

Supplementary Data S4: FAIR Genomes MOLGENIS walkthrough

The MOLGENIS FAIR Genomes app can be used by visiting the FAIR Genomes public demo, by running the Docker install, or by performing a regular server install. See: <https://molgenis.org> and <http://fairgenomes.org>.

After launch, you are taken to the homepage. Start by clicking on a module such as Sequencing.



The Sequencing table is now shown in the Data Explorer. Here, you can select which columns you want to see, and apply filters to reduce the number of rows. Click the green [+] to create a new row.

The screenshot displays the FAIR Genomes Data Explorer interface. At the top, there is a blue header with the FAIR Genomes logo and a menu icon. Below the header, the main content area is titled "sequencing" and includes a breadcrumb link to "FAIR Genomes metadata schema", a search input field containing "sequencing", and a red "Delete" button.

On the left side, there is a sidebar with a search bar and two main sections: "Data item filters" with a "Wizard" button, and "Data item" with a "Deselect all" button and a list of data items. The "Data item" list includes:

- Sequencing identifier
- Belongs to sample
- Sequencing date
- Sequencing platform
- Sequencing instrument
- Sequencing type
- Average read length
- Observed read length
- Observed insert size
- Percentage Q30
- Percentage T>30
- Other quality metrics

The main table area is titled "Data" and "Aggregates". It contains a table with the following columns: "Sequencing identifier", "Belongs to sample", and "Sequencing date". The table has 5 rows of data:

	Sequencing identifier	Belongs to sample	Sequencing date
<input checked="" type="checkbox"/>	Pat001seq	Pat001bloodsampleprep	Feb 2, 2021
<input checked="" type="checkbox"/>	U123456CF2021D12346		Apr 2, 2021
<input checked="" type="checkbox"/>	U123457CM2021D12347		Apr 2, 2021
<input checked="" type="checkbox"/>	U123458PF2021D12348		Apr 2, 2021
<input checked="" type="checkbox"/>	U123459PM2021D12349		Apr 2, 2021

Below the table, there is a "Rows per page:" dropdown set to "20" and a "Download" button. A status message indicates "5 items found".

A form view is shown where values for a new row can be entered. Click Save when you are done.

sequencing



Sequencing identifier *

A unique identifier assigned to raw data from a performed nucleic acid sequencing assay.

Belongs to sample

 ▾ +

Link to the prepared sample i.e. source that was sequenced.

Sequencing date

 📅 ✕

Date the sequencing run was performed.

Sequencing platform

 ▾ +

A sequencing platform (brand) is a name of a company that produces sequencer equipment.

Sequencing instrument model

 ▾ +

A product name and model number of a manufacturer's genomic (dna) sequencer.

Sequencing type

 ▾ +

Sequencing distinguishable as an identifiable class based on common qualities.

Average read depth

The average number of times a particular locus (site, nucleotide, amplicon, region) was sequenced.

