

Quick Start Guide

The geneXplain platform

<https://platform.genexplain.com/bioulweb/>

An inside look at how to create your personal research space on the geneXplain platform.

What can I learn from this Quick Start Guide?

Whether you have just started to use the geneXplain platform or you have already experienced it, this Quick start guide will help you with some more detailed questions you may have.

You may wonder: How can I get started with my data? How to upload? What can I analyze? How to save my results?

This Quick start guide will answer those questions and more. We are going to take you deep in to the platform.

- ✕ How do I get started?
- ✕ How is the geneXplain platform organized?
- ✕ How to work with the geneXplain platform?
- ✕ How do I upload my data?
- ✕ How to open/delete a file?
- ✕ Which types of analyses are supported?
- ✕ What is a workflow?
- ✕ Where are the workflows on the platform?
- ✕ How to start a workflow?
- ✕ Are there examples on the platform?
- ✕ How to organize my project and my data?
- ✕ How can I get information about my account?



How do I get started?

Log in to the geneXplain platform

After registration you will receive an email from BioStore with your password (subject is "BioStore: registration").

For the log in to the platform you need your email address and the password.

Please, log in here to the platform:

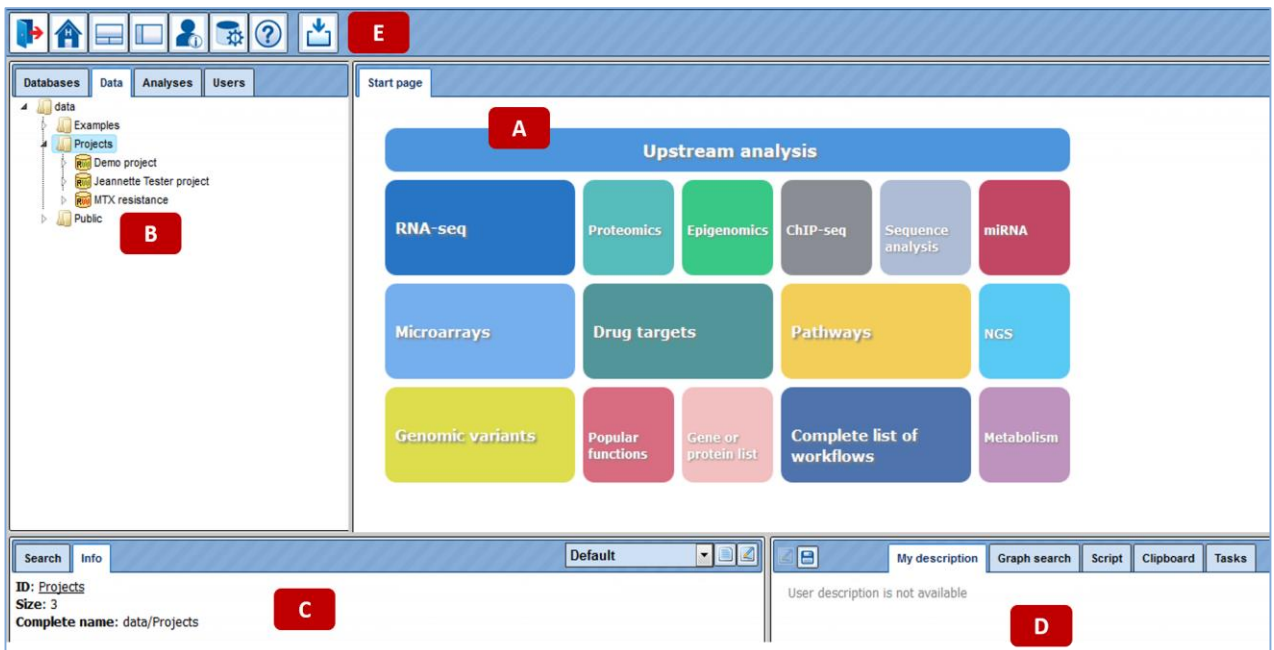
<https://platform.genexplain.com/bioumlweb/>



The image shows a screenshot of a web browser window titled "Login". The window content includes the text "Platform: geneXplain web edition 4.3", followed by the instruction "Enter e-mail and password:". Below this, there are two input fields: "E-mail:" with the value "MisterX@google.com" and "Password:" with the value "12345678". At the bottom of the window, there are three buttons: "Register", "Demo", and "Login".

How is the geneXplain platform organized?

When you login into the geneXplain platform for the first time, a window opens that contains the following areas A-E:



A - The **Work Space** is the main part of the window. The Start page presents a couple of predefined workflows and methods.

