

cBiT manual

1. What is cBiT?

The Compendium for Biomaterial Transcriptomics, or cBiT, is a repository that incorporates data archives containing material science and transcriptomics-based cell biology (i.e. “materiomics”). One of the main goals of cBiT is to describe all meta-information necessary for reproducibility and downstream analysis of experiments with respect to the culturing conditions, transcriptomics procedure, and biomaterial fabrication and properties.

This is a step towards:

- A better comparability between biomaterial-based experiments, allowing for new and more efficient data analysis strategies
- A sustainable infrastructure (well annotated experiments can also be uploaded to public repositories like GEO or ArrayExpress)
- Maximum and effective use of public data

Data archives in cBiT are prepared using ISAcreator (<http://isa-tools.org/>) which is essentially dedicated spreadsheet software that enables the inclusion of transcriptomics data (or other Big Data types) and allows for an unambiguous characterization of meta-information.

cBiT was designed to be intuitive in use and with this manual users can get all the information they need to start looking for data, downloading interesting data sets, and analyzing them!

2. Main lay-out

cBiT is available at <https://cbit.maastrichtuniversity.nl/> and should work fine in recent versions of Chrome, Firefox, and Safari. cBiT is not compatible with Internet Explorer and for those users we suggest switching to Chrome. cBiT consists of three sections divided into tabs, the Welcome tab, the About tab, and the Browse tab. The Browse tab is the most important for cBiT users but we will first briefly go into the Welcome and About tabs.

2.1 Welcome tab

The Welcome tab is cBiT's homepage and contains a short explanatory video of what cBiT is all about. Please have a look and don't hesitate to distribute the video amongst colleagues. The more people know about cBiT, the better!

2.2 About tab

The About tab describes the idea behind cBiT in a bit more detail, offers the cBiT manual for download, provides instructions on how you can contribute your data, and shows our contact details.

2.3 FAQ tab

This tab contains a list of frequently asked questions (FAQ) that you may find useful.

2.4 Browse tab

The Browse tab is what cBiT is all about. Here, you can browse through the available data sets, look for specific data using either the search box or the quick selection menu, and download data of interest for further analysis. In the next paragraph the Browse tab will be explained in more detail and examples will be given on how to use cBiT.

3. Using cBiT

The Browse tab is divided into a results section on the right side of the screen showing the cBiT studies, a search and selection menu section on the left to perform queries, and a download box in the upper right part to download selected studies and samples (**Figure 1**). Before showing how to search the cBiT repository, let's first look into the “study cards” in the results section.

