

Phenotype Database

Quick start user manual for entering study (meta)data in
the Generic Study Capture Framework & Assay Modules

Last updated: September 2014

How to store your study information in a quick manner in Phenotype Database

This is a quick start manual showing users the required steps to create a study in Phenotype Database, from uploading of the subjects to the attachment of measurements to samples from those samples. For a full description on the principle of Phenotype Database users are referred to the more in-depth user guide available on the website.

Of note should be that it is easier for the user to:

- Create templates for each part of the study before starting with the actual data upload as fields in the templates represent data to be entered. Creating them beforehand will enable the user to directly match their data columns to the particular template fields (see chapter 3 of the user manual for description on templates).
- Prepare files to be uploaded before actually starting with data entry; files should adhere to a specific structure to allow matching of columns headers to the desired template field (see paragraph 2.3 of the user manual for some examples).

Example walkthrough: creating a study with a sampling-event and measurements.

I: Generic study capture framework section: create study, templates, (sampling-)event and assay

II: Module section: creating platforms and features

III: Adding numerical results/measurements to an assay

I: Generic study capture framework section: create study, templates, (sampling-)event and assay

Create new study

1. Go to tab 'Create' and select 'Create a new study'.
2. 'Page 1: Start'. In the first template dropdown, select (add/modify) option to create template → describe design of study.
 - a. Study template: contains standard items such as 'title', 'description', 'code' and 'startdate'.
 - b. Add other variables to describe study, e.g. study aim, study protocol, study conclusion, participating institute, consortium, etc...
3. If publications are available of this study, can search and insert PubMed reference.
4. Add contact of the study. Can assign person for data entry or reader access.
5. Quick save (instead of manual uploading and entering of the data, will next use the importer wizard to upload excel file describing subjects).

Importing subjects using an excel file

1. Go to tab 'Import' and select 'A part of the study design'.
2. 'Page 1: Import file'. Select Excel file with subject information.
 - a. Select date format: dd/MM/yyyy
 - b. Select sheet containing data: ...
 - c. Indicate columns header: ...
 - d. Select data type being uploaded: Subject
 - e. Select study: ...
 - f. Select type of data template: should be a subject template. If not created beforehand, select add/modify and create subject template:
 - i. Subject template: only two fields are obligatory, 'subject identifier' and 'species'.
 - ii. Add other variables to describe subjects, e.g. gender, age, tumour pathology, etc... (clinical variable that user may be interested in; though take care not to completely copy data of e.g. OpenClinica).
 - g. Press next.
3. 'Page 2: Assign properties'. Match fields in subject template to column headers present in Excel file.
 - a. E.g. template field 'name (IDENTIFIER)' should be matched to Excel column with header 'Subject ID', 'species' should be matched to excel column with header 'species'.
4. 'Page 3: Check imported data'. Overview of the to-be-imported subjects. Check variables and where necessary make corrections/additions.
5. 'Page 4: Confirmation'. Are you sure everything has been matched all right? If so, press next, otherwise, press previous and make changes.
6. 'Page 5: Done'. Subjects have been imported to the study. (Can start uploading other study parts now or view uploaded data)

Define (sampling-)events

In this part you need to state what happened to the subject, i.e. event (e.g. treatment with a drug, diet challenge) or sampling-event (e.g. mRNA expression biomarker was measured on material from a subject).

Go back to study for which you've imported subjects.

1. Browse/edit your study and go to tab 3: 'Events timeline'. Press edit.
 - a. Both for events and sampling-events, templates are required with descriptive fields.
2. Adding a sampling-event: select relevant sampling-event template (e.g. mRNA expression). If not created beforehand, select add/modify and create sampling-event template:
 - a. Standard template contains: 'startTime', 'duration' and 'sampleTemplate'*.

*sampleTemplate: contains standard fields 'name' and 'material'

possible to add more fields to the sample templates (e.g. organ/tissue on which measurements for that subject was performed.) Press add.

7. Below a description will be given of a group. Change name of the Group to be more informative (e.g. name of measured marker). Check the box and press next.
8. Overview of subjects and sampling-events. Mark for which subjects the mRNA expression was determined. Press next.

Samples

Continuing on directly from the previous step above.

9. Overview screen will appear with the to-be-generated sample-names (long names which are a combination of 'subject ID', 'sampling-event template' and 'Group'**). Press next.

**which in this case was the name of the measured marker

Assays

10. Need to create an assay (a place in which the measurements for this particular experiment are stored).
 - a. Select template 'assay'.
 - b. Name the assay (e.g. name of the measured biomarker).
 - c. Select the module in which to store this assay (different modules are available, in this case, qPCR could be selected).
 - d. Press add. In the field below an assay will appear. Press next.
 - e. Overview screen showing sample type and sample names. Check the relevant boxes to indicate which samples belong to what assay. Press next.

You have now created a study with subjects, sampling-events, samples and an 'empty' assay awaiting data.