Observed read length

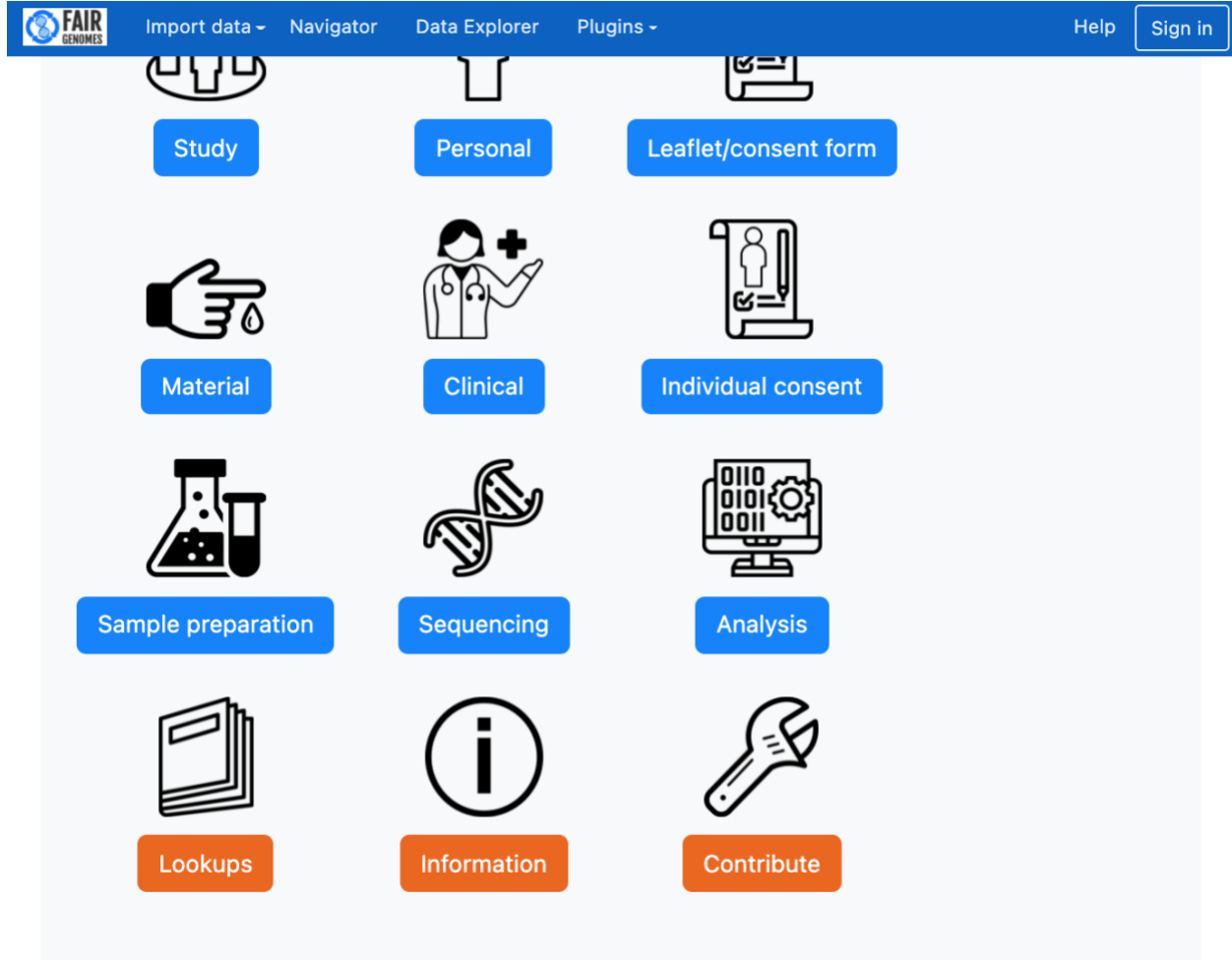
After saving, you are returned to the Data Explorer. The various modules in FAIR Genomes are linked to each other. You can click on hyperlinks to follow references in the database. For example, Sequencing is performed on a Sample. The details for the linked Sample can be viewed.

The screenshot displays the FAIR Genomes Data Explorer interface. A modal window titled "Belongs to sample" is open, showing a table with one row of data. The table has columns for "Sampleprep identifier", "Belongs to material", and "Input amount". The row contains the values "Pat001bloodsampleprep", "Pat001blood", and "100". Below the table, there is a "Rows per page" dropdown set to "20" and a "1 item found" status. An "Ok" button is visible at the bottom right of the modal. In the background, a table of sequencing data is visible, with columns for "Sample", "Sequencing", and "Date". The table shows five rows of data, each with a "Sample" ID, a "Sequencing" ID, and a date of "Apr 2, 2021". A "Download" button is located below the table. On the left side of the interface, there are sections for "Data item filters" and "Data item", both with expandable menus.

Sampleprep identifier	Belongs to material	Input amount
Pat001bloodsampleprep	Pat001blood	100

Sample	Sequencing	Date
Pat001seq	Pat001bloodsampleprep	Feb 2, 2021
U123456CF2021D12346		Apr 2, 2021
U123457CM2021D12347		Apr 2, 2021
U123458PF2021D12348		Apr 2, 2021
U123459PM2021D12349		Apr 2, 2021

You can find all of the FAIR Genomes lookups, i.e. underlying ontological code systems, by clicking Lookups on homepage. This will take you to the Navigator view with all database tables. These underlying data can be modified if necessary.