B- In the **Tree Area** you find the collection of Databases, the uploaded data files and the available analyses methods under the corresponding tabs.

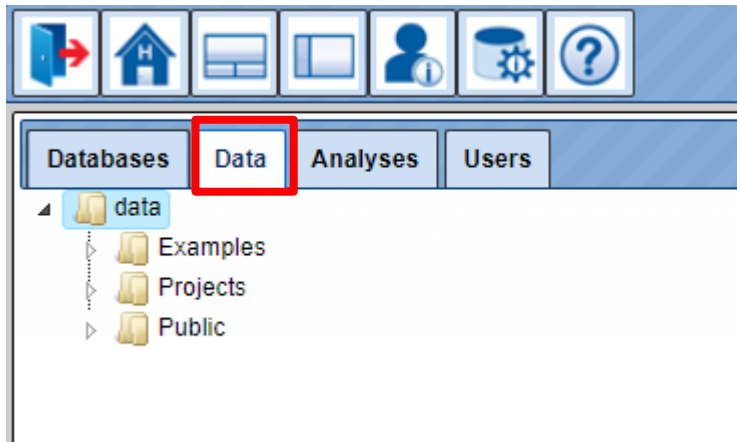
C - The **Info Box** will give you information about the data file or analysis method that you select with a single click in the Tree Area. You can also select the data resource to search in.

D - The **Operations Field** provides additional analysis options under the different tabs in a context-dependent manner.

E - The general **Control Panel** (tool bar), on top of the different areas, shows a context-dependent set of icons for the available operations.

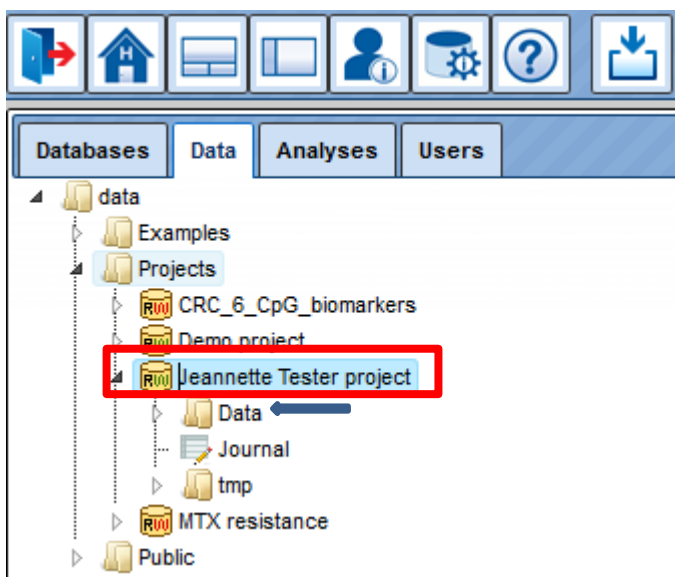
How to work on the geneXplain platform?

When you enter your account for the first time, you can see the following three folders in the Tree Area (B) under the **Data** tab (red box): Examples, Projects and Public.



Each of these folders can be expanded (opened) by clicking on the small triangle on the left side of the folder name and symbol.

You can find **your** created **project** (red box; your project name from the registration) after opening the folder Projects.



Your project contains the subfolder **Data** (blue arrow), where all your uploaded data files and analyses results reside.

