

ODM Basic Training

Maria Borodaenko
Senior Solution Consultant



Program

Welcome to ODM training

- Learning goals
- Brief overview of what ODM does and why
- Training
- Q&A Session

Learning Goals

Learning Goals

- Get an impression how ODM can help me and my colleagues
- Get understanding of ODM capabilities
 - Know how to find relevant data
 - Learn how to import a study
 - Get knowledge about curation
 - Know how to share, export and delete a study

ODM Introduction



Making Data FAIR and Action-Ready for Both Consumers & Curators

DATA CURATORS



Metadata Curation

Harmonise thousands of samples
Enforce minimum metadata model

RESEARCHERS



Integrated Data Catalog

Find data across sources
Explore study-sample-data links

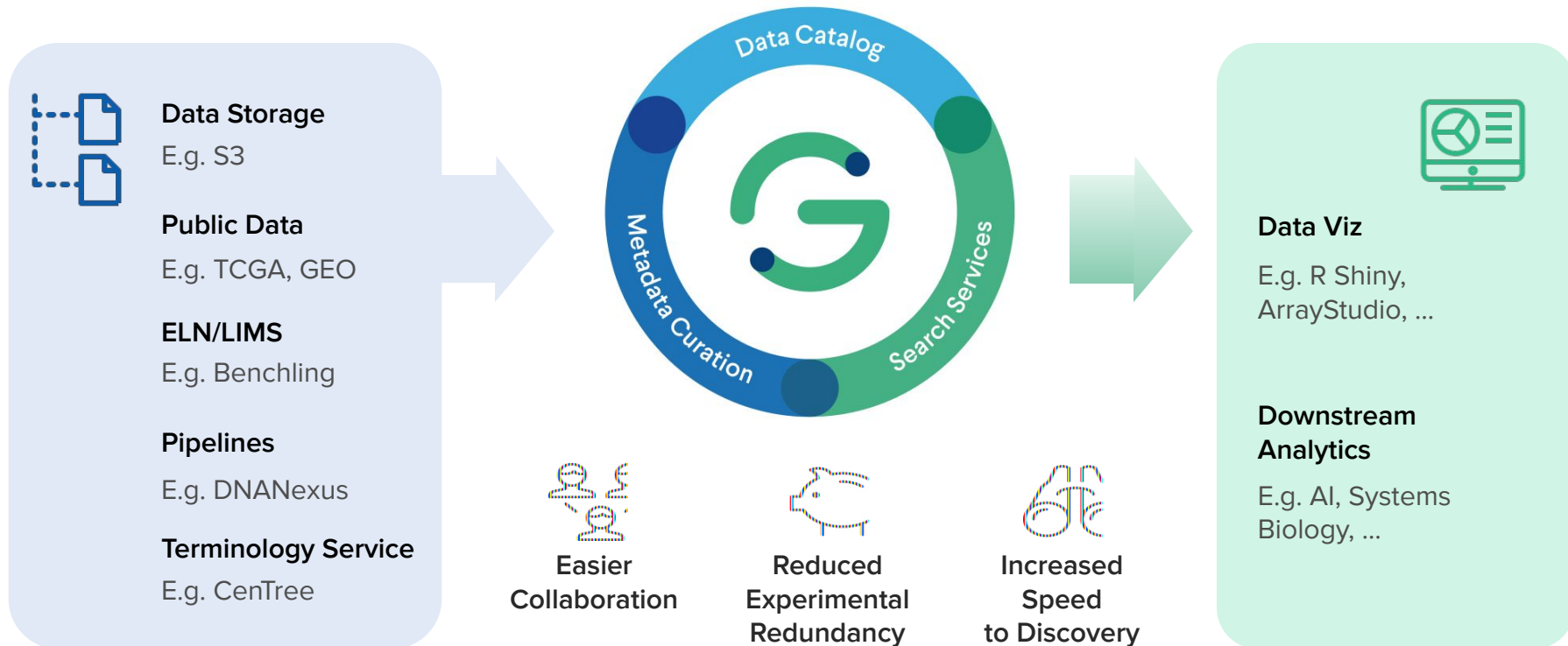
DATA SCIENTISTS



Search Services

Slice-and-dice analysis-ready data
Write R Shiny apps rapidly

A Single Point of Truth for Your Data



Training

Training Plan

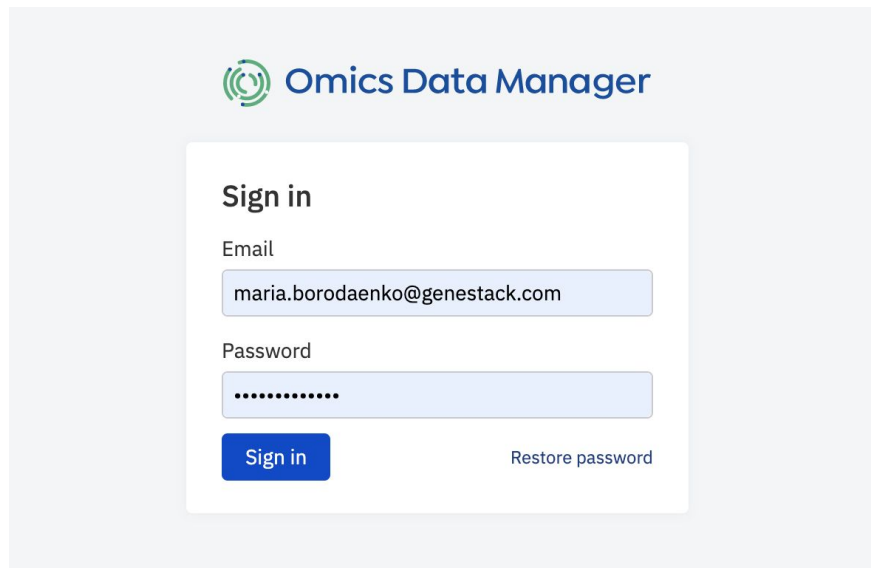
- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- Metadata curation
- Access to studies
- Study export
- Study deletion

Training Plan

- **Getting started**
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- Metadata curation
- Access to studies
- Study export
- Study deletion

Login

The users are able to login into ODM via SSO



The screenshot shows the login interface for Omics Data Manager. At the top, there is a logo consisting of three concentric green circles with a stylized 'v' inside, followed by the text 'Omics Data Manager' in blue. Below this is a white rectangular box containing the login form. The form has a title 'Sign in' in bold. It includes an 'Email' label and a text input field containing 'maria.borodaenko@genestack.com'. Below the email field is a 'Password' label and a password input field filled with dots. At the bottom left of the form is a blue 'Sign in' button, and at the bottom right is a 'Restore password' link.

Omics Data Manager

Sign in

Email

maria.borodaenko@genestack.com

Password

.....

