

QUICK START GUIDE

version 1.0

Introduction • First Steps • Sets •
Tracks • Settings • Search • Sample Scenario

*Gene-centric
Content
Management
System*

Contents

Preface.....	2
Features.....	2
Workflow	3
Guest Login.....	4
Main Tabs and Menu	4
Registration.....	5
Login and Logout	6
Selecting the Set of Genes	7
Browsing the Gene Set	8
Selecting Genes into the New Set	11
Sharing the Set	12
Choosing the Tracks.....	14
Creating the Track (Curators).....	15
Change Display Settings	16
Order and Separators.....	16
Style of Gene Naming	17
Quick and Extended Search on Matrix/Table View	19
Search the Set or the Track.....	19
User and Creator Setting of the Set.....	20

Introduction

Preface

Gene-centric Content Management System (GenoCMS) is developed for integration of postgenomic data using Web-based heat map or traffic-light matrix data representation. The system is destined to foster the collaborative filtration of the underlying data scope, by creating, annotating and sharing the sets of protein-coding genes.

Features

GenoCMS provides the following features for the Guest User:

- browsing of the annotated current set of protein-coding genes;
- switching between published sets;
- selecting the tracks (features) for annotating the protein-coding genes of the current set;
- selecting genes into a separate set.

Registered GenoCMS user provides an option to save the created sets with assigned tracks and to publish these sets on Web.

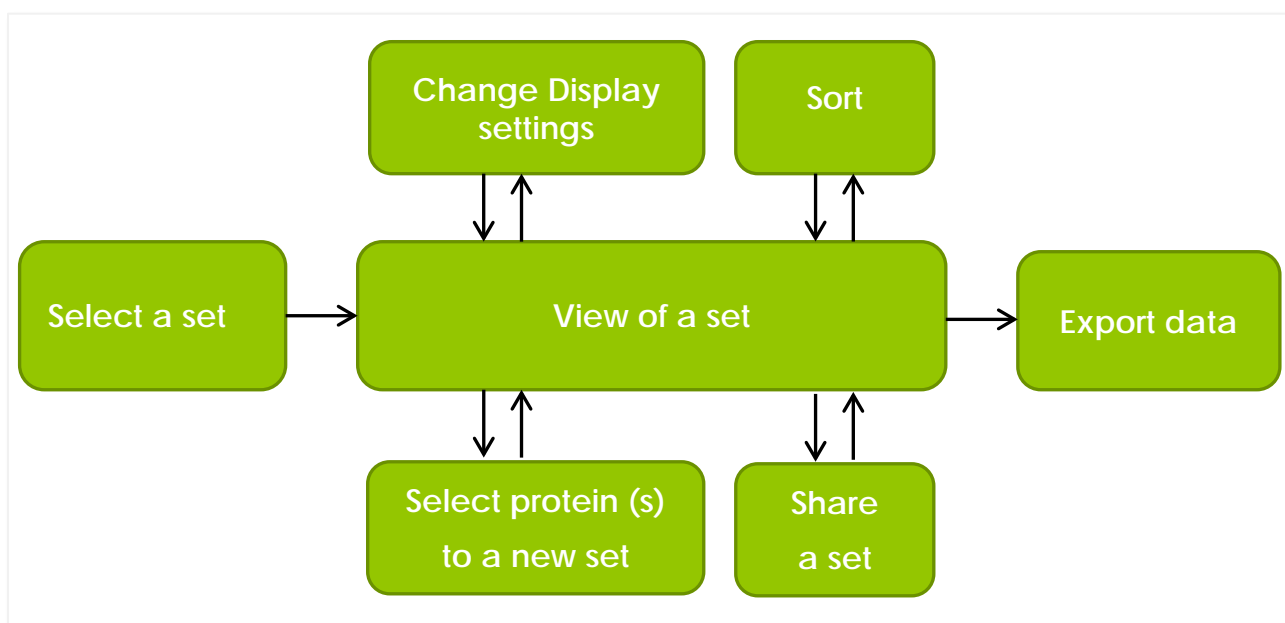
As a registered user, you are able to request to the administration of privilege Curator (in more detail below, section Curators). You will be established track (s) and you will get the following features:

- managing track settings;
- update the values of the track;
- adjusting the color-coding of the track.

Workflow


Typical workflow consists of seven steps:

1. Select a set
2. View of a set
3. Change Display settings
4. Sort
5. Select protein (s) to a new set
6. Share a set
7. Export data



First Steps

Guest Login

Use Guest Login from the front page of GenoCMS to start working with the system (Figure 1). After selecting Guest Login the default set of genes is loaded into the current Workspace. At the top right part of the Web-page the Guest Login is indicated and Logout hyperlink  is also provided.

