InSyBio Datastore

October 2025

Insybio Suite v3.4

InSyBio

Intelligent Systems Biology

User Manual

www.insybio.com

Introduction

Data Store is a tool that helps you upload input files to the other four InSyBio tools (Interact, ncRNASeq, BioNets and Biomarkers) and stores the datasets created by them.

USING INSYBIO DATA STORE YOU CAN:

- upload a file, and
- view, download or delete:
  - Your uploaded ncRNA Sequences Files
  - Your uploaded mRNA Sequences Files
  - Your uploaded miRNA Sequences Files
  - Your uploaded Biological Networks Files
  - Your uploaded Biological Network Clusters Files
  - Your uploaded SOFT Files
  - Your uploaded or generated Molecules Quantification Files
  - Your uploaded or generated Differential Expression Analysis Biomarkers Files
  - Your uploaded or generated Network Comparison Biomarkers Files
  - Your uploaded or generated Merged Biomarkers Files
  - Your uploaded or generated Biomarkers Datasets
  - Your uploaded or generated Biomarkers Labels Files
  - Your generated Biomarkers Model Files
  - Your uploaded or generated FASTQ Files
  - Your generated SAM Files
  - Your generated BAM Files
  - Your generated Read Counts Files
  - Your generated DESeq2 Report Files
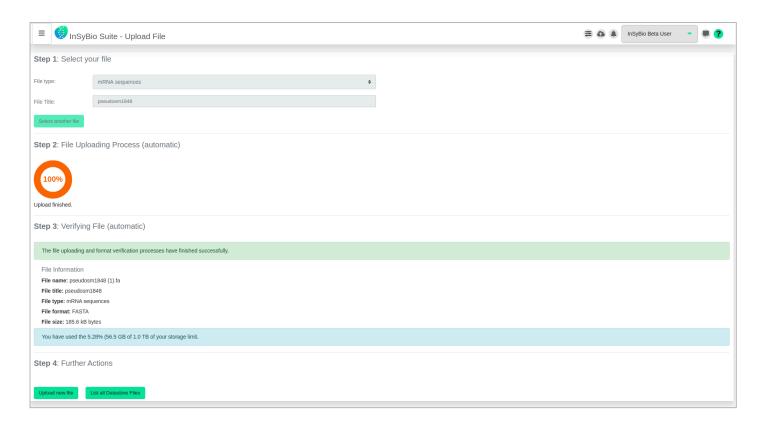 - Your uploaded Genome Files
  - Your uploaded
  - Your uploaded

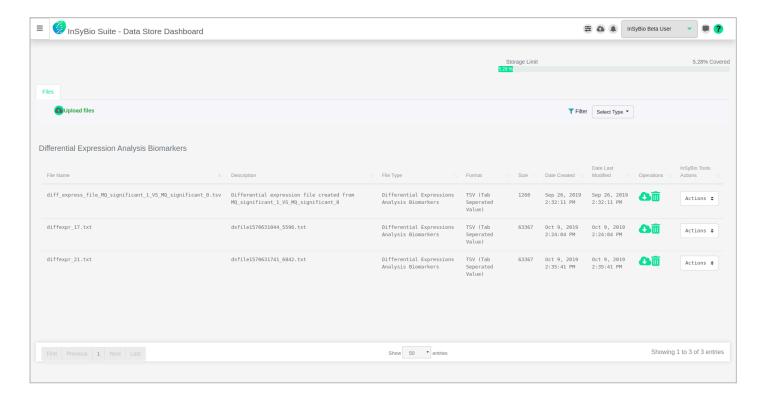# Upload New File

In order to upload a new file you have to:

A. select one of the available file types (ncRNA Sequences, miRNA Sequences, mRNA Sequences, Biological Networks, Biological Network Clusters, SOFT Files, Molecules Quantification, Differential Expression Analysis Biomarkers, Network Comparison Biomarkers, Merged Biomarkers, Biomarkers Datasets, Biomarkers Labels Files, FASTQ Files),
B. provide a title for your file,
C. select a file from your computer and
D. upload the selected file.

When the uploading process finishes, the file is verified automatically to check it has the right format.

After the verification process, the uploaded file information is presented (file name, file title, file type, file format, file size), together with a message informing you about your storage use. Your total storage limit depends on the edition of the InSyBio Suite you are using.

The upload new file process ends with Datastore asking you to upload another file or list your respective datasets.

NCRNA Sequences Files

You can upload a new ncRNA Sequences File in fasta format with one or multiple sequences. The RNA sequences datasets which can be handled by the current version are those derived from RNA sequencing experimental techniques such as RNA-seq Next Generation sequencing.