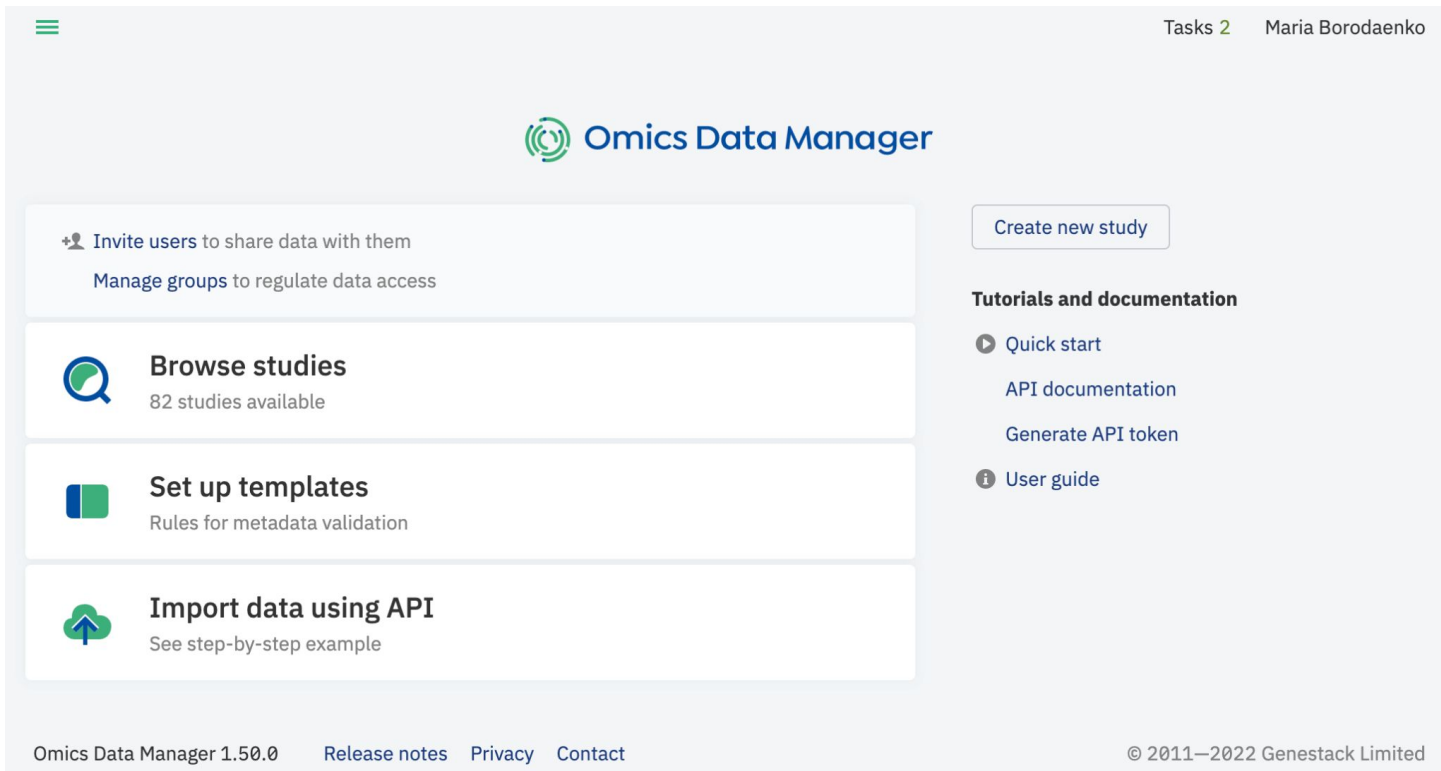
Sign in

[Restore password](#)

Starting Page

Navigation

-  [Dashboard](#)
-  [Study Browser](#)
-  [Template Editor](#)
-  [Users and permissions](#)
-  [Groups](#)
-  [User Guide](#)



The screenshot shows the Omics Data Manager starting page. At the top right, it displays 'Tasks 2' and the user name 'Maria Borodaenko'. The main header features the 'Omics Data Manager' logo. Below the header, there are several sections: a top section with 'Invite users to share data with them' and 'Manage groups to regulate data access'; a 'Browse studies' section with a magnifying glass icon and '82 studies available'; a 'Set up templates' section with a document icon and 'Rules for metadata validation'; and an 'Import data using API' section with a cloud upload icon and 'See step-by-step example'. On the right side, there is a 'Create new study' button and a 'Tutorials and documentation' section with links for 'Quick start', 'API documentation', 'Generate API token', and 'User guide'. The footer contains the version 'Omics Data Manager 1.50.0', links for 'Release notes', 'Privacy', and 'Contact', and the copyright notice '© 2011—2022 Genestack Limited'.

Tasks 2 Maria Borodaenko

 Omics Data Manager

 Invite users to share data with them
[Manage groups](#) to regulate data access

 **Browse studies**
82 studies available

 **Set up templates**
Rules for metadata validation

 **Import data using API**
See step-by-step example

Create new study

Tutorials and documentation

-  [Quick start](#)
- [API documentation](#)
- [Generate API token](#)
-  [User guide](#)

Omics Data Manager 1.50.0 [Release notes](#) [Privacy](#) [Contact](#)

© 2011—2022 Genestack Limited

Starting Page

Tasks 2 Maria Borodaenko

Omics Data Manager

+ Invite users to share data with them
Manage groups to regulate data access

Browse studies
82 studies available

Set up templates
Rules for metadata validation

Import data using API
See step-by-step example

Create new study

Tutorials and documentation

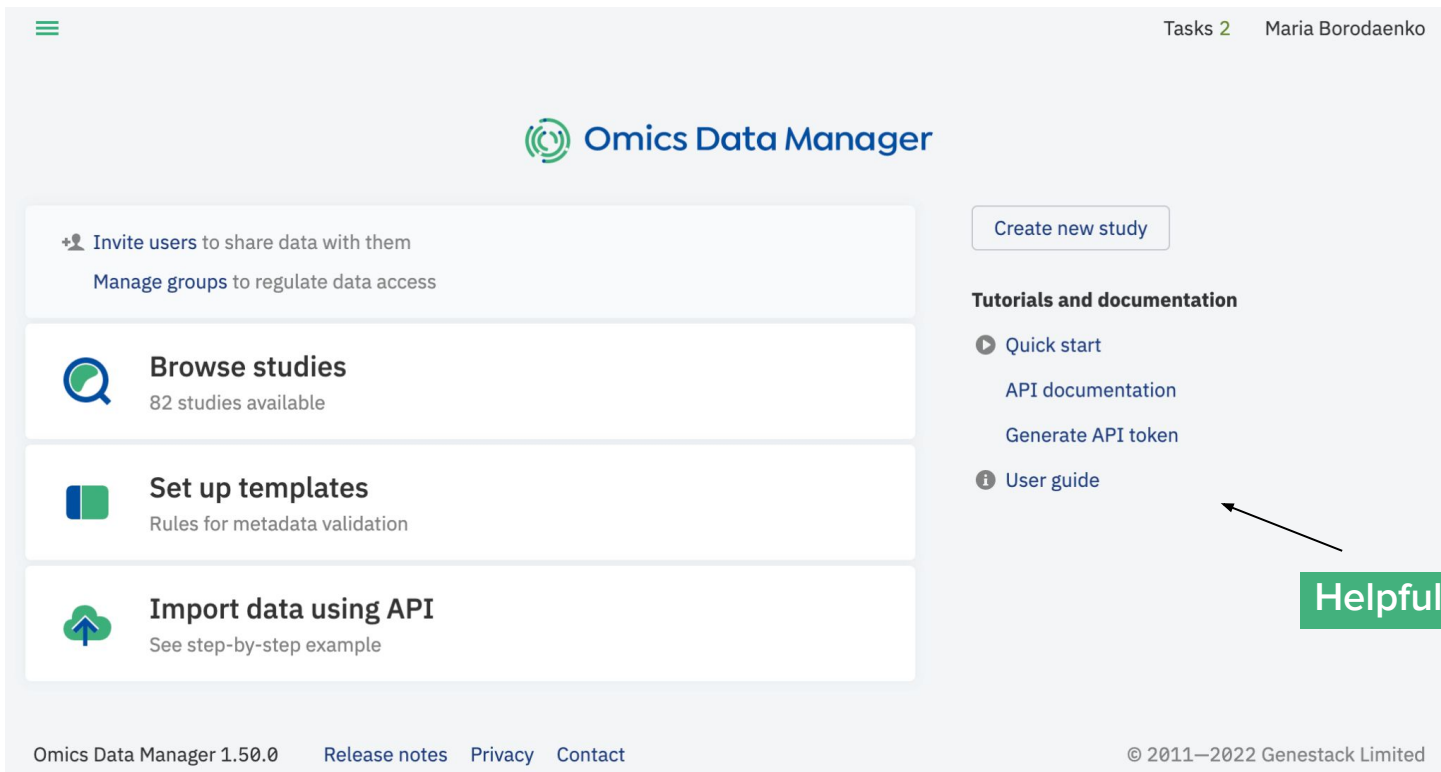
- Quick start
- API documentation
- Generate API token
- User guide