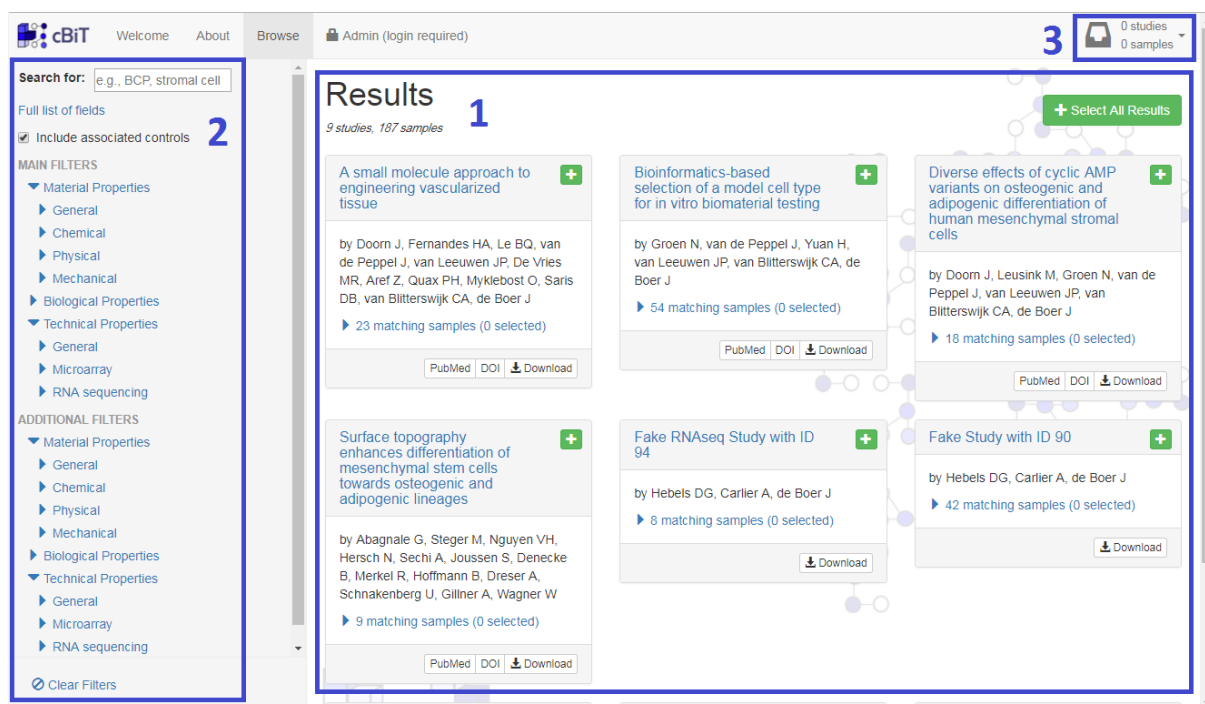


Figure 1: The three sections of the Browse tab, i.e. the results section (1), a search and selection menu section (2), and a download box (3).

3.1 The study cards

The results section shows all studies as “study cards” with the study title, the authors/researchers involved and a collapsible overview of the matching samples in each study (**Figure 2**). Pubmed and DOI links are also available for studies that have been published already. By default, all studies present in cBiT are shown. Clicking on a study title will show a pop-up with additional details about the study, including a study abstract, experimental protocols and sample information. It is a good way to get a grasp of what the study was about and whether it is of interest to you.

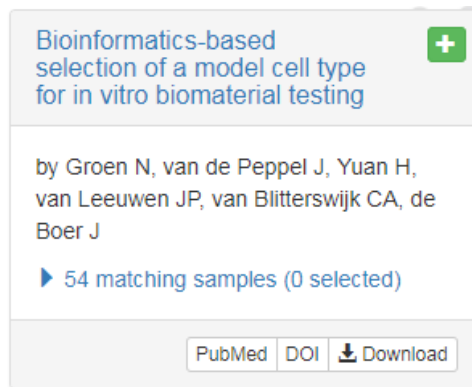


Figure 2: Example of a study card.

3.2 Performing a query

In order to query the cBiT database, the underlying search engine will find matches in any of the recorded study properties of the studies and samples in cBiT. For a list of study properties available in cBiT (Material, Biological, Technical) click on “Full list of properties” just underneath the search box. A specific cBiT query can be performed using either the search box or the selection menu (**Figure 1**, section 2) in the following way:

1. Use the search field in the upper left corner of the screen to look for specific study terms. Most study properties are indexed when new data sets are integrated into cBiT and all these properties are being searched for query matches. For example, searching for “MG-63” will return all studies using the MG-63 cell line (included in the “Cell strain” biological property). The results section on the right immediately updates in real-time the studies and samples matching the query.
2. Use the quick selection menu on the left of the screen to narrow down studies. The menu items correspond with study properties and selecting/de-selecting items will immediately update the studies and samples in the results section that match the selection. Properties are divided into Material, Biological, and Technical Properties and further sub-categories per item. Properties which are popular are included in the “MAIN FILTERS” while less popular properties can be found under “ADDITIONAL FILTERS”.

Most properties have values that can be checked or unchecked (e.g. “Material class”) but there are also properties containing numerical values where a range of values can be selected using a slider (e.g. “Culture duration”) (**Figure 3**). For properties with the slider option the unit of measurement can also be selected (e.g. hour, day, week, etc.). The number of samples matching each property value is shown on the right.

An overview of all available biomaterial properties included in cBiT can be viewed by clicking “Full list of fields” just underneath the search box.

The search box and selection menu can also be used in parallel, e.g. first search for a specific term and then narrow down further by using the selection menu. To start a query from scratch, click on “Clear Filters” on the bottom left.