You can find a list of your uploaded ncRNA sequences. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For an ncRNA sequences file type you can select one of the two available actions:

- Feature Calculation at ncRNASeq, or
- miRNA Prediction at ncRNASeq

and be guided to the respective functionality of ncRNASeq.

MRNA Sequences Files

You can find a list of your uploaded mRNA sequences. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a mRNA sequences file, you can select one of the two available actions:

- miRNA Target Feature Calculation at ncRNASeq, or
- miRNA Target Prediction at ncRNASeq

and after also selecting the matching miRNA sequences file, be guided to the respective functionality of ncRNASeq.

MIRNA Sequences Files

You can find a list of your uploaded miRNA sequences. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a miRNA sequences file, you can select one of the two available actions:

- miRNA Target Feature Calculation at ncRNASeq, or
- miRNA Target Prediction at ncRNASeq

and after also selecting the matching mRNA sequences file, be guided to the respective functionality of ncRNASeq.

Biological Network Files

You can upload a new Biological Network File. The Biological Network datasets which can be handled by the current version are those derived from experimental techniques for the prediction of protein-protein interactions and protein complexes such as Yeast-2-Hybrid, Tandem Affinity Purification, Protein Microarrays, Co-immunoprecipitation experiments etc.

You can find a list of your uploaded biological networks. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a Biological Network file type you can select one of the four available actions:

- Visualization at BioNets,
- Network Clustering at BioNets,
- Network Analysis at BioNets

- Networks Comparison at BioNets

and be guided to the respective functionality of BioNets.

For the Networks Comparison at BioNets action, you also have to select a second biological network file. The uploaded network could be weighted. That means that each edge of the network must have a weight assigned. If the network is unweighted, you should assign equal weight (for example 1) to all edges.

Biological Network Clusters Files

You can upload a new Biological Network Clusters file. The Biological Network Clusters datasets which can be handled by the current version are those derived from experimental techniques for the prediction of protein-protein interactions and protein complexes such as Yeast-2-Hybrid, Tandem Affinity Purification, Protein Microarrays, Co-immunoprecipitation experiments etc.

You can find a list of your uploaded biological network clusters. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a Biological Network Clusters file type you can select:

- Visualization at BioNets

and be guided to the respective functionality of BioNets.

SOFT Files

You can upload a new SOFT file. Simple Omnibus Format in Text (SOFT) is the format supported by Gene Expression Omnibus database (http://www.ncbi.nlm.nih.gov/geo/) and it is the prevalent format for gene expression

experiments. It is a simple line-based, plain text format, meaning that SOFT files may be readily generated from common spreadsheet and database applications. A single SOFT file can hold both data tables and accompanying descriptive information for multiple, concatenated Platforms, Samples, and/or Series records. Affymetrix CHP files and supplementary data files like Affymetrix CEL or GenePix GPR scan files can also be provided - just bundle them together with the SOFT file into a .zip, .rar, or .tar archive using a program like WinZip at the time of submission.

You can find a list of your uploaded SOFT files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a SOFT file you can select:

- Extract Gene Expression

and be guided to the respective functionality of BioNets.

Molecules Quantification Files

You can upload a new gene expression (or more generally a molecules quantification) file. Gene expression files are tab delimited files in the following format:

gene_name_1 expr1 ... exprn

gene_name_2 expr1 ... exprn

...

You can find a list of your uploaded or generated from InSyBio BioNets Tool gene expression files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a gene expression file you can select one of the two available actions:

- Create Gene Expression Network, or
- Differential Expression Biomarkers from Analyzing Gene Expressions

and be guided to the respective functionality of BioNets.

For the Differential Expression Biomarkers from Analyzing Gene Expressions action, you also have to select a second gene expression file.

Differential Expression Biomarkers Files

You can upload a new differential expression biomarkers file. Differential expression biomarkers files are tab delimited txt files with two columns: the first column includes the names of the biomarkers and the second their p-values.

You can find a list of your uploaded or generated from InSyBio BioNets Tool differential expression biomarkers files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a differential expression biomarkers file you can select:

- Merge it with network comparison biomarkers

and be guided to the respective functionality of BioNets.

For the Merge biomarkers from network comparison biomarkers and differential expression biomarkers files action, you also have to select a network comparison biomarkers file.

Network Comparison Biomarkers Files

You can upload a new network comparison biomarkers file. Network comparison biomarkers files include the network comparison biomarkers. These biomarkers have been inferred by InSyBio's novel approach for comparing biological networks and locating nodes whose roles in the network has been significantly altered. Network comparison biomarkers files are tab delimited txt files with two columns: the first column includes the names of the biomarkers and the second a confidence score for this biomarker. This confidence score takes values from 0.5 to 1 and indicates the significance of the biomarker.