Main tasks

Omics Data Manager 1.50.0 [Release notes](#) [Privacy](#) [Contact](#)

© 2011—2022 Genestack Limited

Starting Page



The screenshot shows the Omics Data Manager starting page. At the top right, it displays 'Tasks 2' and the user name 'Maria Borodaenko'. The main header features the 'Omics Data Manager' logo. Below the header, there are several action cards on the left: 'Invite users to share data with them' with a sub-link 'Manage groups to regulate data access'; 'Browse studies' with a magnifying glass icon and '82 studies available'; 'Set up templates' with a document icon and 'Rules for metadata validation'; and 'Import data using API' with a cloud upload icon and 'See step-by-step example'. On the right side, there is a 'Create new study' button and a 'Tutorials and documentation' section containing links for 'Quick start', 'API documentation', 'Generate API token', and 'User guide'. A green box labeled 'Helpful links' with an arrow points to the 'User guide' link. The footer contains the version 'Omics Data Manager 1.50.0', links for 'Release notes', 'Privacy', and 'Contact', and the copyright notice '© 2011—2022 Genestack Limited'.

Tasks 2 Maria Borodaenko

Omics Data Manager

+ Invite users to share data with them
Manage groups to regulate data access

 **Browse studies**
82 studies available

 **Set up templates**
Rules for metadata validation

 **Import data using API**
See step-by-step example

Create new study

Tutorials and documentation

- Quick start
- API documentation
- Generate API token
- User guide

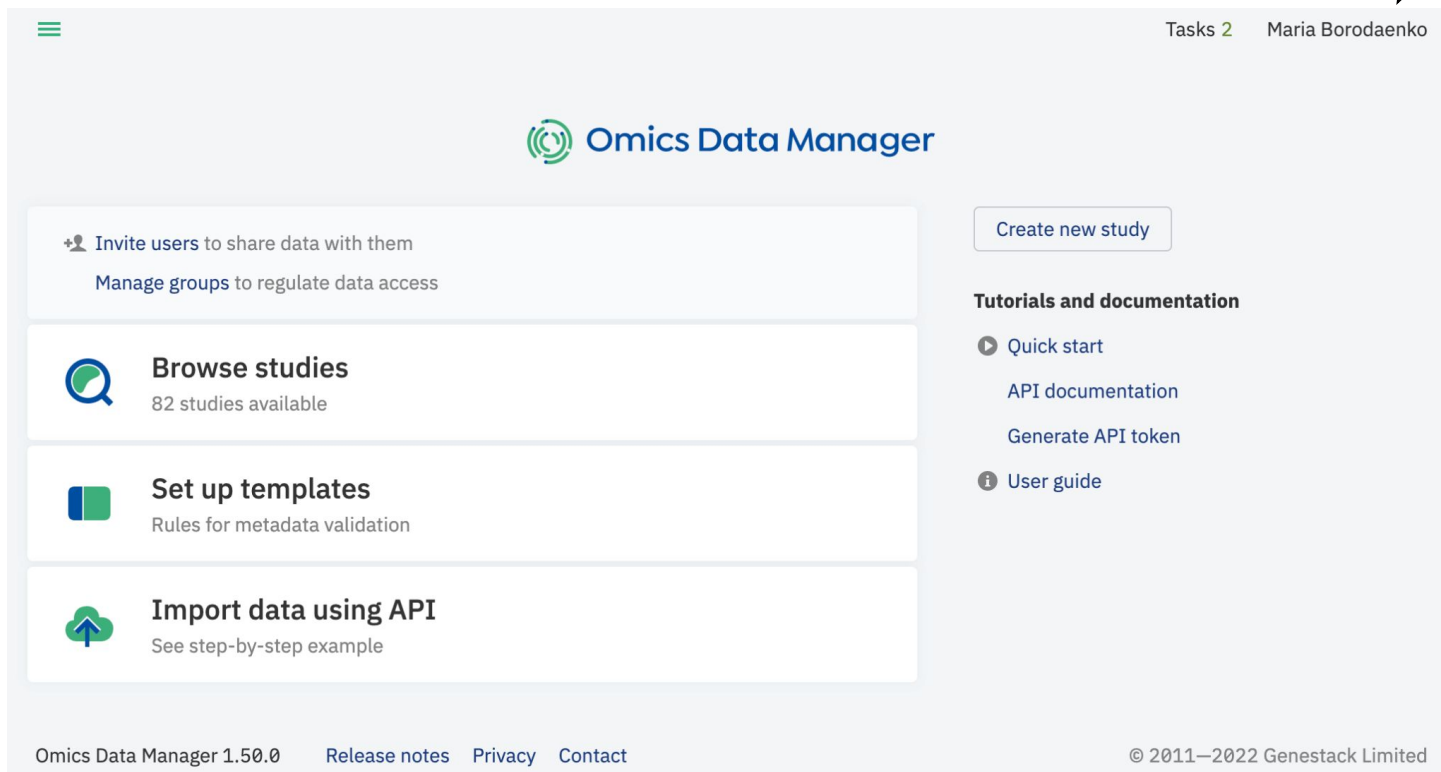
Helpful links

Omics Data Manager 1.50.0 [Release notes](#) [Privacy](#) [Contact](#)

© 2011—2022 Genestack Limited

Starting Page

Your profile



The screenshot shows the Omics Data Manager web application. At the top right, a green box labeled "Your profile" has an arrow pointing to the user name "Maria Borodaenko" in the top navigation bar. The navigation bar also includes a hamburger menu icon and a "Tasks 2" indicator. The main content area features the "Omics Data Manager" logo and a list of actions: "Invite users to share data with them" (with a sub-link "Manage groups to regulate data access"), "Browse studies" (82 studies available), "Set up templates" (Rules for metadata validation), and "Import data using API" (See step-by-step example). A "Create new study" button is located on the right. Below this is a "Tutorials and documentation" section with links for "Quick start", "API documentation", "Generate API token", and "User guide". The footer contains the version "Omics Data Manager 1.50.0", links for "Release notes", "Privacy", and "Contact", and the copyright notice "© 2011—2022 Genestack Limited".

Tasks 2 Maria Borodaenko

 Omics Data Manager

 Invite users to share data with them
[Manage groups](#) to regulate data access

 **Browse studies**
82 studies available

 **Set up templates**
Rules for metadata validation

 **Import data using API**
See step-by-step example

Create new study

Tutorials and documentation

-  [Quick start](#)
- [API documentation](#)
- [Generate API token](#)
-  [User guide](#)

Omics Data Manager 1.50.0 [Release notes](#) [Privacy](#) [Contact](#)

© 2011—2022 Genestack Limited

Profile

Profile application shows your permission role, the list of available groups, allows you to edit personal information and to get a API access token.

