RESTful BioCurations

Quick Start Guide

2013

Copyright

Document history
RESTful BioCurations/Quick Start Guide (2013)

Warranty
The information contained in this document is subject to change without notice.
ProContent Ltd makes no warranty of any kind with regard to this material, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose.
ProContent Ltd shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this material.
ProContent Ltd assumes no responsibility for the use or reliability of its software on equipment that is not furnished by ProContent Ltd.

How to Obtain Support

ProContent Limited
Hardtwaldallee 18
61381 Friedrichsdorf
Germany
Phone: + 49-179-463-1067
Fax: + 49-80-801-5422

Email: support@pcontent.de
Internet: www.pcontent.de

Contents

1 Introduction

1.1 Preface

RESTful BioCurations (ReBiC) is developed to stimulate the development of bioinformatic Web services using Google AppEngine. ReBiC is a powerful tool to create Web resources to integrate the diverse information accumulated in the molecular biomedicine. With the platform environment, you can create your own databases, post them in the Internet, connect to various information sources and manage the content.

ReBiC provides a programming method you need to easily integrate the targeted vision of the scientific achievements. You can monitor multi-lab research efficacy by mapping updates on a consolidated color-coded matrix.

1.2 Features

The ReBiC platform enables development of knowledge-based resources to answer the following questions:

- what is the current scope of the problem ahead?
- how much is known about the research items under investigation?
- if there are valuable contributions from your research into the field?
- what steps should be taken to have an impact?

ReBiC provides the following features for a guest user:

- browsing the current annotated datasets;
- switching among the published datasets;
- selecting tracks (features) to annotate data in the current set;
- selecting data in a particular set.

A registered ReBiC user is provided with the option to save the datasets created with the assigned tracks and to publish these datasets in the Web.

As a registered user, you are able to request the administration for curator privileges (see the Curators section for more details). You will be able to establish a track(s) and get the following features:

- managing track settings;
- updating track values;
- adjusting track color coding.

1.3 Workflow

A typical workflow consists of seven steps:

1. Selecting a dataset
2. Viewing the dataset
3. Changing the Display settings
4. Sorting
5. Selecting a biomolecule(s) for a new dataset
6. Sharing the dataset
7. Exporting the data

2 First Steps

2.1 Guest User

Use Guest user on the ReBiC home page to start working with the system (Figure 1). After selecting Guest user, the default dataset is loaded into the current workspace. At the top right of the Web page there are the Guest Login indicated and the Logout hyperlink provided.