If you would like to perform a 'Google search', you can use the Search All function available via the Plugins menu.

 Search Clear

No packages found

Matching tables

clinical_clinicaldiagnosis <a>Show in navigator <a>Show in dataexplorer
Data <a>1 rows found
Columns <i>No matching columns found</i>
clinical_medication <a>Show in navigator <a>Show in dataexplorer
Data <a>21 rows found
Columns <i>No matching columns found</i>
clinical_moleculardiagnosisgene <a>Show in navigator <a>Show in dataexplorer
Data <a>40 rows found
Columns <i>No matching columns found</i>
clinical_phenotype <a>Show in navigator <a>Show in dataexplorer
Data <a>99 rows found
Columns <i>No matching columns found</i>
clinical_unobservedphenotype <a>Show in navigator <a>Show in dataexplorer

MOLGENIS offers powerful features such as altering the underlying data structure of the database itself. This can be done in the MetaData manager.

The screenshot displays the MOLGENIS Metadata Manager interface. At the top, a blue navigation bar contains the FAIR Genomes logo and menu items: Import data, Navigator, Data Explorer, Data Integration, Plugins, Admin, and Help. Below the navigation bar, the 'Account' section includes a 'Sign out' button. The main area is titled 'Metadata manager' and shows the configuration for an entity named 'analysis'. The configuration includes fields for 'Extends', 'Label', 'Description', and 'Package'. The 'Label' is set to 'analysis', 'Description' is 'Add a descriptive', and 'Package' is 'FAIR Genomes metadata schema'. The 'ID Attribute' is 'Analysis identifier', and the 'Label attribute' is also 'Analysis identifier'. The 'Lookup attributes' section shows 'Analysis identifier' as a selected attribute. Below the entity configuration, the 'Attributes' section is visible, showing a list of attributes on the left and the configuration for the 'Algorithms used' attribute on the right. The 'Algorithms used' attribute configuration includes fields for Name, Label, Description, Type, Parent, Default value, Computed value expression, Nullable expression, Visible expression, and Validation expression. The 'Name' is 'algorithmsused', 'Label' is 'Algorithms used', 'Description' is 'A defined procedure for', 'Type' is 'text', and 'Parent' is 'Select a compound attribute...'. The 'Computed value expression' is 'attributeX -', 'Nullable expression' is '\$('attributeX').isNull()', 'Visible expression' is '\$('attributeX').value() > 5', and 'Validation expression' is '\$('attributeX')

Metadata manager

analysis

Extends: Select an Entity...
 Abstract:
 Save all changes
 Undo all changes Delete entity

Label: analysis
 Description: Add a descriptive
 Package: FAIR Genomes metadata schema

ID Attribute: Analysis identifier
 Label attribute: Analysis identifier
 Lookup attributes: Analysis identifier

Attributes

- Analysis identifier
- Belongs to sequencing
- Physical data location
- Abstract data location
- Data formats stored
- Algorithms used**
- Bioinformatic protocol used
- Deviation from protocol
- Reason for deviation
- WGS guideline followed

Attribute: Algorithms used

Name: algorithmsused
 Label: Algorithms used
 Description: A defined procedure for
 Type: text
 Parent: Select a compound attribute...
 Default value: Add a default value...

Computed value expression: Example: attributeX -
 Nullable expression: Example: \$('attributeX').isNull()
 Visible expression: Example: \$('attributeX').value() > 5
 Validation expression: Example: \$('attributeX')

The menu structure of the user interface can also be adjusted to any needs.

Drag and drop menu items to update menu, press Save to store the menu. Each menu should contain at least one item.

Home

- Home
- Import data**
 - Quick data import
 - Advanced data import
- Navigator
- Data Explorer
- Data Integration**
 - Metadata Manager
 - Mapping Service
 - SORTA
 - Tag Wizard
- Plugins**
 - Search all data
 - API documentation
 - App manager
 - Feedback
 - Job overview
 - Questionnaires
 - Scripts
- Admin**
 - Log manager
 - Menu Manager
 - Permission Manager

Create Menu

Id *

Name *

Create Menu Item

Plugin *

Name *

Query string

Create

Create

Save the new menu layout

Upload logo

Choose No file chosen
file

Upload logo

MOLGENIS also offers a detailed permission system for groups, users, packages, entities, and rows. These options are available for power users.