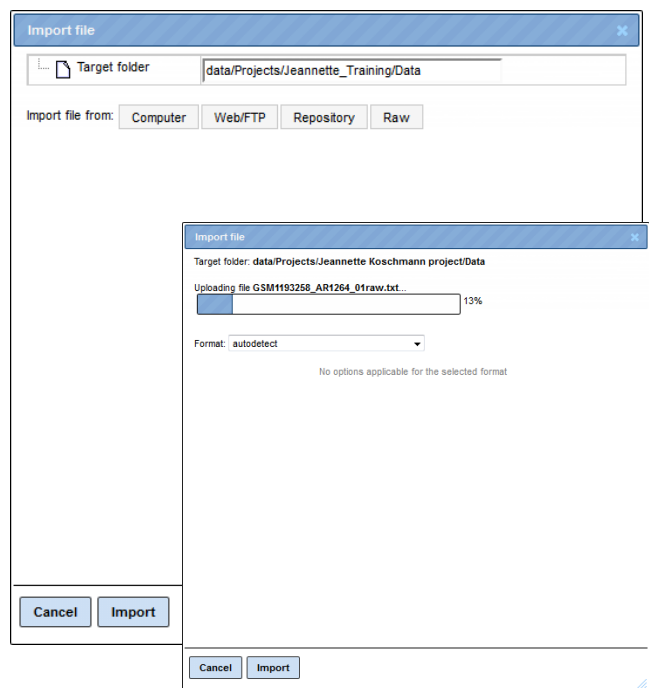
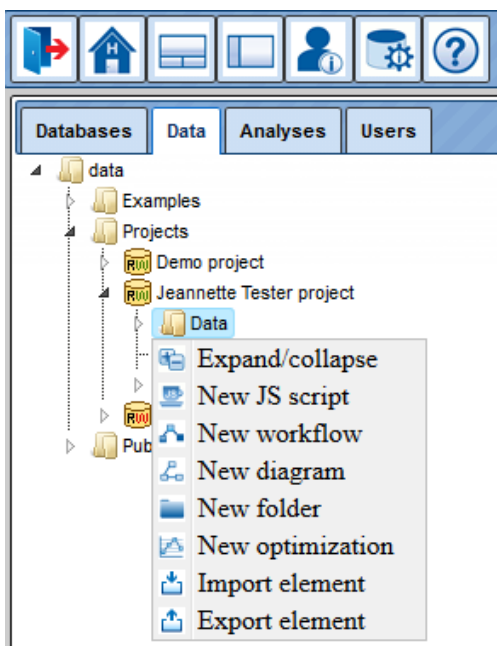
How do I upload my data?



Three ways to **import** your **data**:

1. Go to your Data folder (marked blue) and right click on it; select "Import element" and a window „Import file“ will open:

Select one of the different upload options: Computer, WebFTP, Repository or Raw.

Press the „Import“ button to start loading data.



2. Click the Import button of the tool bar [], to open the window „Import file“.
3. Click on any button of the start page. The first list entry is always „Load Data“ []. Clicking on it will lead you to the window „Import file“.

How to open/delete a file?

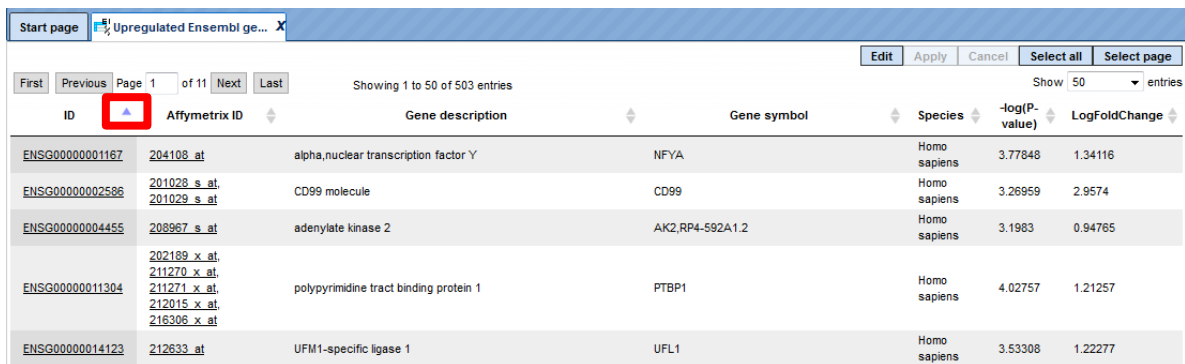
All analysis tools, and likewise all prepared workflows, require input data from a file in the Tree Area.

Double-clicking on a file will **open** it under a new tab in the Work Space.

Files can also be opened by right-clicking on them and selecting the “Open table/track” option.

You can also **delete** files (“Remove”; default value is “No”) this way.

The content of the table is sorted according to the values in one of its columns. Being opened for the first time, a default column is defined for sorting, usually the **ID** column. This default column is indicated by a blue arrowhead (red box).



