openBIS Tutorial

2020-09-23

## Table of Contents

[1.1 Table of Contents 1](#_Toc51603751)

[2 Overview 2](#_Toc51603752)

[3 Registration of samples and protocols in the Inventory 3](#_Toc51603753)

[3.1 Registration of new materials and samples in the Inventory 3](#_Toc51603754)

[3.1.1 Registration of tissue samples 3](#_Toc51603755)

[3.1.2 Copy of tissue sample 4](#_Toc51603756)

[3.1.3 Batch registration of chemical samples 4](#_Toc51603757)

[3.1.4 Deletion of duplicate Objects 5](#_Toc51603758)

[3.1.5 Batch modification of chemical samples 5](#_Toc51603759)

[3.2 Visualization of storage positions with the Storage Manager 6](#_Toc51603760)

[3.3 Registration of a new protocol in the Inventory 6](#_Toc51603761)

[3.3.1 Registration of protocol for RNA extraction 7](#_Toc51603762)

[4 Laboratory Notebook 8](#_Toc51603763)

[4.1 Registration of a Project 8](#_Toc51603764)

[4.2 Registration of an Experiment 9](#_Toc51603765)

[4.3 Registration of Experimental Steps 9](#_Toc51603766)

[4.3.1 Registration of RNA extraction experimental step 9](#_Toc51603767)

[4.3.2 Register RNA extracts in inventory as children of Experimental Step 11](#_Toc51603768)

[4.3.3 Register RNA sequencing Experimental Step 11](#_Toc51603769)

[4.3.4 Data upload to an Experimental Step 12](#_Toc51603770)

[4.3.5 Data visualization 12](#_Toc51603771)

[4.3.6 Data and metadata exports 13](#_Toc51603772)

[**4.3.7** Data analysis Entry registration 13](#_Toc51603773)

[4.4 Access rights assignment to a personal notebook or to a single Project 14](#_Toc51603774)

[5 Searching the ELN 15](#_Toc51603775)

[5.1 Generic text search 15](#_Toc51603776)

[5.1.1 Saving and re-using searches 15](#_Toc51603777)

[5.2 Data Set Files search 16](#_Toc51603778)

[6 Freezing entities 16](#_Toc51603779)

[7 Jupyter Notebook 16](#_Toc51603780)

# Overview

In this tutorial we will learn how to use the openBIS inventory and lab notebook.

We will work on a RNA sequencing study of 8 different dog breeds, which aims at identifying genetic intra-breed differences in reference to the Beagle dog. This is the most commonly used dog in pre-clinical studies.

RNA stands for RiboNucleic Acid. RNA is essential in many biological processes, such as coding, decoding, regulation and expression of genes.

The study involves RNA extraction from dog’s tissue samples, RNA sequencing and analysis of the sequenced data. We will see how to capture the information and data generated in each of these experimental steps into openBIS.

The tutorial covers the registration of samples, protocols and experiments:

1. Registration of chemicals, tissues, RNA extracts samples
2. Registration of RNA extraction protocol
3. Description of experiments in the lab notebook
4. Data upload
5. Searching the ELN
6. Freezing entities
7. Jupyter Notebook

We will see different ways of registering and updating samples and we will see how to keep track of connections between entities in the system.

Additional documentation and video tutorials can be found at: <https://openbis.ch/>.

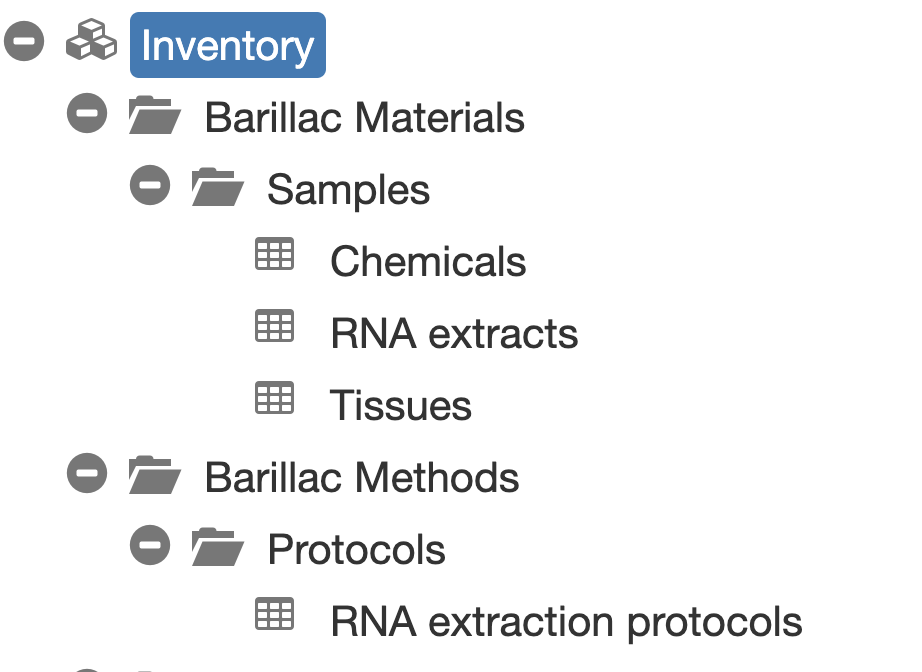
# Registration of samples and protocols in the Inventory