II: Module section: creating platforms and features

Go to the Module tab and select the module section in which you want your assay to appear: qPCR.

Create platform

Note: a platform may be created to distinguish between types of techniques, e.g. qRT-PCR SYBR Green

1. Select Browse.
2. Select Platform.

3. Select Create new platform (e.g. for real-time experiments, a platform could be qRT-PCR SYBR Green).
4. Use the shown fields to enter description of platform. Standard fields shown are: 'Name', 'Platformversion', 'Platformtype', 'Comments' and 'Template'*.
 - a. *Platform template: if not created beforehand, select add/modify and create.
 - i. Extra fields can be added to the template for completing description of the platform.
5. Press create.

Create feature(s)

Note: a feature describes the measured item (e.g. an mRNA marker), therefore it is important to add as many fields as required (minimal information necessary to repeat the experiment).

Method 1: manual creation of a feature

1. Select Browse.
2. Select Features.
3. Feature template: if not created beforehand, select add/modify and create template.
 - a. Standard fields are 'name' and 'unit'.
 - b. Important to add additional fields describing the specifics of the measured item. E.g. for a real-time quantitative experiment, fields in the feature template could/should include:
 - i. Target gene full name(s): full name(s) of the target gene, e.g. NCBI definition.
 - ii. Target gene symbol (HGNC): name of the target gene using HGNC term (for human subjects)
 - iii. Target gene ontology: NCBI code, GI code
 - iv. Target gene forward primer: forward primer sequence; orientation given as standard: 5'-3' (in nucleotides)
 - v. Target gene reverse primer: reverse primer sequence; orientation given as standard: 5'-3' (in nucleotides)
 - vi. Annealing temperature of target gene primers (in °Celsius)
 - vii. Target gene primer manufacturer
 - viii. Reference gene full name(s): full name(s) of the reference (housekeeping or control) gene, e.g. NCBI definition.
 - ix. Reference gene symbol (HGNC)
 - x. Reference gene ontology
 - xi. Reference gene forward primer: forward primer sequence; orientation given as standard: 5'-3' (in nucleotides)
 - xii. Reference gene reverse primer: reverse primer sequence; orientation given as standard: 5'-3' (in nucleotides)
 - xiii. Annealing temperature of reference gene primers (in °Celsius)

- xiv. Reference gene primer manufacturer
- xv. Reporter dye (+ manufacturer)
- xvi. Description: description on how the qRT-PCR was performed. Can be reference to a published article or a reference to the protocol. It is also possible to enter a field with experimental protocol and upload the protocol file itself.
- xvii. Machine used for qRT-PCR (+ manufacturer)
- xviii. Software used for analysis of results (software, manufacturer and version)
- xix. Analysis/calculation + unit value of the measured feature

Method 2: importing of features

When creating multiple features of the same type of experiment, importing of features will provide a quick way of uploading them to the module section. Of note should be that the list of descriptive fields shown above cannot be imported in this manner, but should be created beforehand. However, the *content* of these fields can be uploaded in one session for multiple markers.

1. Select Import → Features.
2. Choose Excel file containing information on the features.
3. Choose data template (= Feature template).
 - a. Though stated that data template is not required, leaving this empty results in standard feature template to be used and then the descriptive fields cannot be matched in the subsequent screen. Avoid leaving the data template empty.
4. Select platform. Press next.
5. Match the headers in the Excel file to the descriptive features created for the feature template.
6. Press next. You can now use these features to add measurements to your study.

III: Adding numerical results/measurements to an assay

Note: at the moment the Phenotype Database does not yet handle results in text format, if possible, dichotomise such output to e.g. unmethylated, $U = 0$, methylated, $M = 1$.

Go back to the **Generic study capture framework section** and view your study.

1. Go to tab 'Assays'. You will now see an overview with assayname, module, samples and in the column Link a hyperlink with details.
2. Click on details.
3. You'll see a message here. Click on 'Use the **importer** to import your data' to upload results into the assay.
4. Select the correct assay. Press next.
5. Select the correct platform.
6. Select the Excel file containing the results/measurements. Press next.
7. Select sample layout. Press next.

8. Check that samples were matched correctly and that features are matched as well, otherwise, make adjustments. Press next if all is matched correctly. Your data has been added to the assay and can be viewed and exported.

Export

It may be desired to export data in order to perform calculations or visualisations on the data, or ready the data for export to e.g. the data integration platform tranSMART.

1. To export the study (meta)data, go to the tab 'Export' → 'Export studies to SimpleTox file'. Select the study of interest and press Export. The downloaded file will contain information on subjects, sampling-events and sample-names.
2. To export the study (meta)data as well as the measurements, go to the tab 'Export' → 'Export assay data to file'. In the next screen, mark the items of interest and submit. Press ok in the next screen. The downloaded file will contain information on subjects, sampling-events, sample-names, event groups and the measurements.