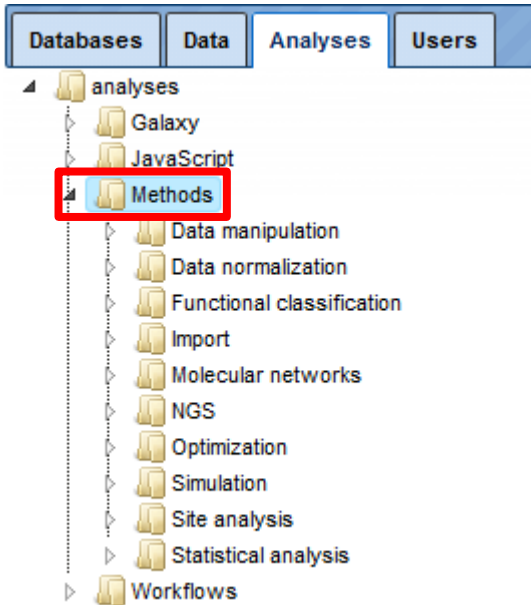
ID	Affymetrix ID	Gene description	Gene symbol	Species	-log(P-value)	LogFoldChange
ENSG0000001167	204108_at	alpha,nuclear transcription factor Y	NFYA	Homo sapiens	3.77848	1.34116
ENSG00000002586	201028_s_at 201029_s_at	CD99 molecule	CD99	Homo sapiens	3.26959	2.9574
ENSG00000004455	208967_s_at	adenylate kinase 2	AK2,RP4-592A1.2	Homo sapiens	3.1983	0.94765
ENSG00000011304	202189_x_at 211270_x_at 211271_x_at 212015_x_at 216306_x_at	polypyrimidine tract binding protein 1	PTBP1	Homo sapiens	4.02757	1.21257
ENSG00000014123	212833_at	UFM1-specific ligase 1	UFL1	Homo sapiens	3.53308	1.22277

You may **sort** the table according to the values of any column in ascending or descending order by clicking on the up- or downwards pointing arrowhead on top of this column, respectively.

On top of the table, you can **navigate** between the individual pages of the table; it is also shown on which page of the table you are, and in the right top corner, the page size in terms of number of entries (rows) is shown and can be adjusted.

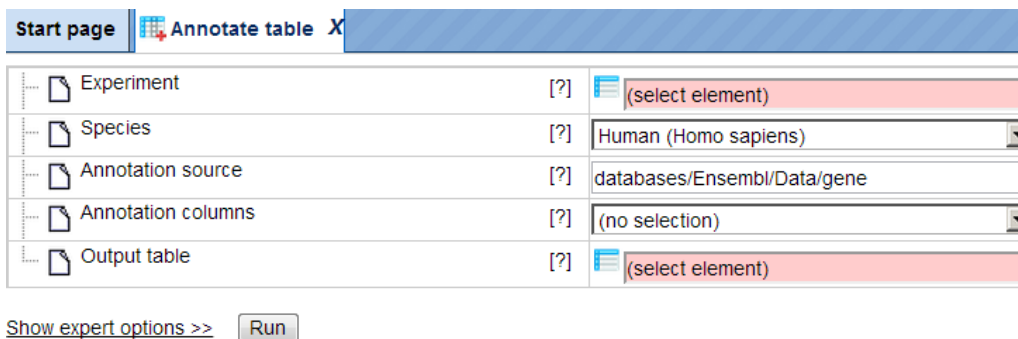
Which analyses are supported?

In the Tree Area under the tab Analyses, you find all **Methods** (red box) you may apply using the geneXplain platform.



Upon double-clicking on a method name in the Tree Area the input mask for this method will be opened in the work space.

The **input form** of one method looks like this:



Field	Value
Experiment	(select element)
Species	Human (Homo sapiens)
Annotation source	databases/Ensembl/Data/gene
Annotation columns	(no selection)
Output table	(select element)

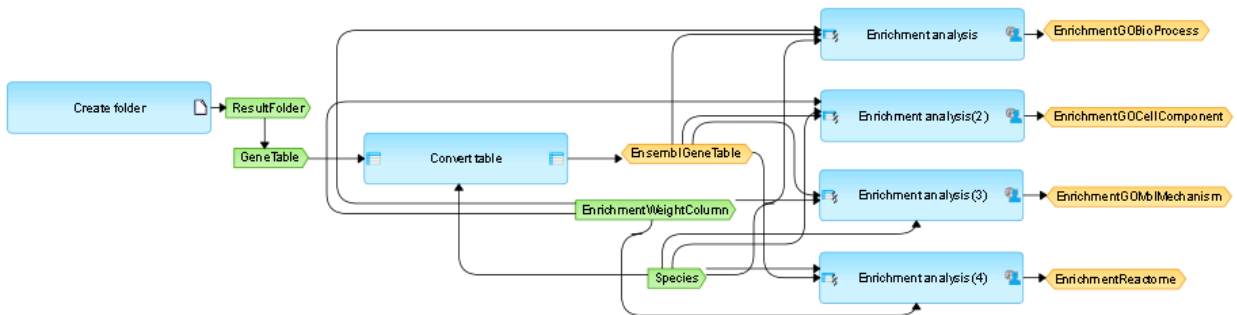
Show expert options >> Run

For starting a method just drag and drop your input file, verify your settings and press the **Run** button.