This part of the tutorial covers the registration of new samples and protocols in the lab inventory. The lab inventory is usually shared by all lab members.

## Registration of new materials and samples in the Inventory

We will start by registering a few samples in the **Materials** inventory: *tissues* and *chemicals*. Topics covered:

1. How to register single samples
2. How to copy samples
3. How to batch register samples
4. How to batch update samples
5. How to assign storage positions to samples



**SESSION 1**

### Registration of tissue samples

We will now register a few tissue samples in the **Materials** Inventory. These are the dog tissue samples (liver and kidney) from which we want to extract RNA. We will also assign storage location to the samples.

1. In the **Tissues** folder, inside your **Samples** folder (in **Username Materials**), register a kidney sample with the following specs:
2. **Sample ID**: 12345
3. **Tissue type**: Kidney
4. **CRO**: Covance Maddison
5. **Breeder**: Covance Research Products
6. **Storage location**: -80°C freezer, position 1,1

Note: Remember to **Save** the sample after filling in the form.

1. In the **Tissues** folder, inside your **Samples** folder (in **Username Materials**), register a liver sample with the following specs:
2. **Sample ID**: 12346
3. **Tissue type**: Liver
4. **CRO**: Charles River
5. **Breeder**: Marshall (North Rose, NY)
6. **Storage location**: -80°C freezer, position 1,2, position A1 in a 9x9 box. In the B**ox name** field, enter username\_box (so if your username is KG12345, you would enter KG12345\_box)

After registering the two samples, click on the **Tissues** folder. From the **Columns** drop down in the table you can choose which fields to show in the table. Select **Sample ID, Tissue type**, **CRO**, **Breeder**. This information is stored per user, so the selection only has to be made once.

### Copy of tissue sample

We now want to register a third sample, which is a replica of **sample 12345**. All fields are the same, but the sample ID is different. Instead of entering all the information again from scratch, we can copy sample 12345 and change the sample ID to 12347.

Procedure:

1. Select the sample with ID **12345** from the table in the Tissues folder (note that fields in blue, such as Name/Code, are links, so they can be clicked)
2. The **Copy** functionality is in the **More…** dropdown menu
3. After copying the sample**, Edit** the copied sample and change the **Sample ID** to **12347**

**SESSION 2**

### Batch registration of chemical samples

Now we want to register 8 chemicals. We will do this by *Batch Registration*, using a .tsv file. For batch registration it is necessary to use the template file provided by openBIS. We will use the file **SC-template-READY.tsv** as reference: this file contains 8 chemicals, for each of which the name, supplier and article number are specified.

Procedure:

1. Go to the **Chemicals** folder in the **Samples** folder.
2. Select **Batch Register Objects** from the **More…** drop down menu.
3. Select **Chemical** from the **Object Type** dropdown menu.
4. Download the **Template** file (**SAMPLE-CHEMICAL-template.tsv).**
5. Open the file with Excel (*Numbers* and *LibreOffice* can also be used).
6. Remove all comments lines (the first 7 rows of the files).
7. Remove the first column (=**identifier**) from the file. We do this, so that identifiers will be automatically generated by openBIS. This is not mandatory. It is possible to provide your own identifiers (example of openBIS identifier: /BARILLAC\_MATERIALS/SAMPLES/CHE1).
8. Open the file **SC-template-READY.tsv.**
9. Copy and paste the corresponding fields from **SC-template-READY.tsv** to **SAMPLE-CHEMICAL-template.tsv** *($Name, Supplier, Article number*).
10. Save the file **SAMPLE-CHEMICAL-template.tsv.** You can simply use “Save” in Excel and it will be saved as .tsv. If you encounter issues with this, you can alternatively use the “Save as” option and save the file as tab separated file (.txt).
11. Upload the file you just saved to the ELN (*Batch Register Objects -> Chemical*).
12. **Accept.**
13. Select **Name/Code**, **Art. Number, Supplier** from the **Columns** drop down in the table.