The screenshot displays the Genestack Profile application interface. At the top, there is a navigation bar with a hamburger menu icon, the word "Profile", and the user's name "Maria Borodaenko" next to a "Tasks 1" indicator. Below the navigation bar, the user's profile information is shown, including a circular avatar with the initials "MB", the name "Maria Borodaenko", and two role badges: "ADMIN" and "CURATOR". The email address "maria.borodaenko@genestack.com" and the groups "Curator, Maria test group" are listed below. Two buttons are present: "Change password" (blue) and "Edit personal information" (light blue). Below this, the "Genestack API tokens" section is visible, featuring a table with one token entry: "Summer token odm-demos" with a creation date of "Jun 16, 2021 14:05". To the right of the table, a text block explains that personal access tokens allow data access via the REST API and are permanent. At the bottom, a "Create new token" button is shown, with a note stating "You will be emailed a link to download token as plain text".

Profile

Tasks 1 Maria Borodaenko

MB Maria Borodaenko ADMIN CURATOR

maria.borodaenko@genestack.com

Groups: Curator, Maria test group

Change password Edit personal information

Genestack API tokens

⋮ Summer token odm-demos	Jun 16, 2021 14:05	Personal access tokens allow you to access your data via the REST API. The tokens are permanent. You can have multiple tokens and revoke them at any time
--------------------------	--------------------	---

Create new token You will be emailed a link to download token as plain text

Training Plan

- Getting started
- **Exploring a Study in Metadata Editor**
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- Metadata curation
- Access to studies
- Study export
- Study deletion

Exploring a Study

Study menu

The screenshot displays the Genestack interface for a study titled "Multiplexed enrichment and genomic profiling of peripheral blood cell specific immune signature". The interface includes a "Metadata Editor" tab and a "Study menu" dropdown. The dropdown menu is open, showing options: "Share", "Export data", "Rename", "Copy accession" (highlighted), "Save to bookmarks", "More info", "Default Template", and "Template". The study details table shows the accession number GSF051531, source GEO, and ID GSE120442. The study description is also visible.

Metadata Editor Multiplexed enrichment and genomic profiling of p... ▾

Multiplexed enrichment and genomic profiling of peripheral blood cell specific immune signature

Study Samples 80 Libraries 80 Preparations 80 EXP Expression

genestack:accession	GSF051531
Study Source	GEO
Study Source ID	GSE120442
Study Description	The human immune system consists of many specialized cell subsets that simultaneously carry out a diverse range of functions using overlapping pathways and signals. Subset-specific immune profiling can resolve immune activity in autoimmune disease, cancer immunity, and infectious disease that may not be discoverable or detectable in analyses of crude blood samples. Here, we present a low-input microfluidic system for sorting immune cells into subsets and profiling their cellular states by gene expression analysis using full-length RNA-seq. We validate the device's technical performance by benchmarking its subset enrichment and genomic profiling performance against gold standard benchtop protocols and make the added value of subset-resolved profiling over crude sample analysis clear through in vitro experiments that show subset-specific stimulated responses. Our results show that even the

View only Published by Maria Borodaenko on Nov 21, 22:07

Exploring a Study

Metadata Editor Multiplexed enrichment and genomic profiling of peripheral specific immune signature

Study Samples 80 EXP Expression VAR Variant

Export dataset

genestack:accession GSF316814

Experimental Platform Illumina Genome Analyzer II

Data Processing Method VarScan

Scale No value

Genome Version GRCh38

Tabs with linked data

genestack:accession GSF316814

Study Source GEO

Study Source ID GSE120442

Study Description The human immune system carry out a diverse immune profiling of infectious disease samples. Here, we subsets and profiled seq. We validate the and genomic profile added value of sub experiments that

Study Samples 80 EXP Expression VAR Variant

Filters

Sex	Disease	Age	Age Unit
2 distinct values	2 distinct values	8 distinct values	1 distinct value
☐ female 72	systemic lup... 40	59 24	☒ year 80
☐ male 8	Healthy 40	57 8	

Published by Maria Borodaenko on Aug 8, 2020

Exploring Sample Filters

Open filters to analyse data

Metadata Editor Multiplexed enrichment and genomic profiling of p... Tasks 5 3 Maria Borodaenko

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature

Export

Study Samples 80 Libraries 80 Preparations 80 EXP Expression

Filters 41 samples: Disease systemic lupus eryth... x

Sex	Disease	Age	Age Unit	Tissue
2 distinct values	2 distinct values	8 distinct values	2 distinct values	2 distinct values
female 37	☒ systemic lup... 41	☐ 59 11	year 22	Blood 22
male 4	Healthy 39	☐ 57 6	No value 19	No value 19
		☐ 51 5		
		☐ 52 5		
		☐ 24 4		
		☐ 26 4		
		☐ 81 4		
		☐ 29 2		

Search

Apply

Select values of interest

Apply to explore a subset of samples

Study with Multiple Datasets

List of available datasets

Dataset in focus

Metadata Editor Multiplexed enrichment and genomic profiling of p... Tasks 5 3 Maria Borodaenko

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature

Export

Study Samples 80 Libraries 80 Preparations 80 EXP Expression

Datasets

- training.2.0-tx-expression.gct
- training.px-expression.gct
- training.rpkm-tx-expression.gct**

Export dataset

genestack:accession	GSF051699
Experimental Platform	Illumina NextSeq 550
Data Processing Method	Bowtie2, RSEM
Scale	RPKM
Genome Version	GRCh37
Pipeline ID	No value
Data Files / Raw	No value

View only Published by Maria Borodaenko on Nov 21, 22:24

Experiment's metadata

Multiple Versions of The Same Dataset

If few versions of dataset are available, they will be displayed. The latest version is active and focused by default, previous version available for exploring and exporting.

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature

Export

Study

Samples 80

Libraries 80

Preparations 80

EXP Expression

Datasets

training.2.0-tx-expression.gct

training.px-expression.gct

training.rpkm-tx-expression.gct

v. 2 ACTIVE

v. 1

Export dataset

genestack:accession

GSF052036

Experimental Platform

Illumina NextSeq 550

Data Processing Method

Bowtie2, RSEM

Scale

TPM

Version History

The screenshot displays the Genestack Metadata Editor interface. The main title is "Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature". The interface includes tabs for "Study", "Samples 80", "Libraries 80", "Preparations 80", and "EXP Expression". A table lists metadata fields: "genestack:accession" (GSF051531), "Study Source" (GEO), "Study Source ID" (GSE120442), and "Study Description" (The human immune system consists of many specialized cells that carry out a diverse range of functions using overlapping patterns of gene expression. Immune profiling can resolve immune activity in autoimmune and infectious disease that may not be discoverable or detectable in bulk samples. Here, we present a low-input microfluidic system for isolating and profiling their cellular states by gene expression). A green box labeled "History of changes" points to the "Study Description" field. A dark overlay titled "Version history" lists three changes: "Fixing Sex attribute" (Published by Maria Borodaenko on Jun 16, 14:33), "uncurated" (Published by Fatima Sabitova on Jun 6, 14:33), and "organism and sex was changed" (Published by Fatima Sabitova on May 16, 16:57). A "Back to latest version" button is visible. The bottom status bar shows "View only" and "Published by Maria Borodaenko on Nov 21, 22:07".

Metadata Editor Multiplexed enrichment and genomic profiling of p... Tasks 7 Maria Borodaenko

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature Export

Study Samples 80 Libraries 80 Preparations 80 EXP Expression

genestack:accession	GSF051531
Study Source	GEO
Study Source ID	GSE120442
Study Description	The human immune system consists of many specialized cells that carry out a diverse range of functions using overlapping patterns of gene expression. Immune profiling can resolve immune activity in autoimmune and infectious disease that may not be discoverable or detectable in bulk samples. Here, we present a low-input microfluidic system for isolating and profiling their cellular states by gene expression.