As a **result** of this analysis, a new table will be created and automatically stored in your data folder.

What is a workflow?

Workflows on the geneXplain platform are linked chains of methods to perform a bioinformatic analysis with one click.



The picture shows a small workflow for the gene set enrichment analysis (GSEA) using four different ontologies: Gene Ontology (GO) Biological Process, GO Cellular Compartment, GO Molecular Function, and Reactome's functional assignments (see right side).

The larger light **blue boxes** are analysis functions (program modules, "Bricks"). **Green boxes** stand for input files, especially user-defined inputs. **Yellow boxes** represent automatic delivery or output files.

The platform provides workflows for different topics e.g.

- ✕ Analyze ChIP-seq/RNA-seq data
- ✕ Analyze Microarray data
- ✕ Proteomics/Epigenomics/Metabolomics
- ✕ Discover functional enrichment (Gene Ontology)
- ✕ Analyze networks
- ✕ Identification of regulatory regions
- ✕ Finding drug targets
- ✕ Analyze genomic variants/mutations
- ✕ Expression and binding sites of miRNAs
- ✕ Working with NGS data

If you have valid licenses for e.g. [TRANSFAC®](#), [TRANSPATH®](#) or [HumanPSD™](#) databases the number of applications is higher.

Where are the workflows on the platform?

1. In the **Tree Area** under the tab Analyses, you find all workflows you may apply using the geneXplain platform. Upon double-clicking on a workflow name in the Tree Area (A) the method will be opened in the work space (B). Additional information is shown in the Info box (C). A schematic overview of the workflow is presented in the operation field (D).

The screenshot displays the geneXplain platform interface. On the left, a 'Tree Area' (A) shows a hierarchy of workflows under 'Analyses' > 'Common'. The main workspace (B) shows the 'Start page' for 'Compute differentially expressed genes (Affymetrix probes)', with various input fields and a 'Run workflow' button. Below the workspace, the 'Info' box (C) provides details for the selected workflow, including its title, size, and description. To the right, the 'Operation field' (D) shows a flowchart of the workflow steps, such as 'Experiment normalized', 'Up and Down Identification', and 'Intersect tables'.

2. The new platform **Start page** arranges workflows according to major areas of interest. Select one button with your favorite analysis topic. You can expand the list of available workflows by clicking. The workflows from the start page are marked as blue links.

The screenshot shows the 'Start page' of the platform, which is organized into a grid of colored buttons representing different analysis areas: RNA-seq, Proteomics, Epigenomics, ChIP-seq, Sequence analysis, miRNA, Microarrays, Drug targets, Pathways, NGS, Genomic variants, Popular functions, Gene or protein list, Complete list of workflows, and Metabolism. A red arrow points from the 'RNA-seq' button to a detailed view of RNA-seq workflows on the right. This view lists various steps and resources, including 'RNA-seq data', 'seq preprocessing', 'differentially expressed genes (DEG)', 'Gene set enrichment analysis (GSEA)', 'GO categories and metabolic pathways', 'Functional classification', and 'Networks of DEG'.

How to start a workflow?

Open the workflow input form from the Tree Area or the Start page (see e.g. red box) by clicking.

The image shows the Galaxy web interface. On the left is a tree view with categories: Databases, Data, Analyses, and Users. Under 'Analyses', there is a 'Workflows' folder containing several workflow icons: Common, GTRD, PROTEOME, TRANSFAC, and TRANSPATH. The 'TRANSFAC' workflow is highlighted. On the right is a preview of the 'Microarrays' workflow input form. The form has a title 'Microarrays' and a list of steps: Load microarray data, Normalize data, Detect differentially expressed genes, Discover functional enrichment, Analyze networks, and Analyze regulatory regions. Under 'Analyze regulatory regions', there are several sub-steps, including 'Identify enriched TF sites in promoters'. A red box highlights the option 'with TRANSFAC(R) with GTRD'. On the right side of the form, there are several icons representing different data types or analysis tools.

For starting a workflow just drag and drop your input file(s), verify your settings and press the **Run** button.