### Deletion of duplicate Objects

In the step above we registered 8 chemicals, but two of them are the same. We want to delete one of them:

1. Go to the **Chemicals** folder in **Samples** in the main menu
2. Select the entry **Rneasy Plus Mini kit** from the table, using the tick box in first column
3. Select **Delete selected** from the **Exports and..** drop down in the table
4. Enter **duplicate** in the **Reason** field
5. **Accept**

When Objects (and also Experiments/Collections) are deleted, they are first moved to the trashcan. In order to be completely removed from the database they have to be deleted also from the trashcan. Alternatively, it is also possible to revert deletion from the trashcan, if something was accidentally moved here.

To delete entries from the trashcan:

1. Go to the **Trashcan** under **Utilities** in the main menu
2. Select **Delete Permanently** from the **Operations** drop down in the table
3. Read the warning message!
4. **Accept**

Note: only admins can delete!

### Batch modification of chemical samples

When we registered the chemicals before, we forgot to enter the information about the storage conditions. All samples should be stored at room temperature. To add this info to all samples, instead of editing each single sample, we can use the **batch update** functionality.

Procedure:

1. Select the **Chemicals** folder in the **Samples** folder
2. Select only **Identifier, Name/Code** and **Storage Conditions** from the **Columns** dropdown in the table
3. Select **Export visible columns with all rows** from the **Exports and..** dropdown list
4. Open this file with Excel:
   1. enter **RT** under the Storage column in all rows
   2. **Save** the file (always as .tsv)
5. Go back to the ELN
6. Select **Batch Update Objects** from the **More..** dropdown
7. Select **Chemical** from the **Object Type** list
8. Upload the previously modified file
9. **Accept**

Note: If you encounter issues due to file formats, you can try to upload this file available in the training material: **chemicals-batch-update.tsv.** You will need to change the **identifier** column and replace the provided identifiers with the identifiers of your chemical samples (/USERNAME\_MATERIALS/SAMPLES/CHE*n*).

The storage info has now been added to the chemicals.

Storage conditions is a **Controlled Vocabulary** in openBIS, i.e. a list of pre-defined values. The list of available Controlled Vocabularies is available under **Utilities** **-> Vocabulary Browser.** By clicking on one Vocabulary, the list of terms can be visualized. Vocabulary terms have **Codes** and **labels**. The user interface displays the **label**, but when you use batch upload/update, the **Code** of the Vocabulary term needs to be entered in the template file. This is the reason why we enter **RT** in the file, but then we see **room temperature** in the user interface (**RT** is the code of this vocabulary term and **room temperature** is the label of this vocabulary term).

**SESSION 3**

## Visualization of storage positions with the Storage Manager

The openBIS storage manager offers an overview of all storages configured for the lab. Before we stored some tissue samples in the -80°C freezer. If we want to check all that is stored in the -80°C freezer, we can do the following:

* 1. Select **Storage Manager** under **Utilities**
  2. Select the **-80°C freezer** storage from the **Storage** list
  3. Go with the mouse over one of the compounds to see the information related to it

