data, the tool allows calculation pathway enrichment and offers rich visualizations to gain insights with ease.

Fast Functional Annotation

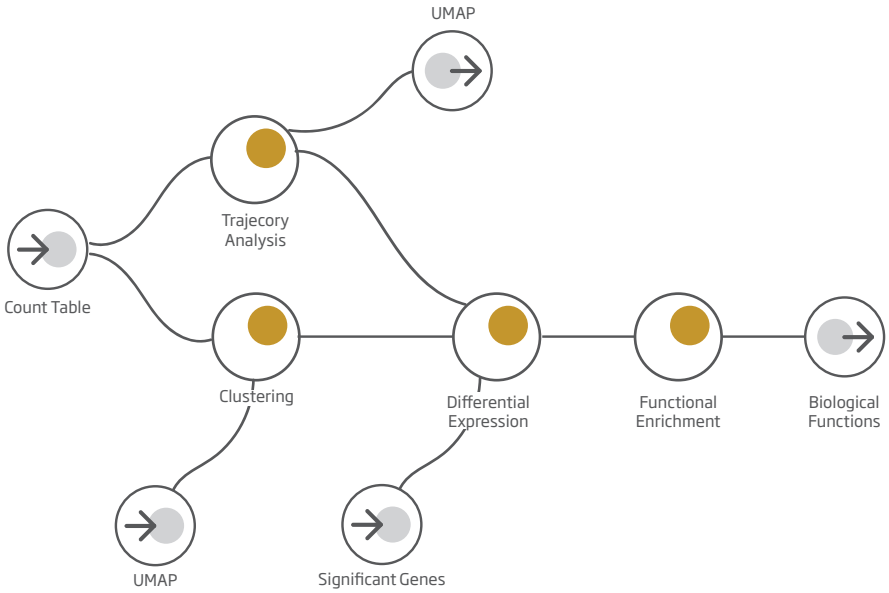
Combined Pathway Analysis

Example Workflows

Genetic Variation Data Analysis



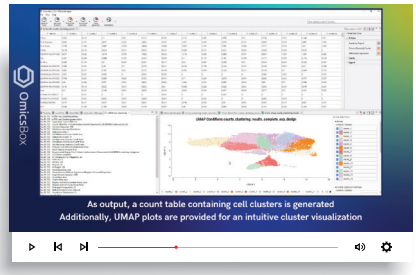
Single Cell RNA-Seq Data Analysis

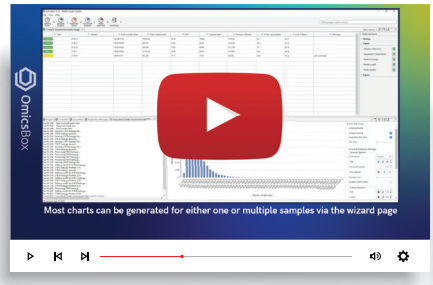


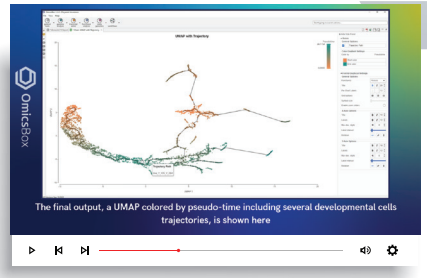
Example Data and Use Cases

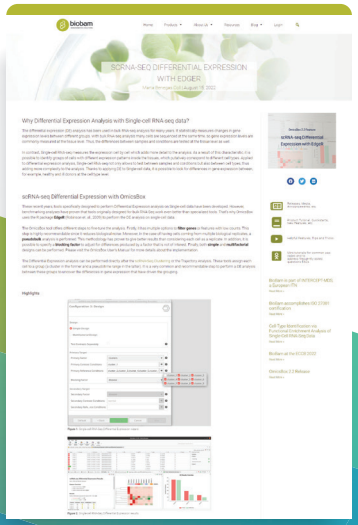
In our blog you can find many video tutorials, example analysis and datasets as well as information about the latest developments.