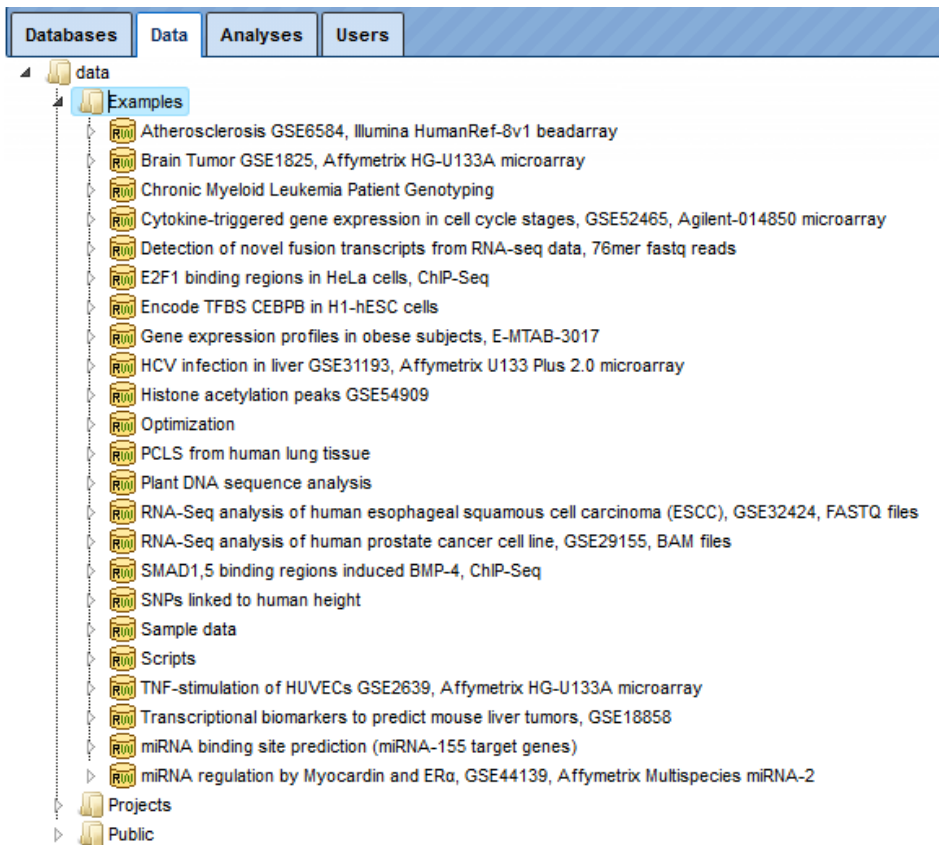
The image shows the input form for the workflow 'Identify enriched motifs in promoters (TRANSFAC(R))'. The form has a title bar with 'Start page' and a tab for the workflow. The form contains several input fields:

Input Yes gene set	(select element)
Input No gene set	...sets/Data/Housekeeping genes (Human) 300
Profile	...013.4/Data/profiles/vertebrate_human_p0.001
Species	Human (Homo sapiens)
Filter by TFBS enrichment fold	1.0
Start promoter	-1000
End promoter	100
Result folder	(select element)

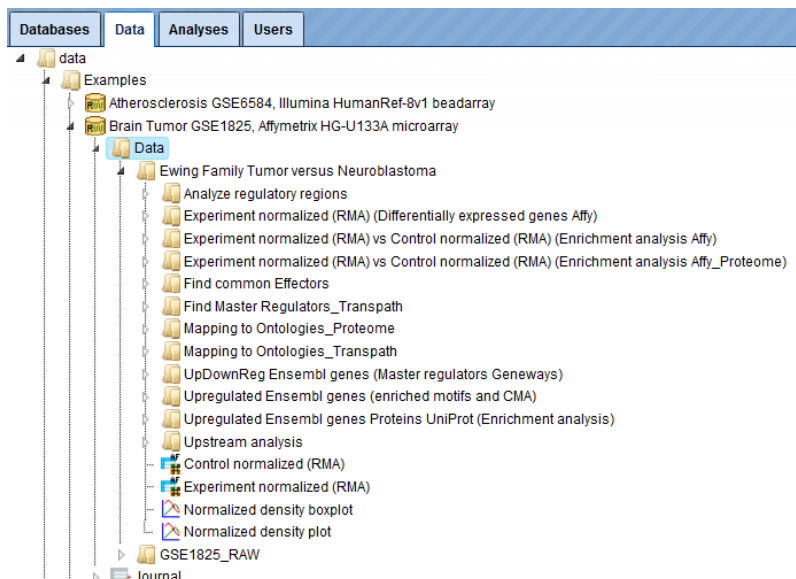
At the bottom of the form are two buttons: 'Run workflow' and 'Edit workflow'.