Plugin Permissions

Package Permissions

Entity Class Permissions

Row-Level Security

Roles

Users

Select User: anonymous

security-ui	<input type="radio"/>	<input checked="" type="radio"/>
void	<input type="radio"/>	<input checked="" type="radio"/>
importwizard	<input type="radio"/>	<input checked="" type="radio"/>
searchAll	<input type="radio"/>	<input checked="" type="radio"/>
jobs	<input type="radio"/>	<input checked="" type="radio"/>
data-row-edit	<input type="radio"/>	<input checked="" type="radio"/>
home	<input checked="" type="radio"/>	<input type="radio"/>
questionnaires	<input type="radio"/>	<input checked="" type="radio"/>
background	<input type="radio"/>	<input checked="" type="radio"/>
appmanager	<input type="radio"/>	<input checked="" type="radio"/>
one-click-importer	<input type="radio"/>	<input checked="" type="radio"/>
metadata-manager	<input type="radio"/>	<input checked="" type="radio"/>

Save

*This database was created using the open source [MOLGENIS software](#) version 8.7.2 built on 2021-05-11 14:44 UTC.
Please cite [Van der Velde et al \(2018\)](#), [Swertz et al \(2010\)](#) or [Swertz & Jansen \(2007\)](#) on use.*

The home page can also be redesigned by clicking the Edit button on the homepage itself.



*This database was created using the open source MOLGENIS software 8.7.2 built on 2021-05-11 14:44 UTC.
Please cite [Van der Velde et al \(2018\)](#), [Swertz et al \(2010\)](#) or [Swertz & Jansen\(2007\)](#) on use.*

An advanced data import wizard allows uploading of bulk data using Excel or zipped TSV files.

1 Upload file > **2 Options** > 3 Packages > 4 Validation > 5 Result

Metadata options

- Create new metadata / update existing metadata**
Importer adds new metadata or updates existing metadata
- Ignore metadata**
Importer ignores metadata

Data options

- Add entities**
Importer adds new entities or fails if entity exists
- Add entities / update existing**
Importer adds new entities or updates existing entities
- Update entities**
Importer updates existing entities or fails if entity does not exist

[← Previous](#)

[Restart](#) [Next →](#)

Lastly, MOLGENIS offers a variety of APIs to connect to other systems and programming environments, including:

- Data API
- Metadata API
- REST api v1
- REST api v2
- Files api
- Import api
- Permission api
- Python-api client
- R-api client
- Beacon api
- FAIR api
- RSQL operators

Example response for retrieving FAIR Genomes Sequencing data:

```
JSON
{
  href: "/api/v2/fair-genomes_sequencing"
  meta: {
    href: "/api/v2/fair-genomes_sequencing"
    hrefCollection: "/api/v2/fair-genomes_sequencing"
    name: "fair-genomes_sequencing"
    label: "sequencing"
  }
  attributes: {
    labelAttribute: "sequencingidentifier"
    idAttribute: "sequencingidentifier"
  }
  lookupAttributes: {
    isAbstract: false
    writable: true
    languageCode: "en"
  }
  permissions: {
    start: 0
    num: 100
    total: 5
  }
  items: [
    0: {
      _href: "/api/v2/fair-genomes_sequencing/Pat001seq"
      sequencingidentifier: "Pat001seq"
      belongstosample: {
        _href: "/api/v2/fair-genomes_samplepreparation/Pat001bloodsampleprep"
        samplepreidentifier: "Pat001bloodsampleprep"
        sequencingdate: "2021-02-02"
      }
      sequencingplatform: {
        _href: "/api/v2/fair-genomes_sequencing_sequencingplatform/Illumina%20platform"
        value: "Illumina platform"
      }
      sequencinginstrumentmodel: {
        _href: "/api/v2/fair-genomes_sequencing_sequencinginstrumentmodel/Illumina%20HiSeq%20X%20Five"
        value: "Illumina HiSeq X Five"
      }
      sequencingtype: {
        _href: "/api/v2/fair-genomes_sequencing_sequencingtype/Whole%20Exome%20Sequencing"
        value: "Whole Exome Sequencing"
      }
    }
  ]
}
```