Version history

- Fixing Sex attribute
Published by Maria Borodaenko on Jun 16, 14:33
- uncurated
Published by Fatima Sabitova on Jun 6, 14:33
- organism and sex was changed
Published by Fatima Sabitova on May 16, 16:57

History of changes

Back to latest version

View only Published by Maria Borodaenko on Nov 21, 22:07

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- **Searching for available data in Study Browser**
- Data loading
- Metadata templates
- Metadata curation
- Access to studies
- Study export
- Study deletion

Searching for Available Data

The screenshot displays the Genestack Study Browser interface. At the top, a search bar contains the query "GSE66728, brain, Parkinson's Disease, GTEx, any metadata". The user's name, Maria Borodaenko, and the number of tasks (1) are shown in the top right. The left sidebar contains filters for Access (My studies, For all users, Curator, Maria test group), Signal type (Expression 47, Flow Cytometry 2, Variant 8), Organism (Homo sapiens 16, Human 3, Camellia sinensis 2, Mus musculus 2, koala metagenome 2), Data Type (RNA-seq 6), and Study Source (Arvados 10, GEO 10, ArrayExpress 2). The main content area shows 47 studies found, filtered by Signal type and Expression. The first four studies are listed:

- GTEx version 3** (22951 EXP)
GTEx
Created by Fatima Sabitova on Aug 5, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [gene expression levels] - CTA** (373 EXP)
Single-cell, Homo sapiens, Basal pore, Not reported, Alveolar +3 more
Created by Fatima Sabitova on Jul 24, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [gene expression levels] - WTA** (403 EXP)
Single-cell, Homo sapiens, Basal pore, Not reported, Alveolar +3 more
Created by Fatima Sabitova on Jul 24, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [protein levels]** (330 EXP)
Single-cell, Homo sapiens, Basal pore, Seromucinous glands, Alveolar +6 more
Created by Fatima Sabitova on Jul 23, 2022

The fifth study is partially visible:

- [UNCURATED] A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity** (156 EXP)
Homo sapiens, saliva, nasal, Blood, COVID, Control
Created by Fatima Sabitova on Jul 12, 2022

Faceted Search

Access

- My studies
- For all users
- Curator
- Maria test group

Signal type

- Expression 47
- Flow Cytometry 2
- Variant 8

Organism

- Homo sapiens 16
- Human 3
- Camellia sinensis 2
- Mus musculus 2
- koala metagenome 2

Data Type

- RNA-seq 6

Study Source

- Arvados 10
- GEO 10
- ArrayExpress 2

47 studies found

Filtered by: Signal type Expression

- GTEx version 3
GTEx
Created by Fatima Sabitova
- A single-cell and spatial targets of SARS-CoV-2 [g]
Single-cell, Homo sapien
Created by Fatima Sabitova
- A single-cell and spatial targets of SARS-CoV-2 [g]
Single-cell, Homo sapien
Created by Fatima Sabitova
- A single-cell and spatial targets of SARS-CoV-2 [p]
Single-cell, Homo sapien
Created by Fatima Sabitova on Jul 23, 2022
- [UNCURATED] A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity
Homo sapiens, saliva, nasal, Blood, COVID, Control
Created by Fatima Sabitova on Jul 12, 2022

Expression atlas 2

ArrayExpress 1

TCGA 1

Therapeutic Area more >

- Infectious Disease 3
- Immunology 2
- Oncology 2
- Crohn's Disease 1
- Immuno-oncology 1

Disease less <

Sex

- female 8

Disease

Search

- Not applicable 3
- COVID-19 2
- Healthy 2
- lung adenocarcinoma 2
- lung squamous cell carcinoma 2
- acinar cell carcinoma 1
- adenocarcinoma 1
- Adenocarcinoma with mixed subtypes 1
- adenosarcoma 1
- Anesthesia 1

Full-Text Search

The screenshot displays the Genestack Study Browser interface. At the top, a search bar is highlighted with a green box and contains the text "GSE66728, brain, Parkinson's Disease, GTEx, any metadata". Below the search bar, the interface shows "47 studies found" and "Filtered by: Signal type Expression". A list of studies is displayed, including "GTEx version 3", "A single-cell targets of SARS-CoV-2", and "A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [protein levels]".

A dropdown menu is open, showing search results for "human". The results include "Homo sapiens preferred label of humans", "Human herpesvirus 5", "human granulocytic anaplasmosis", and "human intestinal bacterium SNU-Julong732". A red arrow points to the text "Terms from ontologies" next to the search results.

The left sidebar contains filters for Access, Signal type, Organism, Data Type, and Study Source. The "Signal type" filter is set to "Expression 47". The "Organism" filter shows "Homo sapiens" with 16 studies, "Human" with 3 studies, "Camellia sinensis" with 2 studies, "Mus musculus" with 2 studies, and "koala metagenome" with 2 studies. The "Data Type" filter shows "RNA-seq" with 6 studies. The "Study Source" filter shows "Arvados" with 10 studies, "GEO" with 10 studies, and "ArrayExpress" with 2 studies.

Search Results

Study Browser Tasks 1 Maria Borodaenko

47 studies found

Filtered by: **Signal type** **Expression** ×

- GTEx version 3** 22951 EXP
GTEx
Created by Fatima Sabitova on Aug 5, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [gene expression levels] - CTA** 373 EXP
Single-cell, Homo sapiens, Basal pore, Not reported, Alveolar +3 more
Created by Fatima Sabitova on Jul 24, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [gene expression levels] - WTA** 403 EXP
Single-cell, Homo sapiens, Basal pore, Not reported, Alveolar +3 more
Created by Fatima Sabitova on Jul 24, 2022
- A single-cell and spatial atlas of autopsy tissues reveals pathology and cellular targets of SARS-CoV-2 [protein levels]** 330 EXP
Single-cell, Homo sapiens, Basal pore, Seromucinous glands, Alveolar +6 more
Created by Fatima Sabitova on Jul 23, 2022
- [UNCURATED] A multi-tissue study of immune gene expression profiling highlights the key role of the nasal epithelium in COVID-19 severity** 156 EXP
Homo sapiens, saliva, nasal, Blood, COVID, Control
Created by Fatima Sabitova on Jul 12, 2022

Search results

Access

- ☒ My studies
- ☐ For all users
- ☐ Curator
- ☐ Maria test group

Signal type

- Expression 47**
- Flow Cytometry 2
- Variant 8

Organism more >

- ☐ Homo sapiens 16
- ☐ Human 3
- ☐ Camellia sinensis 2
- ☐ Mus musculus 2
- ☐ koala metagenome 2