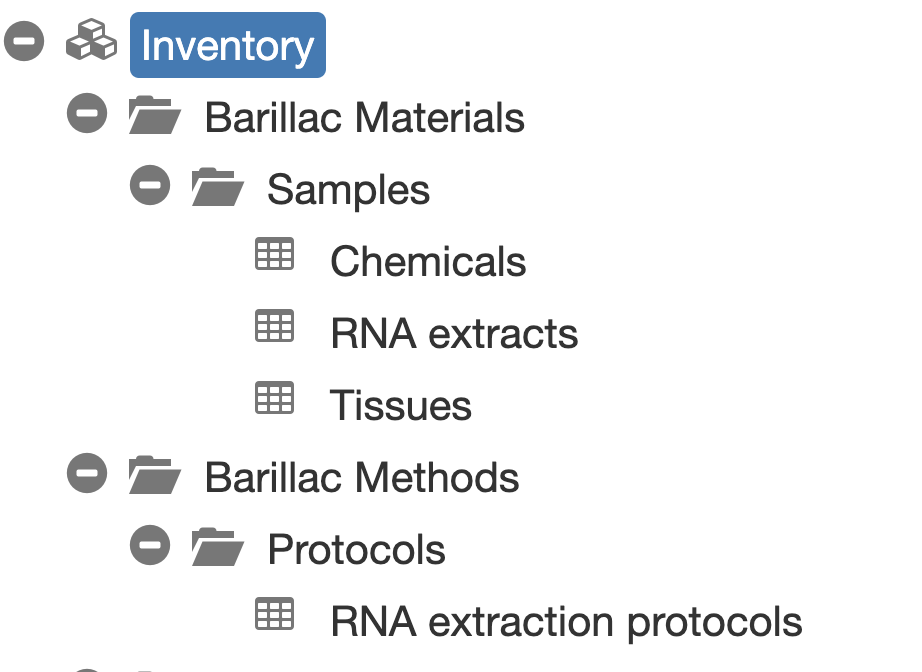
If we have boxes in the freezer, we can click on the box and the content of the box will be displayed.

It is possible to drag & drop boxes to change the position of a box inside a storage (or to a different storage). In the same way the position of a sample inside a storage or a box can be changed (note that this is not supported for multiple positions for the same sample).

**SESSION 4**

## Registration of a new protocol in the Inventory

We will now register a protocol for RNA extraction in the Methods Inventory. This is a standard procedure to follow every time an RNA extraction experiment is performed. We want to keep it in the common lab inventory, so every lab member can access it when needed.



### Registration of protocol for RNA extraction

We provided an example protocol document in the training material: **RNA\_extraction\_protocol.docx.** We want to enter the information contained in this document in openBIS, in our **RNA extraction protocols** in the **Methods** folder.

Steps:

1. Open a **New General Protocol** in the **RNA extraction protocols** folder (in **Username Methods -> Protocols**)
2. Copy and paste the content of the **RNA\_extraction\_protocol.docx file** in the **Name** and **Procedure** fields of the General Protocol form in the ELN.

**SESSION 5**

# Laboratory Notebook

In the Laboratory Notebook, usually each lab member has a personal folder (=*Space*) to organize *Projects* and *Experiments*. *Experiments* can be further divided in *Experimental Steps*. An openBIS *Experiment* is a specific scientific question and it contains *Experimental Steps* that are individual attempts to answer it.

*Experimental Steps* can be linked to each other (if needed) and they can also be linked to materials and methods stored in the Inventory.

In this tutorial we are working on an example project study of intra breed genetic variability of dogs. The study is based on the analysis of RNA sequencing data of different dog’s tissues. RNA is extracted from the dog’s tissues; the obtained samples are sequenced and the data obtained from sequencing is analyzed.

To model this in openBIS we can register 1 *Project*, 1 *Experiment* and 3 *Experimental Steps* in the lab notebook:

1. **Project**: *Intra breed genetic variability of dogs*
   1. **Experiment**: *RNA analysis of 8 dog breeds*
      1. **Experimental Step 1**: *RNA extraction from dog tissues*. In this step we want to establish links to tissue samples stored in the inventory, to the RNA extraction protocol and to the chemicals stored in the inventory. This *Experimental Step* generates RNA extract samples, which we will register in the Inventory of Materials as “children” of this step.
      2. **Experimental Step 2**: *Sequencing of previously extracted RNA*. In this step we want to establish the connection to the RNA extract samples produced in the previous step. We will upload an example fasta file that we would obtain from sequencing. In reality, sequencing data are fairly large and several hundreds of files are generated from a single sequencing experiment.
      3. **Experimental Step 3**: A*nalysis of sequencing data*. In this step we want to give an overview of the different methodologies used to analyze the data and we want to establish the connection to the previous Experimental Step (child).