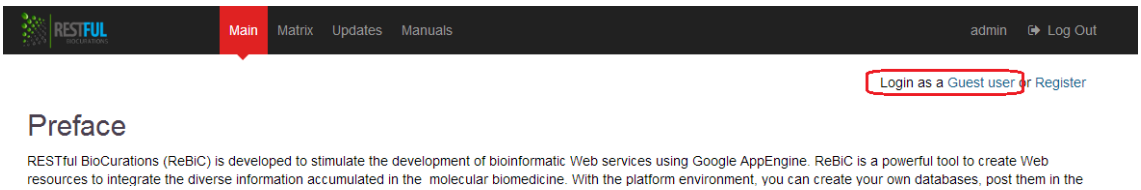

Figure 1 – Guest User

2.2 Main Tabs and Menu

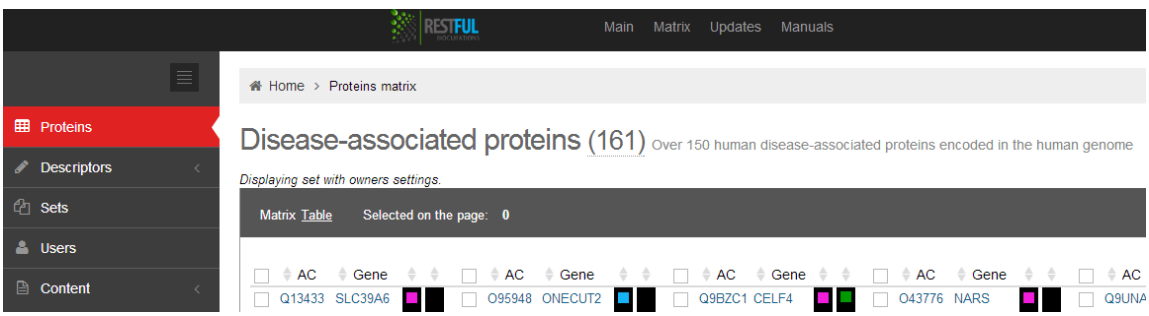

Figure 2 – Main tabs

There are three tabs available to switch among the ReBiC workspace panels (Figure 2). The Proteins tab contains a matrix or corresponding table. The Descriptors tab provides an access to the list of available tracks and enables selection of appropriate tracks for the current dataset. The Sets tab enables selection of a dataset for browsing.

Navigation through accompanying information, e.g. manuals, news, etc., is provided by the menu at the top right corner (Figure 3). To return to the workspace, use the Matrix hyperlink.

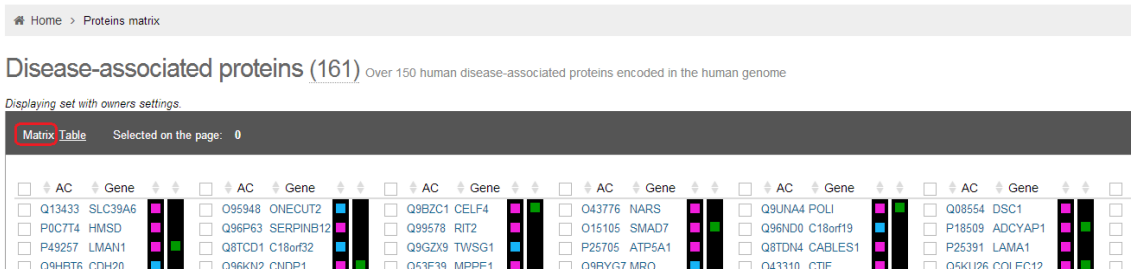

Figure 3– Menu

2.3 Registration

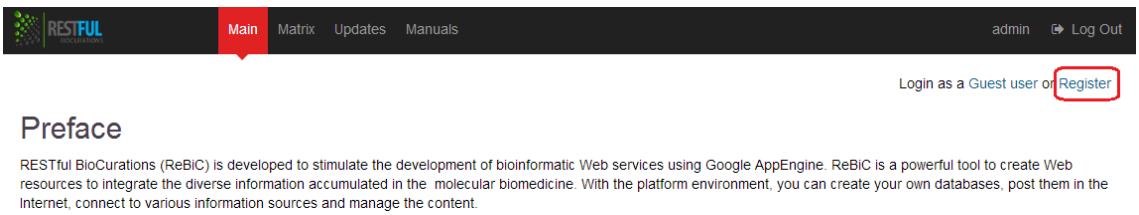

Figure 4 – Registration links

Use the Register or Sign up hyperlink on the ReBiC home page to obtain the option to save the sets created with the assigned tracks and to publish these sets in the Web (Figure 4). Fill your e-mail, user name and password in the registration form and click on the Sign up button. The system will check your password and highlight the result (weak, medium, strong or very strong). We recommend using a strong password at the least.

2.4 Login and Logout

After registration, you can authorize in the system using the Login hyperlink. A successful authentication results in redirect to the workspace. At the top right of the Web page there are the User Name indicated and the Logout hyper-

link  provided. User name is a link to your profile where you can change the registration information, including your password.

3 Sets

3.1 Selecting Sets

Select the Sets tab on the current workspace page. The dataset list will be provided (Figure 5). Each set has a short title and description. The next column shows the set type: reference sets are provided by the administrator while users' sets can be shared with the users of the one ReBiC server. The owner (creator) of the set is also indicated on the tab followed by the number of entries (biomolecules, etc.) in the set. Select a particular set and click on its title to load it into the workspace.