Data Type

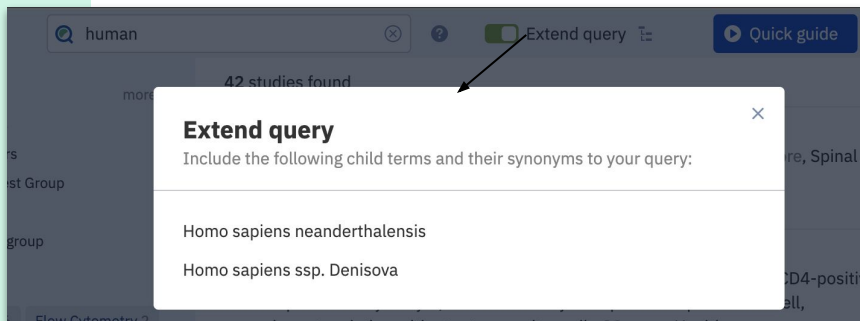
- ☐ RNA-seq 6

Study Source more >

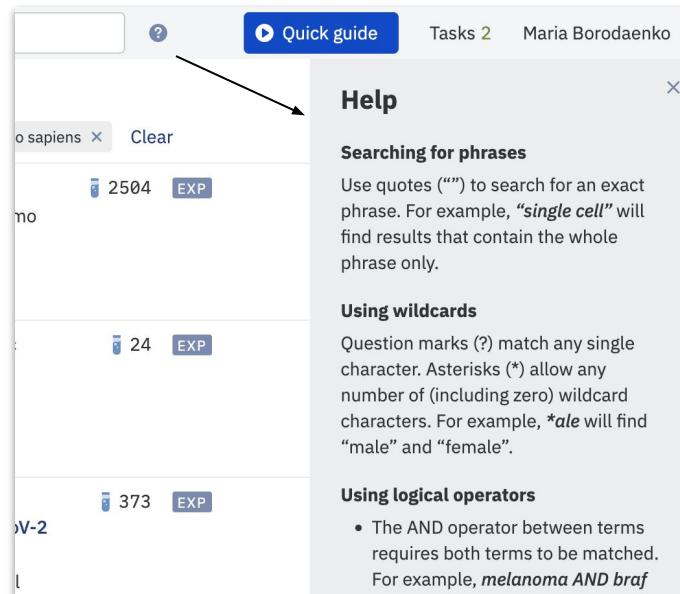
- ☐ Arvados 10
- ☐ GEO 10
- ☐ ArrayExpress 2

Advanced Search Capabilities

The query can be extended based to include child terms from the ontology



Help button for using logical operators, wildcards and more



Bookmarks

In case you have some studies which you regularly work with, it could be easier and faster to find them among bookmarks.

Find it quickly
next time

Save an interesting
study

Bookmarks

⋮  Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature

Created by Maria Borodaenko on Nov 21, 2022

⋮  GTEx

Created by Genestack Superuser on Nov 10, 2022

Study Browser

 Saved

Access

 My studies

 For all users

☐ Maria Test Group

Signal type

Expression 7

Flow Cytometry 1

Variant 1

Therapeutic Area

 GSE66728, brain, Parkinson's Disease, GTEx, any m

7 studies

⋮  Multiplexed enrichment and g

 Share

 Export data

Copy accession

 Save to bookmarks

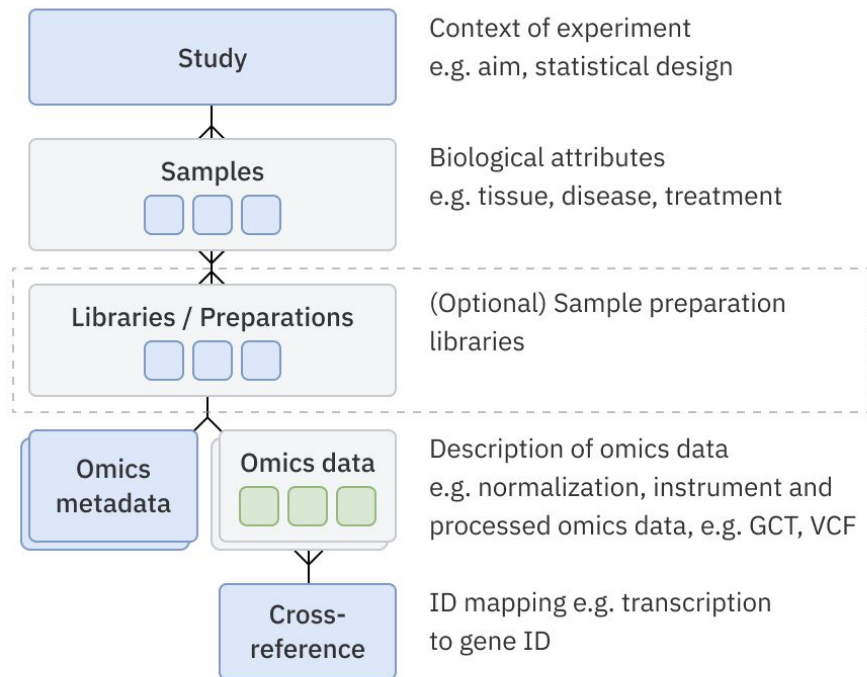
More info

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- **Data loading**
 - Metadata templates
 - Metadata curation
 - Access to studies
 - Study export
 - Study deletion

Data Model and Supported formats

Metadata Data



Import	GUI	Python Script
Study	Manual	TSV
Samples	Manual or TSV import	TSV
Libraries / Preparations (optional)	X	TSV
Omics metadata	X	TSV
Omics data	X	GCT, VCF, FACS

Linking Attributes

Headers in the expression file should correspond to sample metadata so the system recognise what column relates to what sample.

During import user can specify which attribute from the template should be used for linking during import.

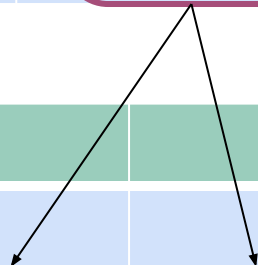
If nothing is specified the “Sample Source ID” attribute is used.

Can be any template attribute



Sample file		
Name	Sample Source	Sample Source ID
Sampe_1	Lab_1	Sample_1_ID
Sample_2	Lab_1	Sample_2_ID
Sample_3	Lab_1	Sample_3_ID

Expression file			
#1.2			
24571	60		
Name	Description	Sample_1_ID	Sample_2_ID
Gene_1		2.23451	0



Data Loading via Script

During study import a study itself and associated data objects are created sequentially and linked to each other based on the data model.

This is done by the import data script which calls REST APIs. The user should be a curator and should provide an access token, links to the files and the server address. More details about the parameters can be found in the script's help.

Example command

```
python3 import_ODM_data.py \
--token YOUR_TOKEN \
--server YOUR_SERVER \
--study s3://data_source/study.tsv \
--samples s3://data_source/samples.tsv \
--libraries s3://data_source/libraries.tsv \
--expression s3://data_source/expression.gct \
--em s3://data_source/expression.gct.tsv
```

```
study GSF052044 was added successfully
samples were added successfully (sample group accession is GSF052044)
Successfully linked: [samples_to_study]
libraries were added successfully (libraries group accession is GSF052044)
Successfully linked: [libraries_to_samples]
Successfully linked: [libraries_to_samples]
expression data were added successfully (group accession is GSF052210)
Successfully linked: [expression_to_libraries]
Execution is finished!
```

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- **Metadata templates**
- Metadata curation
- Access to studies
- Study export
- Study deletion

Metadata Templates

For effective metadata curation the Templates are used. Template are set by users with “Manage templates” permission in Template Editor.

They describe the attributes and select ontologies for metadata validation. There could be more than 1 template on the instance to have set of attributes specific to a study type. The default template can be set.

The screenshot displays the 'Template Editor' interface. On the left, a sidebar lists various metadata attributes: Study, Sample, Library, Preparation, Expression, and Variant. The main area is titled 'Baseline Biospyder Template' and contains a table with columns for Name, Required, and Meta. The table lists attributes like Accession, SAMPLE_ID, Sample Name, Organism, and Tissue, each with a 'Yes' or 'No' requirement and a 'Text' meta type. A modal window titled 'Template Editor' is open, showing a 'Default template' section with a 'Default Template' button and a description: 'The default template is applied to a study when no other template is selected.' To the right of the modal, a 'Templates' section lists available templates: 'Template for BZ' and 'Agri'.

Name	Required	Meta
Accession	Yes	Text
SAMPLE_ID	Yes	Text
Sample Name	Yes	Text
Organism	Yes	Text
Tissue	No	Text

Template Editor

Default template

Default Template ...

The default template is applied to a study when no other template is selected.

Templates

Name

- Template for BZ**
- Agri**

Metadata Validation

Curators can choose a template for their studies.
Default template is applied by default.

The study will be validated against the template.
Any inconsistencies will be shown.

