Annotating Metabolomics analysis with STATegra EMS.

Experimental design

The Experiment is compound by 2 different conditions (treated and untreated sample), for each condition 3 samples (biological replicates) were isolated at different time-points in batch mode. Finally for each biological replicate, 2 analytical samples were isolated and used later for both GC-MS and LC-MS measurements.

Figure 1 Experimental design scheme
Annotation of Sample information

At Home panel choose the option Browse Samples. This option will show a list of all the annotated samples in the system grouped by Biological condition. From this panel users can inspect, edit and annotate sample information.

The STATegra EMS offers 2 ways for Sample annotation: manually, filling out the appropriate forms, or by using XLS files provided by the application. In this tutorial we will show how to record samples manually.

Click at the Add new Biological Condition button, you will see the Sample Module input form. The first of four sections of this form includes fields for General information on the sample such as sample ID (auto-generated), sample name (e.g. K562 or GM12878) and title, which is a more extended description of the sample (e.g. Chronic myelogenous leukemia / Lymphoblastoid). The second section of the form is used to describe the Biomaterial (e.g. lymphocyte cell type, blood tissue, human, normal or cancer variation...). At the Experimental condition section we describe the treatment’s particular details. The next section records the associated Biological Replicates, the corresponding batch, when applicable, and its derived analytical samples.
For this use case, we can differentiate between 6 Biological Conditions:

- treated time-point 0h,
- treated time-point 6h,
- treated time-point 12h,
- untreated time-point 0h,
- untreated time-point 6h,
- untreated time-point 12h.
And for each biological condition we will annotate 3 biological replicates, the corresponding batch information, and 2 Analytical samples per each biological replicate (characterized by an extraction protocol and a name) as follows.

1. Starting with the Biological Condition *treated time-point 0h*, fill the Sample form fields.

2. At the Biological Replicates section, annotate the information about Biological Replicates. For each biological replicate do:
   a. Click at *Add new Biological Replicate* button. This will open new dialog with the Biological Replicate form. Fill the replicate name field and at the *Batch options* menu (left-top corner) choose the option *Change the associated batch*.
b. In the new dialog you can register a new batch process or choose a previously annotated one. In this case, choose *Register new Experimental batch*. This option opens a new form for Batch process annotation. Fill the form and save the changes. The new batch process will be show at the Batch processes list, select it and click at *Use selected Experimental Batch*. 
c. After that, annotate the isolated Analytical Replicates from the current Biological Replicate. Click at Analytical Sample options menu and choose Add more Analytical Samples. At the new dialog, for each Analytical Sample (2 for this use case) fill the Analytical Sample Name column and choose the Protocol followed for the sample isolation. New protocols can be annotated using the option
Add/Remove/Edit protocols located at the left-bottom corner of the dialog. Add a new Protocol entry describing the followed process to Analytical Sample isolation and save the changes. Then fill the column for Analytical Sample Name and choose the new protocol at the Protocol List column.
d. Click at Accept and the new Analytical Samples will be added to the current Biological Replicate. Finally, accept the changes and the new Biological Replicate will be added to the current Biological Condition.
3. Repeat the process for the remainder Biological Replicates and save the changes.

4. Close the Sample Form and return to the Biological Conditions list panel. The new Biological Condition (BC) should be listed. For adding the next Biological Condition, select the previously added BC and choose the option Copy Selected. This option will create a new Biological Condition copying the selected one, including
the information about Biological replicates and Analytical samples. Edit the information to adapt to the new biological conditions.
Annotation of Analysis pipelines

Once all the samples are annotated, we can continue with the annotation of the Analysis performed during this example experiment.

As we isolated 2 Analytical samples for each Biological Replicate (one for GC-MS and one for LC-MS), we will annotate 36 different pipelines as follows.

1. First, at main menu choose the option “Browse Analysis” and click at Register new analysis option. At the new dialog choose the new Analysis type (i.e. Metabolomics)

2. After that, we can start annotating steps for our pipeline.
3. First, click at Annotate new step and choose the option New Raw Data step. This option opens a new dialog with the form for the annotation of Raw data acquisition steps. There are 4 main fields that must be specified:
   a. The Analytical Sample that was measured (and previously annotated at the sample’s module).
   b. The Step name, necessary to identify the step uniquely.
   c. File location, specifies where are located the files resulting files (if any).
   d. Technology, describes the technology used for data acquisition such as RNA-seq, Chip-seq, GC-MS, LC-MS...

![Raw data annotation form]

4. For Analytical Sample selection, click at the browse button 📦. This button opens a new window listing all the biological conditions annotated in the system. Select the biological condition to which the Analytical sample belongs to and press Next. The next panel shows all the Analytical samples annotates under the selected Biological condition, choose the appropriate sample and press Accept.
5. Fill the name and the file location fields and choose the extraction technology (i.e. GC-MS), depending on the selected technology, the form will be extended with new fields specific for each omic family.
6. Fill the form and press Accept. After that, the new step will be drawn at the pipeline diagram. The information for the annotated steps can be edited and consulted any time using the toolbar options or the contextual menu at the diagram (right mouse button click over a diagram node).
7. Now, you can just save the new Analysis or continue adding new steps until complete the pipeline. The new analysis will be listed at the Analysis browse panel.

8. Annotate the remaining Analysis. Note that in case of having similar pipelines, using the Copy Selected option we can easily clone the annotated analysis and adapt the information to the specific pipeline requirements.