## Registration of a Project

We start by registering our project, **Intra breed genetic variability of dogs**. Please note that Projects only have two fields:

1. **Code**: this can only take alphanumeric characters and no spaces are allowed
2. **Description**

In **your own folder** in the **Lab Notebook** register a project with the following specs:

**Code: Intra\_breed\_genetic\_variability\_of\_dogs**

**Description: This project aims at understanding intra breed genetic variability of dogs with special reference to Beagle dog, since Beagles are used as an animal model for compound testing in the pharma industry. The eight different dog breeds used are**:

1. **Beagle**
2. **GSD**
3. **Golden Retriver**
4. **Terrier**
5. **King Charles**
6. **Poodle**
7. **Rottweiler**
8. **West Highland White Terrier**

Note: If after saving, you have a blank page, go to the **More..** dropdown and select **Show Description**.

## Registration of an Experiment

In the Project registered before, we want to register our Experiment where we will record everything related to the analysis of the RNA of 8 dog breeds. This experiment is divided into several steps, which will be registered later.

In the **Intra breed genetic variability of dogs** Project, register a new **Default Experiment** with the following specs:

1. **Name**: **RNA analysis of 8 dog breeds**
2. **Show in project overview:** check**.** This is a way to mark important experiment, so that they are shown in the Project page. Usually this would be done at the end of the Experiment.
3. **Start date:** **today’s date**
4. **Experimental goals:** **Understand intra breed genetic variability of dogs with special reference to Beagle dog.**
5. **Publication**: **Lindblad-Toh K et al. (2005). Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438, 803–819**.

Note: You can use the **Publication** field to enter the name of the paper where this work was published. You can use the **References** field to add all references you used in this particular experiment (if any).

## Registration of Experimental Steps

### Registration of RNA extraction experimental step

Now we want to register an **Experimental Step**, which is the RNA extraction experiment. This is the first step in our main RNA analysis Experiment. In this experiment we use some tissues and chemical samples stored in the inventory and we follow the procedure explained in the RNA extraction protocol also stored in the Inventory. We want to create links between the Experimental Step we write and the entities in the inventory.

In the **RNA analysis of 8 dog breeds** Experiment, register a new **Experimental step** with the following specs:

1. **Name**: **RNA extraction**
2. **Show in project overview:** check
3. See if you can add the parents by yourself. If you find this difficult, follow the steps provided below.

Add the following Parents:

* 1. The 3 tissue samples registered before in the inventory
  2. The 7 chemicals registered before in the inventory
  3. The RNA extraction protocol registered before in the Inventory.

**Remember to save at the end.**

These are the steps to add Parents:

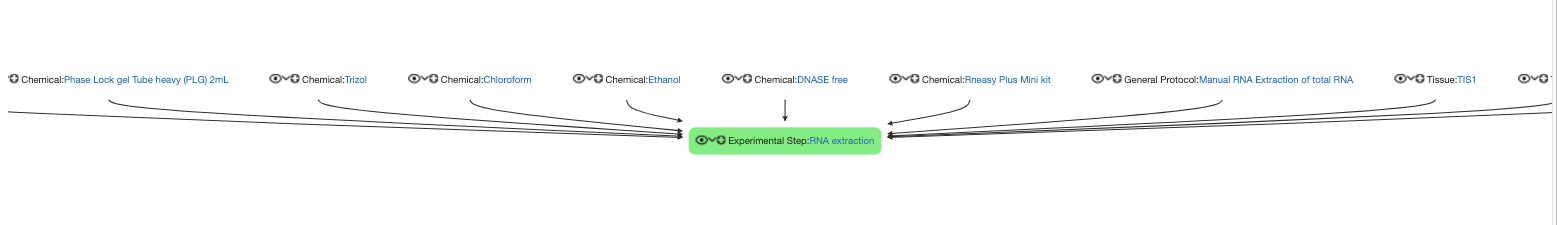
1. Add tissues:
   1. Click the **+** button next to **Parents**
   2. Select **Tissue** from the list
   3. Show the **Sample ID** in the table (Columns dropdown)
   4. Select the 3 tissues available in the table from the tick-box on the left hand-side
   5. Select **Add Selected** from the **Exports and…** dropdown
2. Add chemicals:
   1. Click the **+** button next to **Parents**
   2. Select **Chemical** from the list
   3. Check all chemicals in the table
   4. Select **Add Selected** from the **Exports and…** dropdown
3. Add protocol:
   1. Click +next to **General** **Protocols**
   2. Select the **RNA extraction** protocol by clicking on the table row

Note on protocols

When we link a protocol to an Experimental Step we have two options:

* + 1. Create a link to an existing protocol in the inventory. This is suitable when the protocol is followed “as it is”. This is what we did above.
    2. Copy the protocol to our Experiment folder in the lab notebook in order to modify it. This is suitable when the main protocol is modified during the Experiment. To do this, after adding the protocol as Parent, choose **Use as Template** from the **Operations** dropdown in the table.