Validation summary ✕
Select value to replace

Organism

No value 80
Not filled

Sex

F 72
Should be a preferred label from dictionary "Sex"

M 8
Should be a preferred label from dictionary "Sex"

Cell Type

No value 80
Not filled

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal sub-specific immune signature

Study Samples 80 Libraries 80 EXP Expression

Filters

Invalid metadata



	A	D	E	F	G	H	I	J
	genestack:...	Organism	Sex	Disease	Age	Age Unit	Tissue	Cell Type
1	GSF052047		F		24			
2	GSF052048		F		24			

Required Templates Attributes

There are a few attributes required by the system. These attributes should be presented in any template to allow ODM work properly. The following attributes should not be removed from the attribute or altered in spelling.

ATTRIBUTE	ENTITY	COMMENT
“Accession”	All entities	Mandatory for the template.
“Study Title”	Study	Without these attributes the users will get errors upon opening a study in the GUI
“Sample Source ID” (or alternative linking attribute)	Samples	Linking attribute should present in the template for proper linkage

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- **Metadata curation**
- Access to studies
- Study export
- Study deletion

Metadata Curation

For curating metadata one should be a member of the Curator group.

The studies is opened in a view mode, to check validation and to apply changes the curator should switch to the Edit mode.

Changes made in the Edit mode are available for other curators but they are not seen by the researcher.

Start curation

The screenshot displays the 'Metadata Editor' interface. At the top, there's a navigation bar with a hamburger menu, the title 'Metadata Editor', and a status indicator 'Multiplexed enrichment'. Below this, the main title 'Multiplexed enrichment and genomic subset-specific immune signature' is visible. A tabbed interface shows 'Study', 'Samples 80' (which is highlighted with a blue border), and 'EXP Expression'. Below the tabs, there's a 'Filters' section showing '35 samples:' with filters for 'Disease' and 'systemic lup'. A table with two columns, 'genestack:acc...' and 'Sample Source ID', contains sample data. At the bottom, there's a blue 'Edit' button and a note 'Published by Maria Borodaenko o'.

	genestack:acc...	Sample Source ID
1	GSF316728	P0772_CD14_1
2	GSF316729	P0772_CD14_2
3	GSF316730	P0772_CD19_1
4	GSF316731	P0772_CD19_2

Changing a Template

Prior to curating a study it is better to verify that the proper template is used.

The applied template is available in the study menu. It can be explored and changed if needed.

The screenshot illustrates the process of changing a template in the Genestack application. It features three overlapping windows:

- Multiplexed enrichment and genomic profiling**: The background window, which includes a 'VAR Variant' button and a study menu at the bottom.
- Study Menu**: A dropdown menu that is open, showing options like 'Share', 'Export data', 'Rename', 'Copy accession', 'Save to bookmarks', and 'More info'. The 'Maria Test Template' is highlighted, and a secondary menu is open next to it with options 'Explore' and 'Apply another'.
- Choose template**: A modal window that is open, displaying a list of templates. The 'Astex template' is currently selected and highlighted. Other templates listed include 'Template for BZ', 'Agri', 'Agriculture template', 'Antman', 'Copy of Copy of Default Template', 'Copy of Default Template', 'Copy of Default Template 2', 'Copy of Default Template 3', and 'Copy of Maria Test'. 'Apply' and 'Cancel' buttons are at the bottom of this modal.

Metadata Curation

Curators can bulk replace values and reassign attributes to make metadata valid.

Replace values

Attribute **Sex** not in the template

	Replace with
☒ F 72 ⓘ	female
☒ M 8 ⓘ	M

male

male
metafemale
unknown
preferred label of not determined
female

☐ Only filtered samples

Species ID	Type
Homo sapiens	T Species ID Not from template Copy values to... Bulk replace Reassign attribute Remove attribute
Homo sapiens	
Homo sapiens	
Homo sapiens	
Homo sapiens	
Homo sapiens	
Homo sapiens	

Reassign attribute

Species ID to

Metadata Versioning

- Once the changes are made study should be published.
- The history of published versions is available for each study.
- Any of previous version can be viewed and restored.

Version history ✕

[Restored] Initial version

Published by Maria Borodaenko on Aug 8, 04:03

Edited using curation script

Published by Maria Borodaenko on Jun 17, 16:46

Initial version

Published by Maria Borodaenko on Jun 17, 16:40

Study Design	No value
Therapeutic Area	Immuno-oncology
Edit	Published by Maria Borodaenko on Jun 17, 16:40

17	GSF333659	P1353_CD14_1	10031353	No value
18	GSF333660	P1353_CD14_2	10031353	No value
Back to latest version Restore Published by Maria Borodaenko on Jun 17, 16:40				

Curation Script

Curation script is a python script written for automating frequent operations in the curation process.

The main purpose of the script is to facilitate the curation process, reduce manual operations and reduce errors in metadata curation.

The script could be run on a particular study(ies). Changes done by the script are regulated by JSON file with predefined rules.

```
{
  "object_type": "sample",
  "raw_keys": ["Cell Type", "Cells"],
  "genestack_key": "Cell Type",
  "rules": {
    "CD19+": "B cell",
    "CD14+": "CD14-positive monocyte",
    "CD8+": "CD8-positive, alpha-beta T cell",
    "CD4+": "CD4-positive, alpha-beta T cell"
  }
},
{
  "object_type": "sample",
  "raw_keys": ["Sex", "Gender"],
  "genestack_key": "Sex",
  "dictionary": "Sex"
}
```

Task logs

[Error Log](#)[Output Log](#)

```
Found 82 files
2022-11-23 03:10:14 UTC - Starting to process 82 files in target folder
Creating new version for file GSF052045 with message Edited using curation script
Processing done in 13.10 seconds
Average editing speed: 7.09 files/s

=====
2022-11-23 03:10:27 UTC - Processed 80 files (average editing speed: 6.92 files/s)
COUNTERS:
- Cell Type
  - Successful mapping: 80
- Sex
  - Successful mapping: 80
```

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- Metadata curation
- **Access to studies**
- Study export
- Study deletion

Access to Studies

- The list of groups available in your profile.
- Curator group is a special group granting edit permissions.
- You can share your study with any group you are a member of.
- You can view with which groups a study is shared.

 **Maria Borodaenko** ADMIN CURATOR

maria.borodaenko@genestack.com

Groups: Curator, Maria test group

Change password

Edit personal information

Share study



Select group



Share

Already shared with:



Maria Borodaenko
Owner



All users
Group, can view, access



Maria test group
Group, can view, access

Sharing a Study

To share a study user should be the owner of a study – the one who created it in ODM.

Share option is available in the Study menu.

The screenshot illustrates the process of sharing a study in Genestack. At the top, a green box labeled "Study menu" points to a dropdown menu for a study titled "Multiplexed enrichment and genomic profiling of peripheral blood cells specific immune signature". The menu options include "Share", "Export data", "Rename", "Copy accession", "Save to bookmarks", "More info", and "Default Template". Below this, a "Share study" dialog is shown. It features a "Select group" dropdown and a "Share" button. Under the heading "Already shared with:", three entries are listed: "Maria Borodaenko" (Owner), "All users" (Group, can view, access), and "Maria test group" (Group, can view, access). A second, darker "Share study" dialog is overlaid on the first, showing a search bar "Search by name" and a list of options: "All users", "Curator" (highlighted), and "Maria test group".

Training Plan

- Getting started
- Exploring a Study in Metadata Editor
- Searching for available data in Study Browser
- Data loading
- Metadata templates
- Metadata curation
- Access to studies
- **Study export**
- Study deletion

Exporting a Study

The screenshot illustrates the process of exporting a study from the Genestack interface. It shows the 'Study menu' and 'Study export' options, along with the 'Export Data' page.