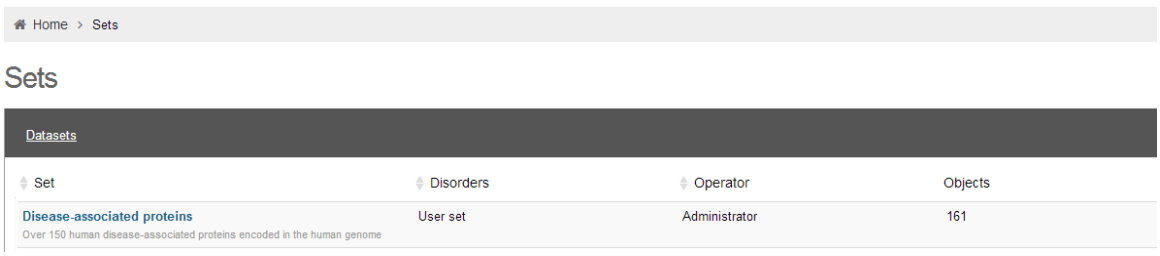

Figure 5 - Selecting datasets on the Sets tab View Experiment

The selector on the Sets tab enables switching between the whole list of all available sets (All sets) and the datasets created by the user (My sets). Use the + Create set button to create a new set: you will be prompted to enter a name and description of the dataset you are creating.

You will be able to save the set only if you have registered and authorized in the system, otherwise the changes will be lost after the ReBiC Web page is closed in the browser.

Special controls for the dataset list (see the Changing Display Settings section for more details) are:

the  icon to get a URL for broadcasting a set. The set will be accessible to any recipient of the URL with the tracks you have selected. The recipient of the broadcasting URL need not to authorize to view it. The user will be automatically recognized as a guest. The user will have to register and authorize only if it is necessary to save the changes in the set configuration.

the  icon to view a set with the tracks and coloring recommended by the set owner.

3.2 Browsing Datasets

3.2.1 Matrix View

The data/biomolecule entries are displayed in several columns. The names of the biomolecules listed can be provided in different notations depending on the user's preferences: Biomolecule Symbol, Biomolecule Name, Accession Code (see the Biomolecule Naming Style section for more details).

The matrix mode supports three tooltip types depending on if the pointer is over a row name (biomolecule name), column name (track label) or colored box (see below). The black background on the matrix view denotes that the information is not available. You can highlight a biomolecule by clicking on the checkbox near the biomolecule name. The biomolecule highlighting is kept while switching between the views. Sorting is available on the matrix view; use the column names for sorting (see below).

Clicking on a biomolecule name results in the opening of a particular biomolecule page where all available features of the selected biomolecule are displayed (Figure 6).

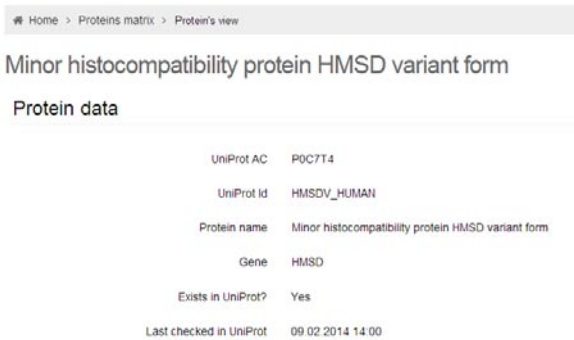

Figure 6 - Particular biomolecule page

3.2.2 Table View

The matrix view can be switched to a tabular format by clicking on the Table hyperlink on the upper toolbar (Figure 7). The table view provides numerical data, color coding is available on the matrix view. The tooltips and sorting are also available (see below). You can highlight a biomolecule by clicking on the checkbox near the biomolecule name. The biomolecule highlighting is kept while switching between the views.