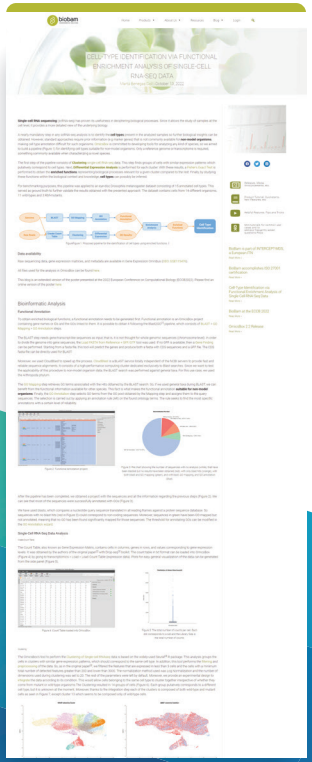
The BioBam's Blog is also a great resource to learn about all the different analysis steps and to stay up to date.











Subscribe to OmicsBox

A subscription to OmicsBox offers a powerful solution for obtaining high-quality data analysis results for your datasets with minimal time and effort. With OmicsBox, you'll have the ability to work independently, leveraging up-to-date resources and the latest technology to analyze your data.

Additionally, you'll have access to professional support throughout the duration of your subscription, ensuring that you have the guidance and assistance you need to get the most out of your data analysis.

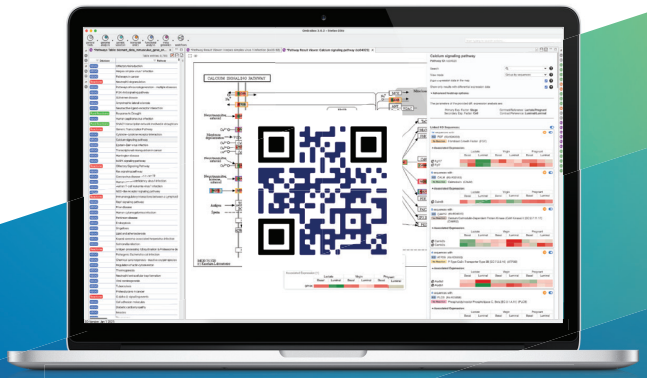
With the platform's user-friendly interface and intuitive tools, you'll be able to quickly and easily process and analyze your datasets, allowing you to focus on interpreting and understanding your results.

1 Module
1 Year
1 User

Your
Personal
Plan

Multiple
Modules
& Users

Try OmicsBox Now!

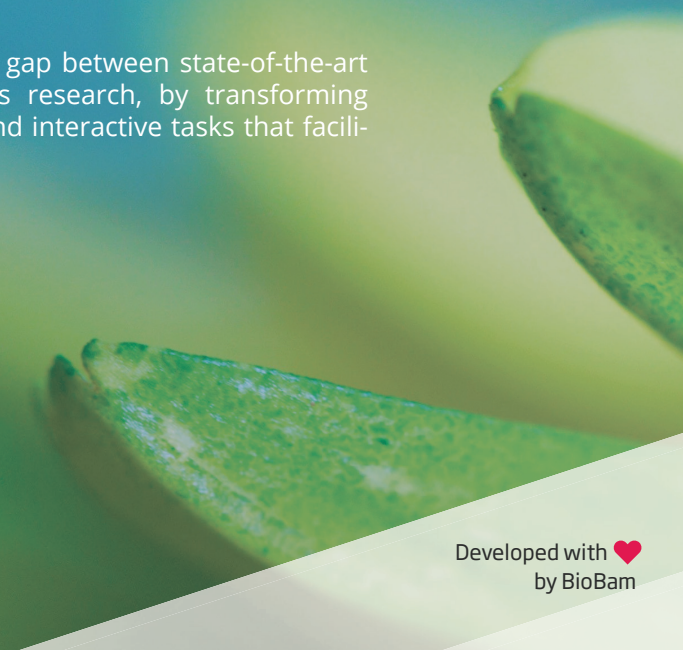


www.biobam.com

BioBam is a leading bioinformatics company that specializes in providing innovative software solutions to accelerate genomics research.

The company is dedicated to developing user-friendly and powerful bioinformatics tools that simplify data analysis for researchers, empowering them to focus on data interpretation and explore new insights.

BioBam aims to close the technology gap between state-of-the-art bioinformatics and applied genomics research, by transforming complex data analysis into intuitive and interactive tasks that facilitate scientific advancement.



biobam
BIOINFORMATICS SOLUTIONS

Developed with ♥ by BioBam