You can find a list of your uploaded or generated from InSyBio BioNets Tool network comparison biomarkers files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a network comparison biomarkers file you can select:

- Merge it with differential expression analysis biomarkers

and be guided to the respective functionality of BioNets.

For the Merge biomarkers from network comparison biomarkers and differential expression biomarkers files action, you also have to select a differential expression analysis biomarkers file.

Merged Biomarkers Files

You can upload a new merged biomarkers file. Merged biomarkers files include the biomarkers which have been inferred from merging differential expression biomarkers with network comparison biomarkers. Merged Biomarkers files are tab delimited txt files with two columns: the first column includes the names of the

biomarkers and the second a confidence score for this biomarker. This confidence score is estimated using the p-value of the differential expression biomarkers and the confidence score of the network based biomarkers.

You can find a list of your uploaded or generated from InSyBio BioNets Tool merged biomarkers files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

Biomarkers Datasets

You can upload a new biomarkers dataset. Biomarker datasets are tab delimited txt files. The rows are the biomarkers (molecules, or else features) and the columns are the samples. Thus, each cell contains the quantification value of a biomarker in a sample. Files with and without headers are both supported.

You can find a list of your uploaded or generated from InSyBio Biomarkers Tool biomarkers datasets. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

Biomarkers Labels Files

You can upload a new biomarkers labels files. Biomarker labels are tab delimited txt files. The labels (each one corresponding to a sample) should be all in one row, separated by tabs.

You can find a list of your uploaded or generated from InSyBio Biomarkers Tool biomarkers labels files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

Biomarkers Models Files

You can find a list with your generated model files from InSyBio Biomarkers Tool. You can always download or delete your dataset.

From the InSyBio Tools Actions menu, you can be guided to Biomarkers in order to apply the trained model to new data.

FASTQ Files

You can upload a new FASTQ file. You can also find a list of your uploaded or generated from InSyBio ncRNASeq Tool FASTQ files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a FASTQ file you can select:

- RNA-Seq Differential Expression Pipeline

and be guided to the respective functionality of ncRNASeq.

SAM Files

You can find a list of your generated from InSyBio ncRNASeq Tool SAM files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

BAM Files

You can find a list of your generated from InSyBio ncRNASeq Tool BAM files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

Read Counts Files

You can find a list of your generated from InSyBio ncRNASeq Tool read counts files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

DESeq2 Report Files

You can find a list of your generated from InSyBio ncRNASeq Tool DESeq2 report files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

Genome Files

You can upload a new Genome file. You can also find a list of your uploaded from InSyBio ncRNASeq Tool FASTQ files. In the list, you can view the file name, the file description, the type and the format of the file, the file size, the date of the file creation and the date of the last modification. You can always download or delete each file listed.

For a Genome file you can select:

- RNA-Seq Differential Expression Pipeline

and be guided to the respective functionality of ncRNASeq.

Node Labels Files

Node labels files are tab-delimited text files (txt) where the labels for each sample are all in one row, separated by tabs. You can upload a new node labels file to the InSyBio Datastore for efficient organization and management of sample files. Once uploaded, you can view, download, or delete your node labels files. The list view includes the file name, description, type, format, size, date of creation, and last modification date.

RDS Files

RDS files (R Data Serialization) are used in R programming to save a single R object to a file, which can be easily loaded back into R later. In the InSyBio platform, RDS files are used for storing complex data structures and analytical results, enhancing data analysis efficiency and reproducibility. After uploading RDS files to the InSyBio Datastore, users can view, download, or delete them. The list view provides detailed

information including the file name, description, type, format, size, date of creation, and last modification date.

Cloud computing Infrastructurer and Security Certifications

InSyBio Suite and all its tools are running over the cloud computing as a service infrastructure of Vultr (https://www.vultr.com), at the Amsterdam (Netherlands) facilities, offering the following security attestations and certifications (SOC 2+ (HIPAA), PCI (Merchant), CSA Star Level 1, ISO/IEC 20000-1:2018, ISO/IEC 27001:2022, ISO/IEC 27017:2015, ISO/IEC 27018:2019).

How to get InSyBio Interact

To request a free one month license of InSyBio Suite please email us at info@insybio.com .

To purchase InSyBio Interact commercial version 3.3 please contact us at sales@insybio.com.

About Us

InSyBio Inc is a bioinformatics pioneer company (www.insybio.com) in personalized healthcare, that focuses on developing computational frameworks and tools for the analysis of complex life-science and biological data in order to develop

predictive integrated biomarkers (biomarkers of various categories) with increased prognostic and diagnostic aspects for the personalized Healthcare Industry.

InSyBio Suite consists of tools for providing integrated biological information from various sources, while at the same time it is empowered with robust, user-friendly and installation-free bioinformatics tools based on intelligent algorithms and methods.

COPYRIGHT NOTICE