Study menu

- Share
- Export data
- Rename
- Copy accession
- Save to bookmarks
- More info
- Default Template
- Template

Study export

Export Data

Multiplexed enrichment and genomic profiling of peripheral blood cells reveal subset-specific immune signature

Study, size before compression is unknown

Exported on 2022-10-26 11:39:06, 2.06 MB ...

Y2F0aW9uL3ppcC9hL3QvZy9jLzIxNjMwMjE=/Exported_on_2022-10-26_11_39_06.zip [Download](#)

Authorization is not required to download with this link

Other exports

These archives are already prepared for export but might contain outdated data.

Exported on 2021-11-18 11:14... [Download](#) ...

To export a study with all related data click on the Export data either in the study menu or in Metadata Editor.

This will navigate you to the Export Data page where the link for loading an archive with available files will be created. Also, the previous exports are mentioned there.

Exporting a Study

The archive contains all available files and a ReadMe file with the description.

[accession]_file_description.s.txt contains the accessions of the exported files, their names and the descriptions of the files.

	GSF000422_file_descriptions.txt
	GSF000422_New Study__metainfo.csv
	GSF000423_AZ_proteomic..._samples_csv__metainfo.csv
	GSF000666_AZ_proteomics_demo1.gct
	GSF000666_AZ_proteomics_demo1.gct__metainfo.csv

```
--
Study title: New Study
Study accession: GSF000422
Exported on: 2021-08-31 13:26:12 UTC

This archive contains the following:

Omics data files:
GSF000666_AZ_proteomics_demo1.gct - Expression data file (as imported to ODM) in GCT
format.

Metadata files:
GSF000422_New Study__metainfo.csv - Study metadata (context of study etc.) table in CSV
format. The first row contains metadata field names and the second row contains values.
GSF000423_AZ_proteomics_demo1_samples_csv__metainfo.csv - Samples metadata (sample
biological attributes etc.) table in CSV format. The first row contains metadata field
names and subsequent rows contain values for each sample.
GSF000666_AZ_proteomics_demo1.gct__metainfo.csv - Expression metadata table (that
accompanies the expression data file of the same name) in CSV format. The first row
contains metadata field names and the second row contains values.
```

Exporting a Dataset

It is possible to export just one dataset of interest. For doing it navigate to a tab with analysis of interest and click on “Export dataset” button.

The data file with the corresponding metadata file and the description will be compress into archive.

Multiplexed enrichment and genomic profiling of specific immune signature

Study Samples 80 **EXP** Expression **VAR** Variant

Export dataset

genestack:accession GSF316814

Experimental Platform Illumina Genome Analyzer II

Preparations 80 **EXP** Expression

v. 2 ACTIVE v. 1 Export dataset

genestack:accession GSF051697

Experimental Platform Illumina NextSeq 550

Study Deletions

To delete a study “wipeStudy” method can be called either from the Browser console or via the python script. The method finds all entities linked to the study and delete them one by one.

brain, Parkinson's Disease, GTEx, any me

1 study found

Filtered by: Access Maria test group

Multiplexed enrichment and genomic profiling

Specific immune signature

s, blood, B cell, CD8-positive

+1 more, Healthy, systemic

borodaenko on Nov 16, 2021

Share

Export data

Copy accession

Save to bookmarks

More info

```
Elements Console  
top Filter  
> window.wipeStudy('GSF316726')
```

Prerequisites: The user should have “Manage organisation” permission.

Users and permissions + New user

All users

mar

4 users

Permissions

Manage organization Manage groups

Create and deactivate users

Change user passwords

Manage user permissions

Manage groups

User	Role	Permissions
Maria Borodaenko maria.borodaenko@genestack.com	ADMIN	☒ Manage organization ☒ Manage groups
Maria Lazarenko maria.lazarenko@genestack.com	DEACTIVATED	☐ Manage organization ☐ Manage groups

Input: Study accession.

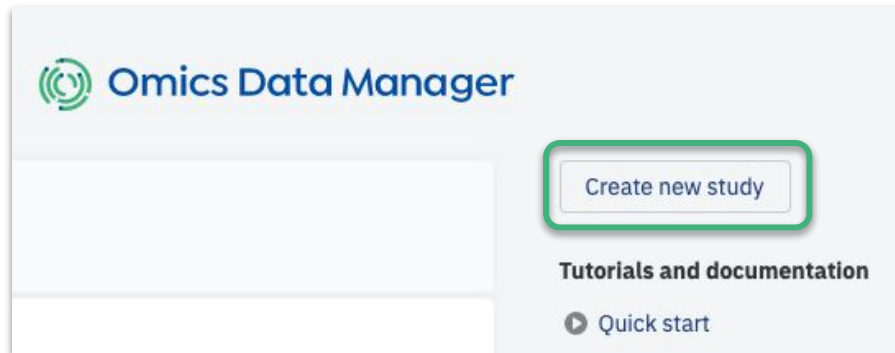
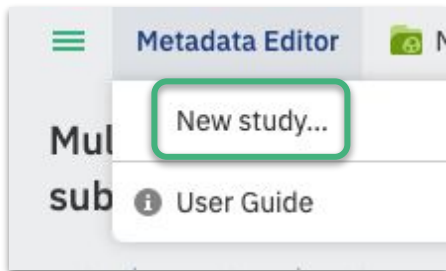
Study Design within ODM

Study Design

A new study can be created directly in ODM without the need of filling in Excel spreadsheets or tsv files. The benefits are:

- Always synced and up-to-date attributes
- Suggestions from the ontologies while filling metadata
- Immediate validation based on the template

New study can be created from Dashboard or from Metadata Editor menu.



Study Design

After clicking on New study, the user will be prompted to enter a name and choose a template.

A new study with the chosen name and blank attributes from the template will be created.

Metadata Editor New Study Tasks 6 Maria Borodaenko

New Study

Export

Study Samples 4 •

genestack:accession GSF052213

Study Source No value

New study

Name

Untitled Study

Choose template:

- ✓ Default Template Default
- Maria Test
- Maria's Template

Create

Adding Samples

A few blank samples will be created with the new study.

New samples can be created by “+” button or imported from .tsv file.

The image shows two overlapping UI elements. In the background is a table with a red error message 'Invalid metadata' at the top. The table has columns labeled F and G. Row 1 has a dropdown arrow and the text 'Disease'. Row 2 has a dropdown arrow and the text 'Age'. In the foreground is a 'New samples' dialog box. It has a text input field containing '25' with a small up/down arrow on the right. Below the input is the text 'Maximum 100 samples at a time'. At the bottom are two buttons: 'Add' (blue) and 'Cancel' (grey).

The image shows the 'New Study' interface. At the top, there's a dark header with 'New Study' and an 'Export' button. Below that, there are tabs for 'Study' and 'Samples 4', with 'Samples 4' being the active tab. A 'Filters' button is on the left, and an 'Invalid metadata' error message with a refresh icon and a '+' button is on the right. The main content area is titled 'Import samples metadata from TSV file'. It features a table with columns A, B, C, D, and an unlabeled column. The rows are labeled 1 through N. A large blue circular overlay with a white cloud and arrow icon and the text 'Select tsv file...' is centered over the table. At the bottom are 'Import' and 'Cancel' buttons.

	A	B	C	D	
	Sample Source ID	Sample Name	Organism	Sex	Dis
1					
2					
3					
4					
5					
6					
7					
8					
...					
N					

Congratulations



You are now
a trained ODM user!

Q&A





Thank You!

Maria Borodaenko
Senior Solution Consultant

maria.borodaenko@genestack.com

www.genestack.com



Unlocking the Power of Life-Sciences Data