Are there examples on the platform?

In the **Tree Area** under the tab Data, you find all stored output examples from different input data sets. Examples include data from different microarray platforms (Affymetrix, Agilent, Illumina), ChIP-seq-data, FASTQ, BAM and BED files, SNP data or sequence collections.

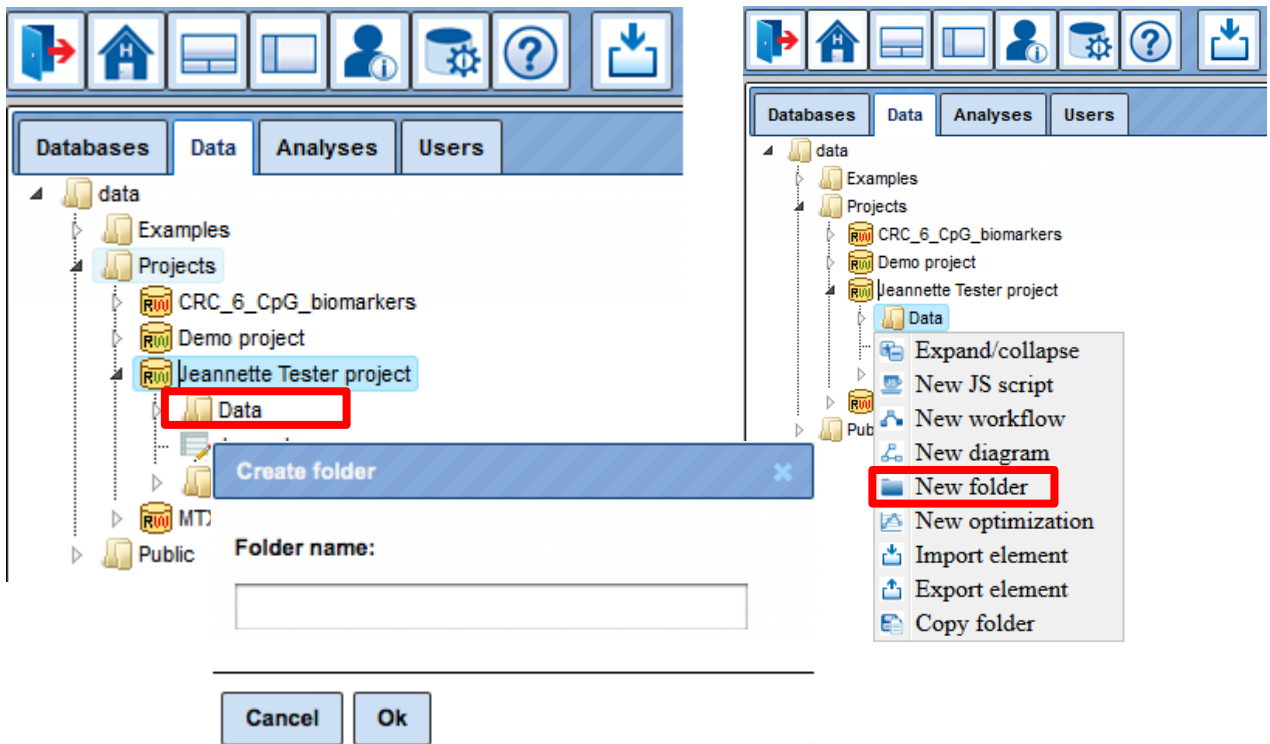


Each example comprises several analyses results, sorted in different folders. You are welcome to open the output files and folders.

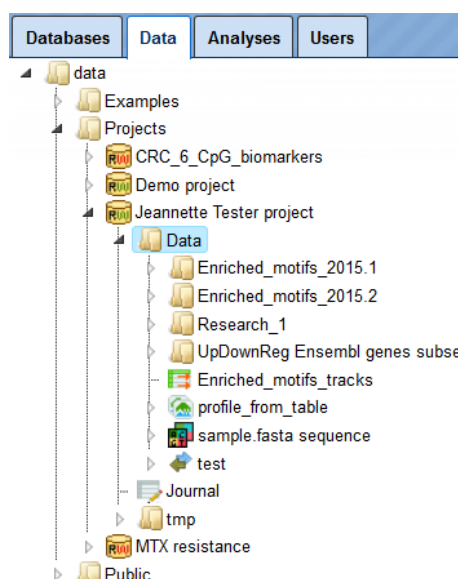


How to organize my project and my data?