You can visualize the connections of this Experimental Step to the samples and protocols by selecting the **Hierarchy graph** in the **More..** dropdown.



**SESSION 6**

### Register RNA extracts in inventory as children of Experimental Step

In the Experimental Step above we extracted RNA from tissue samples. Now we want to register the RNA extract samples in the Inventory and link them to the above Experimental Step. In this case we will only register one sample, as an example.

In the **RNA extracts** folder in the **Inventory** (under **Username** **Materials -> Samples**), register an RNA extract sample with the following specs:

1. **RIN**: **7**
2. **Parent**: the **RNA Extraction** Experimental Step registered before (try to do this by yourself, or follow the steps below)

Steps to add the Experimental Step as parent:

1. Click the **+** next to **Parents**
2. Select **Experimental Step**
3. Click on **RNA extraction**

### Register RNA sequencing Experimental Step

We now want to sequence the RNA samples we extracted before. This is a second Experimental Step in our main RNA analysis Experiment.

In the **RNA analysis of 8 dog breeds** Experiment, register a new Experimental Step with the following specs:

**Name**: **Sequencing read of RNA extracts**

**Description**:

**Illumina HiSeq2000 used**

**Library generated following standard protocols**

**Paired end-reads data generated**

1. **Parents**: the RNA extract sample we registered in the step above (try to do this by yourself, or follow the steps below)

**Spreadsheet** (you can copy/paste):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Sample ID | Dilution ng/ul | Bioanalyzer | RIN | Electrogram |
| Liver 1004 | 1999.43 | Y | 8.1 | OK |
| Liver 1054 | 3677.3 | Y | 8.3 | OK |
| Kidney 1004 | 989.47 | Y | 8.1 | OK |
| Kidney 1054 | 2568.75 | Y | 7.9 | OK |

Steps to add one RNA extract sample as parent of this Experimental Step:

1. Click the **+** button next to **Parents**
2. Select **RNA extract** from the list
3. Click on the only sample in the table (click on the table row)

### Data upload to an Experimental Step

Data can be uploaded to *Experiments* or *Experimental Steps* using the same procedure.

Now we want to upload one example .fasta file to the sequencing Experimental Step we just registered.

In the **Sequencing read of RNA extracts** Experimental Step, upload the **dog\_read1.fasta** file from the training material with the following specs:

1. **Data Set type**: **Raw data**
2. **Name**: **RNA seq data**

Note: In the same way, files can also be added to *Objects* in the Inventory (i.e. to samples or protocols).

**SESSION 7**

### Data visualization

To open data files stored in openBIS we recommend to use software that allow to mount openBIS as a drive on your computer. Examples are:

1. Mountain Duck for MacOS ([https://mountainduck.io](https://mountainduck.io/))
2. NetDrive for Windows (<https://www.nsoftware.com/sftp/netdrive/>).

Files can be opened with the desired application in read-only mode.

Any other FTP solution can also be used (e.g. *Cyberduck, Filezilla*, etc).

Please note that **data files stored in openBIS are** **read-only**!

### Data and metadata exports

It is possible to export a complete lab notebook or only parts of it *(Project, Experiment, Experimental Step, Dataset*s).

In each folder, **the More..** dropdown has an option to **Export metadata only** or **Export metadata & data**. We recommend to export data only if they do not exceed a few GBs.

When you export something, you will receive an email with a link to download your metadata (+data).

The export contains folders with the same structure organization they have in openBIS. Metadata are exported to 4 different file formats*: .docx, .html, .txt, .json*.

**SESSION 8**

### Data analysis Entry registration

After sequencing our RNA samples, we process and analyze the sequencing data. In the lab notebook we want to briefly describe the analysis tools we used. For this, we will use the **Entry** entity in openBIS, which is a form without predefined fields. This can be used as alternative to the Experimental Step, if you prefer to work with unstructured forms.