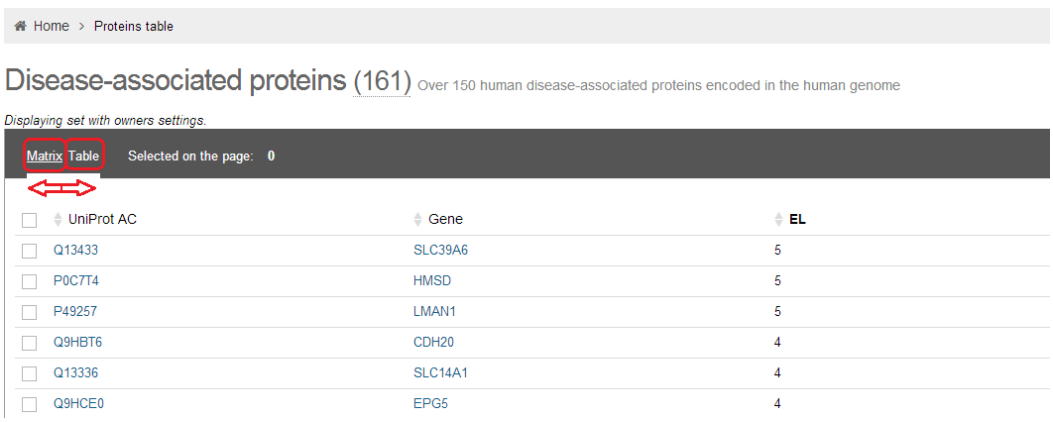

Figure 7 - Switching between the matrix and table view

3.2.3 Tooltips

The matrix/table mode supports three tooltip types depending on if the pointer is over a row name (biomolecule name), column name (track label) or colored box/value (Figure 8). A row name tooltip provides hyperlinks to the biomolecule information in external resources (e.g. in NextProt).

A column name tooltip provides a hyperlink to a particular track page and color coding legend. A colored box tooltip stores a value, some comments and link to an external resource where it is explained how the value was obtained. Both the column name and colored box tooltips include date of the data last update. You can see it clicking on [+]. Some additional information, such as the track description, is also available by clicking on [+] in the track label tooltip.

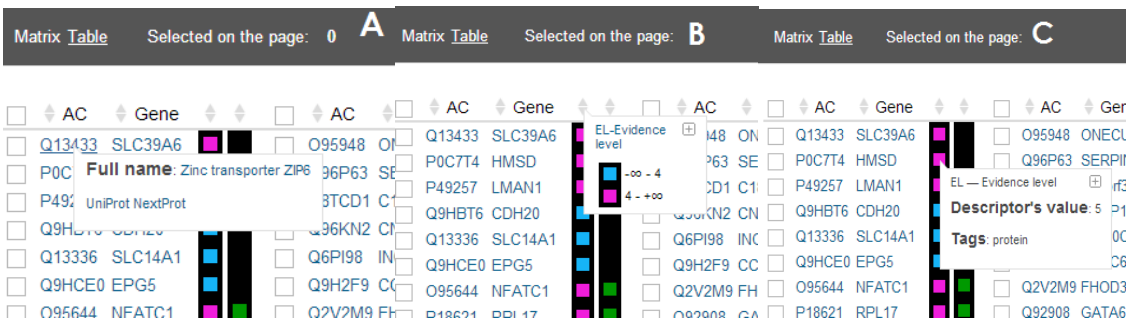

Figure 8 - Browsing the matrix mode tooltips.
A - row name (biomolecule name), B - column name (track label), C - colored box

3.2.4 Sorting

You can sort the matrix by clicking on the column label. The information about the track used to sort the matrix is displayed below the matrix (Figure 9). The name of the selected track, as well as the distribution of the values, is shown according to the color scheme. When sorting the track for first time, there are the biggest values at the top, for the second time there are the lowest values at the top, at the matrix bottom there is always a black background.

If switching to the table view, the sorting will be lost. You can sort the table by clicking on the column name. When sorting the track for the first time, the val-

ues are sorted in descending order, for the second time – in ascending order. The null values on the table view are equal to the highest ones so they will be at the top if descending sorting.

Figure 9 - Statistics for the current track (used for data sorting)

3.3 Selecting Data for a New Set

Select data by clicking on the checkboxes at the left of the biomolecule name in the matrix or table browsing modes. The selected biomolecules will be highlighted. All biomolecules in the matrix column can be selected by clicking on the checkbox in the matrix header. The number of the selected biomolecules is prompted at the top matrix panel (Figure 10).