Upon right mouse click on the folder Data, as in the picture below, you can create a new folder within the selected one.



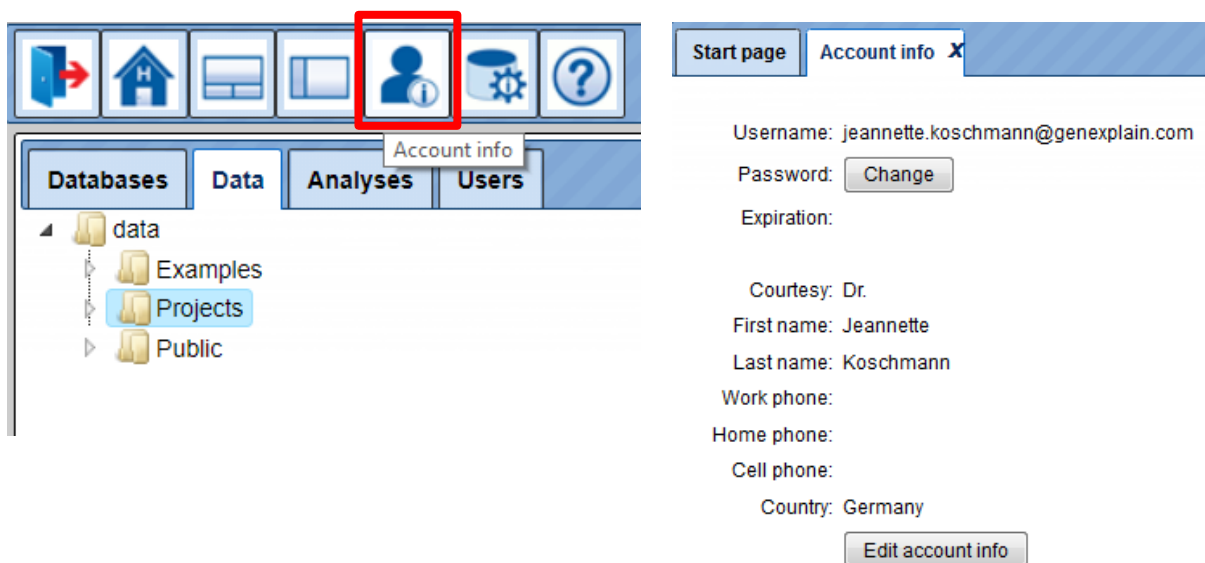
In this way, you can define the **hierarchical organization** of your folders and subfolders within your project, for example as shown below. Every time you run a workflow, you need to specify a location of the results folder, and you can specify any particular location within your project area.



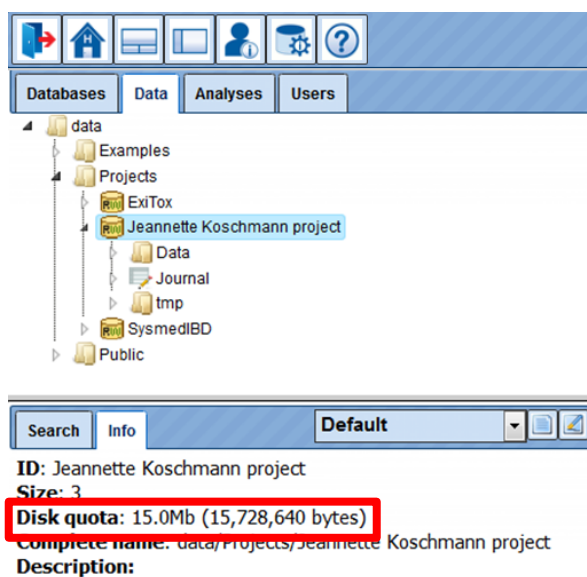
How can I get information about my account?

Your first password for the geneXplain platform is generated automatically and sent to you by email. Once you enter the platform you can change your password and also have an option to edit your personal information.

To get information about your account, select the **Account info** button (red box) on the top menu control panel.



To check your totally available work **space**, click on your project name, so that it is highlighted in blue. In the Info Box you can see information about this project including disk quota (red box). This is the space available for you. If you plan to upload large files, please make sure that you have enough work space available.



Let's start

1. [Log in](#) to the platform.
2. Import your own datasets.
3. Start your workflows.

For any further questions/suggestions /problems please contact:

info@genexplain.com

support@genexplain.com (customer support)

Useful links:

www.genexplain.com

<http://genexplain.com/genexplain-platform/>



We can explain your genes and
put bricks together...