In the **Sequencing read of RNA extracts** Experimental Step, register a new **Entry** with the following specs:

1. **New Title:** **Summary of data analysis**
2. **Body of text**:

**Sequencing pipeline used for data processing and analysis:**

* **BWA: Read mapping, combine read pairs, convert to sam**
* **Samtools: Sam to bam conversion, sorting and indexed bam**
* **Picard: Merging lanes, removing duplicates**
* **IGV                  : Read coverage visualization**
* **Bedtools**
* **DESeq2**
* **Tophat2**
* **RStudio**

1. Drag & drop the file **data-analysis-results-table.png** from the training material in the page, after the description

Remember to **save** the form.

## Access rights assignment to a personal notebook or to a single Project

It is possible to grant collaborators and colleagues access rights to your complete lab notebook or only to selected Project(s). These users need to be already registered in openBIS by an admin. The username of the user needs to be used.

This can be done from the **Manage access** from the **More..** dropdown in the lab notebook folder or in a project folder.

Available roles are:

* 1. **Observer**: has read-only access
  2. **User**: can create and modify entities, but cannot delete anything
  3. **Admin**: can create, modify and delete entities

**Exercise:**

Grant **User** access rights to your complete lab notebook folder to user **hluetcke.**

**SESSION 9**

# Searching the ELN

openBIS offers 3 options for searching:

1. **Global search** on all metadata stored in the database. Searches can be refined using the **Advanced search**.
2. **BLAST** search for sequence comparison across sequences stored in the database.
3. **Data Set Files** search. This should be used to search for files uploaded as datasets by their name.

## Generic text search

We want to find all the RNA extraction Experimental Steps registered by you. We start with a generic search for *RNA* and then we restrict the results.

1. Enter **RNA** in the **Global search** field, on top of the main menu. This search returns a few results. Now we want to narrow down the search.
2. Select **Object** in the **Search For** dropdown
3. Click the + button in the **Criteria** table
4. Select **Property** from the **Field Type** dropdown
5. Select **Object Type** from the **Field Name** dropdown
6. Enter **\*STEP\*** in the **Field Value** field (This is for searching on types containing “STEP” in their name)
7. Click again the + button in the **Criteria** table
8. Select **Property** from the **Field Type** dropdown
9. Select **Registrator** from the **Field Name** dropdown
10. Enter **your username** in the **Field Value** field
11. Run the search

In this case, the search returns two items, one of which is the RNA extraction Experimental Step. In a similar way, you can run complex searches by combining multiple search criteria.

### Saving and re-using searches

To save a search in your Project folder:

1. Select **Save** on top of the page
2. Enter **demo search** in the **Name** field
3. Enter **Queries** in **the search entity to store** **query** field
4. **Save**

To run a saved search:

1. Refresh the **Advanced search** page in the browser to clear the search
2. Select the **demo search** from the list of saved searches
3. Run the search

## Data Set Files search

We want to search for a file that contains the word “dog” in the name.

1. Enter **dog** in the **Global search** field, on top on the main menu
2. Select **Data Set File** from the dropdown next to it

The search returns the dataset that contains the .fasta file we uploaded before to our sequencing experiment.

Note: it is not possible to search the content of files stored in openBIS.

**SESSION 10**

# Freezing entities

In openBIS it is possible to “freeze” entities that should no longer be modified. When freezing one entity, everything connected to it can be frozen too. The user needs to select what should be frozen and what should not.

Please note that freezing is **IRREVERSIBLE**!

At all levels of the lab notebook and inventory, there is a **Freeze Entity** option in the **More..** dropdown.

When you choose this option, you are presented with a list of things connected to the chosen level that can also be frozen. By default, everything is selected. After the selection, you need to enter your password to be able to freeze entities. This is necessary to prevent accidental freezing of entities.

# Jupyter Notebook

openBIS offers the possibility to analyse data stored in openBIS with Jupyter notebooks, which can in turn be saved in openBIS.

It is possible to use a local installation of Jupyter and download the openbis extension: <https://pypi.org/project/jupyter-openbis-extension/>. This allows to connect to openBIS and download and upload data from/to it.

When a JupyterHub server is in place and connected to openBIS (needs to be done by a system admin), at all levels of the lab notebook and inventory, there is anoption to launch Jupyter notebooks in the **More..** dropdown. This is available with this training server and you can try this out. Additional info: <https://openbis.ch/index.php/docs/user-documentation-19-06-4/tools-for-analysis-of-data-stored-in-openbis/jupyter-notebooks/>

This topic is covered more in details in a separate workshop.