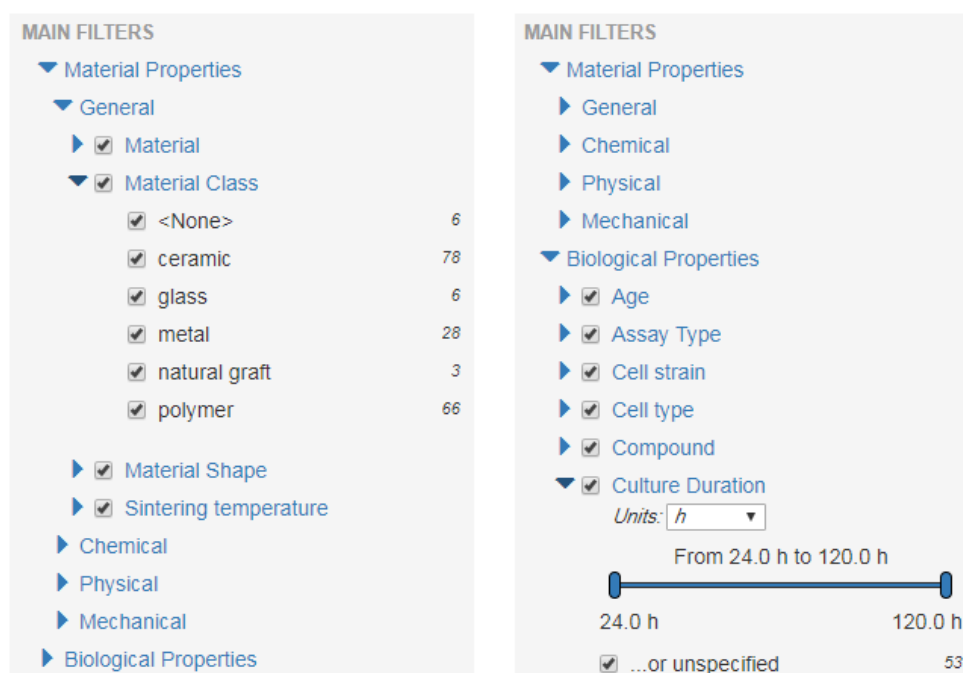


Figure 3: An example of a check box study property (left) and a slider property (right).



3.3 Selecting studies and samples for download

Based on a query (and after exploring the study cards), you may decide to select certain studies or specific samples for download. In most cases, downloading the study will be sufficient, but you may also want to explore the samples in the study first and leave out certain samples which are not of interest to you after all.

3.3.1 Selecting a study for download


Selecting a study for download can be done by clicking on the green “+” button in the study card (**Figure 2**). The button will turn into a red “-” once a study is added to the download box. Clicking it again will remove the study from the download box. Keep in mind that clicking this button will include all samples within that study that match the search query (click on “xx matching samples” for a list). Samples that are part of a study but that do not match the search query will therefore not be


included. However, these can still be selected (see next paragraph). Of course it is also possible to select an entire study for download without first having performed a search query. There is even an option to select the entire database for download by clicking the green “Select All Results” button. To keep track of what has been selected for download, the number of selected studies and samples is shown in the download box in the upper right corner.


A small molecule approach to engineering vascularized tissue  


by Doorn J, Fernandes HA, Le BQ, van de Peppel J, van Leeuwen JP, De Vries MR, Aref Z, Quax PH, Myklebost O, Saris DB, van Blitterswijk CA, de Boer J


▼ 23 matching samples (20 selected)

 **5786049101_A** Control: TRUE
, Donor identifier: D142 , Dose: 0 mM, Sample description: Donor D142_control_rep1

 **5786049101_B** Control: TRUE
, Donor identifier: D142 , Dose: 0 mM, Sample description: Donor D142_control_rep2

 **5786049101_C** Control: TRUE
, Donor identifier: D142 , Dose: 0 mM, Sample description: Donor D142_control_rep3

 **5786049101_D** Compound: DFO , Control: FALSE , Donor identifier: D142 , Dose: 0.150 mM, Sample description: Donor D142_deferoxamine_rep1

 **5786049101_E** Compound: DFO , Control: FALSE , Donor identifier: D142 , Dose: 0.150 mM, Sample description: Donor D142_deferoxamine_rep2


 **5786049101_F** Compound: DFO , Control: FALSE , Donor identifier: D142 , Dose: 0.150 mM, Sample description: Donor D142_deferoxamine_rep3

Figure 4: Example of a study card where some samples have been deselected (green buttons) while the rest of the samples is still selected (red buttons).

3.3.2 Selecting specific samples for download

Before adding a study to the download box, you may first want to explore the samples inside. Click on “xx matching samples” in the study card to have a look. As you can see, every sample has the same selection button as the entire study does. With this button you can add or remove single samples from the download box (**Figure 4**). You may notice that even though you have searched for a specific term (e.g. the compound “dexamethasone”), there also some samples in the “xx matching samples” list that do not match (e.g. no exposure to dexamethasone). These are the control samples and they are included by default since they are usually the group you want to compare your samples of interest to. This option can be switched off by unchecking the “Include associated controls” check box just above “MAIN FILTERS” on the left of the screen.

3.3.3 Downloading full studies

Studies can also be downloaded directly without performing any search queries or property selections. This is basically the default situation when opening the Browse tab. By clicking on the green “+” button in a study card (**Figure 2**), the full study including all samples will be added to the download box.