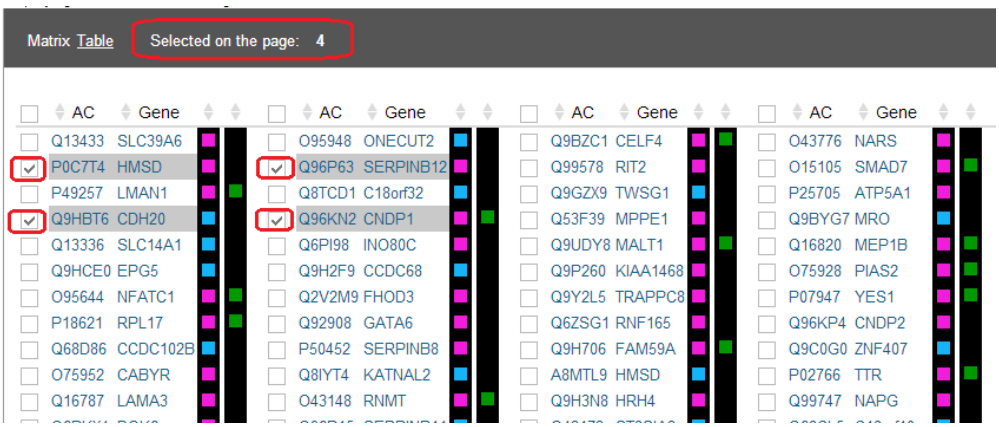

Figure 10 - Selected biomolecules are highlighted on the matrix view

After selecting the biomolecules, you can either create a new set with them or add them to the existing one using the ← Add to set button on the Proteins tab. The dialog box will appear as shown in the figure below (Figure 11). A reg-

istered user can also select one of the pre-existing sets to add new biomolecules there. A guest user should check the Add new set option to proceed.

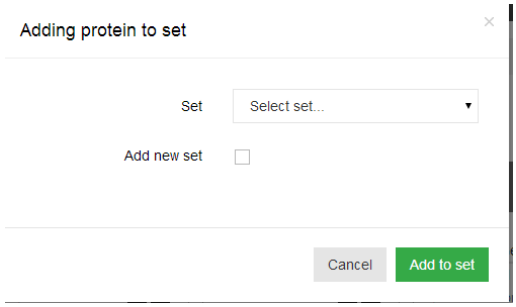

Figure 11 - Dialog Box: Adding biomolecules to a set

3.4 Sharing Sets

To share a set, select the Sets tab and enter editing mode by clicking on the icon in the row of your set. Toggle the Shared checkbox and save the changes. As a result, the set will be available to the users of the one ReBiC server.

Each dataset has the creator's settings (the selected tracks, track outputting order, separator placement and selected name style for biomolecules) but these settings can be changed by any user (see the Changing Display Settings section for more details). If a user has modified the set settings, the initial settings selected by the creator can be viewed after clicking on the icon. Otherwise, if the set should be broadcasted with the user's settings, the appropriate hyperlink is generated after clicking on the icon (a page with an arrow).

4 Tracks

4.1 Selecting Tracks

With some current dataset uploaded, switch to the Descriptors tab. The list of tracks will be displayed (Figure 12). Each track has a short name and type. The Biocuration type indicates manually updated tracks, while Web service type corresponds to the automatically updated tracks. The number of biomolecules with available values is provided for each track.

The currently visible tracks are marked with the green check sign. Click on it to unselect the tracks and exclude them from the matrix. Select other tracks by clicking on the grey check sign to include them into the matrix browsing. The number of the currently selected tracks is indicated above the table. After selecting the track, switch back to the Proteins tab to apply the changes.