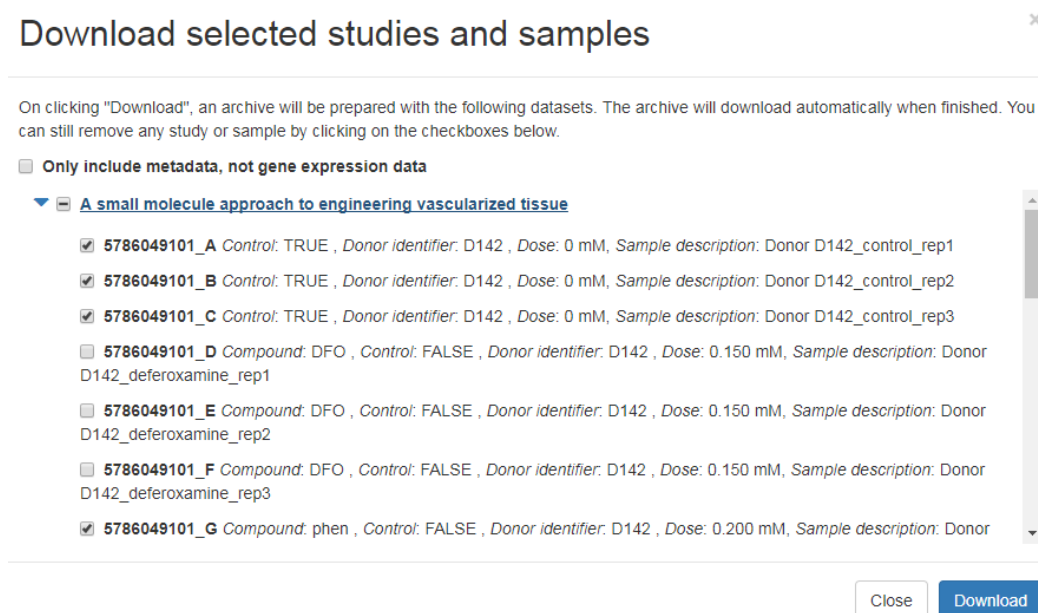


Figure 5: The Download window where studies/samples can still be selected/deselected before downloading the zip archive.

3.4 Downloading samples

After placing studies and samples of interest into the download box, the selection can be downloaded using the download box icon in the top right corner. Clicking this will show two options: “Clear selection” and “Download”. The first option will clear out the download box. Only click this when you want to remove all studies. The second option gives you a small pop-up with an overview of the studies and samples selected for download (**Figure 5**). You can still remove any study or sample here by clicking on the checkboxes. If everything is fine, click on the Download button and a zip archive will be prepared with the selected datasets. The archive will download automatically when finished.

3.5 Exploring the zip archive

Once downloaded, extract the files in the zip archive. This will create a folder containing the raw and normalized transcriptomics data, the study protocols, the study properties, including all measured material properties, and any additional supplementary files (**Table 1**). These data can now be used for data analysis purposes, such as differential expression analysis, correlation analysis, pathway analysis, gene network construction, etc.

Table 1: List of file types typically present in an extracted cBiT zip archive.

File type	Description
field_descriptions.tsv	Alphabetical list of all study properties contained in the download with an explanation of each
i_Investigation.txt	The study protocols and some additional information on associated publications
sample_info.csv	Values of all measured or documented study properties for the selected samples in the study, i.e. the metadata
Biomaterial graphs file*	A separate metadata Excel file, containing more complex study property data, such as graphs or images
Supplementary file*	One or more supplementary files containing additional measured or relevant data, not included in the sample_info.csv file or Biomaterial graphs file.
Raw data file(s)**	The raw data file(s) of the transcriptomics assay corresponding with the selected samples
Processed data file	The processed data file(s) of the transcriptomics assay corresponding with the selected samples and having undergone processing steps such as normalization
Annotation file	Additional annotations for the genes in the raw and processed data files (e.g. EntrezGene ID, Gene Symbol, etc.)

* Not always present; ** Not present for RNAseq data

3.6 Invitation to participate

cBiT's usefulness as a dedicated database increases as more data comes in, not only in terms of the number of datasets, but also their level of completeness and variation. We therefore invite research groups who are interested in depositing their own biomaterial-associated data in a sustainable repository and who are eager to contribute to the cBiT project, to contact us. It is possible to either deposit your own data, or provide us with biomaterials that have clinical potential which we will proceed to investigate, analyze, and incorporate into cBiT. This can also include data analysis of data already present in cBiT. In short, the possibilities offered by cBiT are:

1. Browse, download and analyze data yourself
2. Browse, download and find an external party to analyze the data
3. Contact us with an interesting question and ask us to analyze cBiT data
4. Supply us with an interesting biomaterial and ask us to generate data
5. Upload your own dataset

If you would like to participate, please send an email to Dennie Hebels at:

d.hebels@maastrichtuniversity.nl

4. Questions and feedback

If you have any questions on how to use cBiT which are not covered in this tutorial or if you run into any problems during your queries, please contact Dennie Hebels (d.hebels@maastrichtuniversity.nl). Recommendations for improvements to the cBiT infrastructure, data archives or the tutorial are also much appreciated!