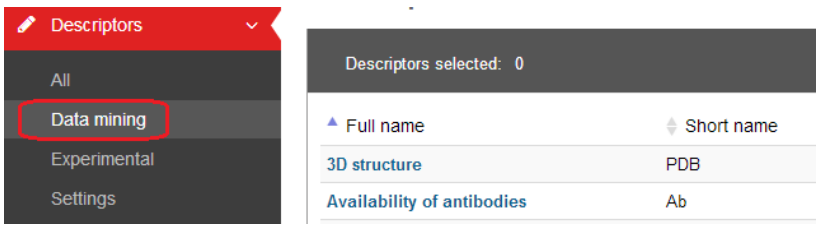

Figure 12 - List of tracks

It is useful to sort the selected tracks by clicking on the last column header (underlined check sign). The checked tracks will be gathered at the top of the list and can be observed on a screen at a time. The Reset settings hyperlink unselects all tracks by a click. That is useful to arrange tracks according to your preferences.

The tracks are organized into several categories displayed on the Descriptors tab. The categories make easier to navigate particular groups of tracks, e.g. the tracks acquired from international resources (Data mining), pre-published or published data (Experimental) or other.

4.2 Creating Tracks (Curators)

To create a track, you should send a request to the administrator (support@pcontent.de). The e-mail must include the user name, e-mail address specified during registration, necessary number of tracks, their names and types (Web service or Biocuration).

The administrator will make you a curator and provide with the curator manual and hyperlinks to manage your tracks.

5 Settings

5.1 Changing Display Settings

There are three types of settings to change:

- visibility of the tracks (see the Selecting Tracks section);
- order of the tracks and separators between track groups (see Order and Separators section);
- visibility and order of the biomolecule names (see Biomolecule Naming Style section).

ReBiC stores the sets of display settings for each user. The owner's dataset display settings are copied to the user when the first visit of the set. So:

- To access to the dataset with the current user's display settings, select the Sets tab and click on its title.
- To get a URL for broadcasting a set, click on the icon in the row of the required set (the Sets tab). The set will be accessible to any recipient of the URL with the display settings you have selected.
- To view the dataset with the owner's (creator's) display settings, select the Sets tab and click on the icon in the row of the required set. The set will be loaded in the workspace.

5.2 Order and Separators

To change the order of the tracks and separators between track groups, go to Descriptors >> Descriptors' setting on the ReBiC Web page (Figure 13). The list of visible tracks will be displayed. Each track has three toggles:

- Separator (show a separator after the track column, a white line on the matrix view, a gray background color on the table view);

- Value (show values on the table view);
- Tags (show tags on the table view).

If both Value and Tags unselected, track will be invisible for the current set. The ✔ sign means the toggle is switched on, ✔ – switched off.

Figure 13 - Descriptors' settings

The order of the tracks can be changed by swapping neighboring tracks in the list. Use the ↑↓ button to swap tracks and arrange them as you desire.

The Default settings hyperlink returns to the selection of the tracks recommended by the administrator. The Reset settings hyperlink unselects all tracks by a click.

5.3 Biomolecule Naming Style

To change the biomolecule naming style, adjust Descriptors >> Settings on the ReBiC Web page (**Ошибка! Источник ссылки не найден.**).

To select any number of biomolecule names to be shown, click on ✔. To unselect, click on ✔.

The order of the biomolecule names can be changed by swapping neighboring rows in the list. Use the ↑↓ button to swap the biomolecule names and arrange them as you desire.

6 Search

6.1 Quick and Extended Search on the Matrix/Table View

The Quick Search is based on the keyword co-occurrence within the current set.

The Extended Search first sends a query to UniProt (http://www.uniprot.org/), retrieves the accession codes for the found UniProt entries and filters the matching proteins from the current set.

The biomolecule highlighting will be lost after searching.

6.2 Searching Sets or Tracks

Searching sets or tracks makes easier to navigate. To find a set or track, enter a part of its title or description. Special controls like the  icons are also available for search results.

7 Sample Scenario

7.1 User's and Creator's Settings of the Set

To understand how it works, follow the scenario:

1. Login as a guest
2. Go to the Sets tab
3. Select the Biomolecules set
4. Go to the Descriptors tab
5. Click on the Reset Settings hyperlink
6. Select some tracks using 
7. Go to the Protein tab
8. Go to the Sets tab
9. Click on the  icon for the set
10. Go to the Sets tab
11. Click on the  icon on the Biomolecules set
12. Copy the link and use it in a web browser