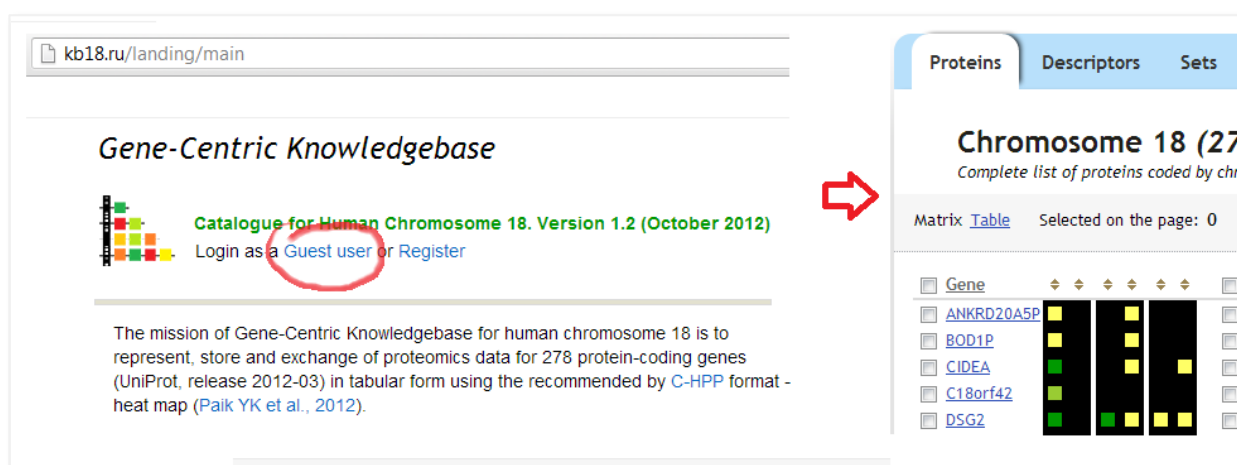


Figure 1 Guest Login.

Main Tabs and Menu

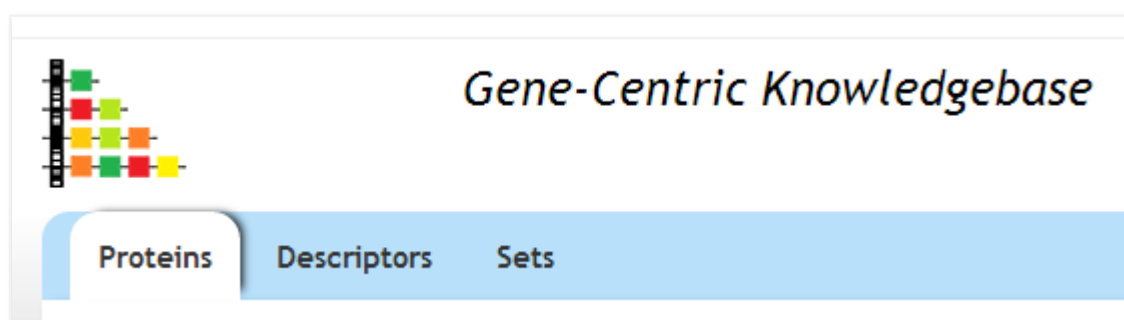


Figure 2 Main tabs.

Three tabs are available to switch the panels of the GenoCMS Workspace (Figure 2). "Proteins" tab contains the matrix or a corresponding table. "Descriptors" tab provides an access to the list of available tracks and enables selection of appropriate tracks for the current set. "Sets" tab enables to select the set of genes for browsing.

Navigation among accompanying information, e.g. manual, news, etc. are carried out by special menu in top right corner (Figure 3). To return to Workspace use the hyperlink "Matrix".

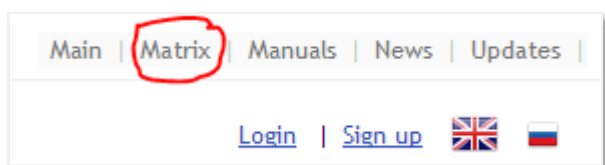


Figure 3 Menu.

Registration

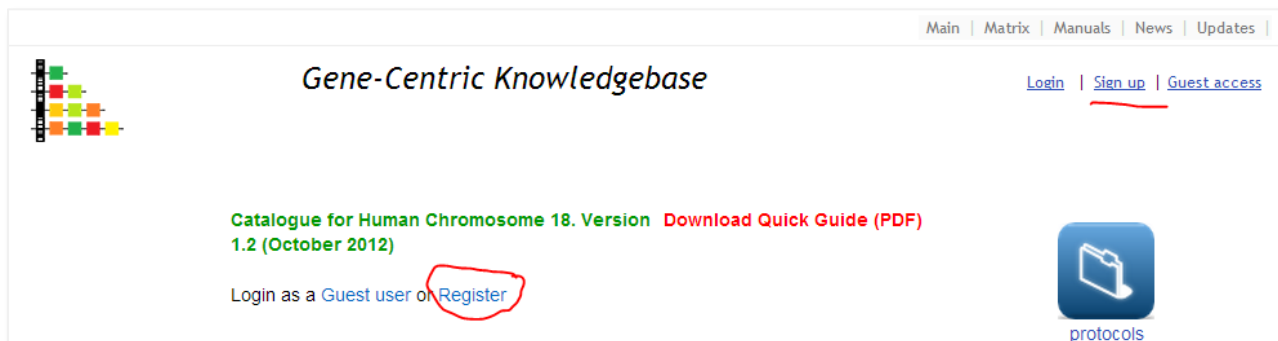



Figure 4 Links to Registration.

Use "Register" or "Sign up" hyperlink from the front page of GenoCMS to obtain an option to save the created sets with assigned tracks and to publish these sets on Web (Figure 4). In the form of registration you need to fill your e-mail, user name and password and click "Sign up". The system will be checking your password and highlight a result (weak, medium, good or strong). We recommend use a good password at the least.

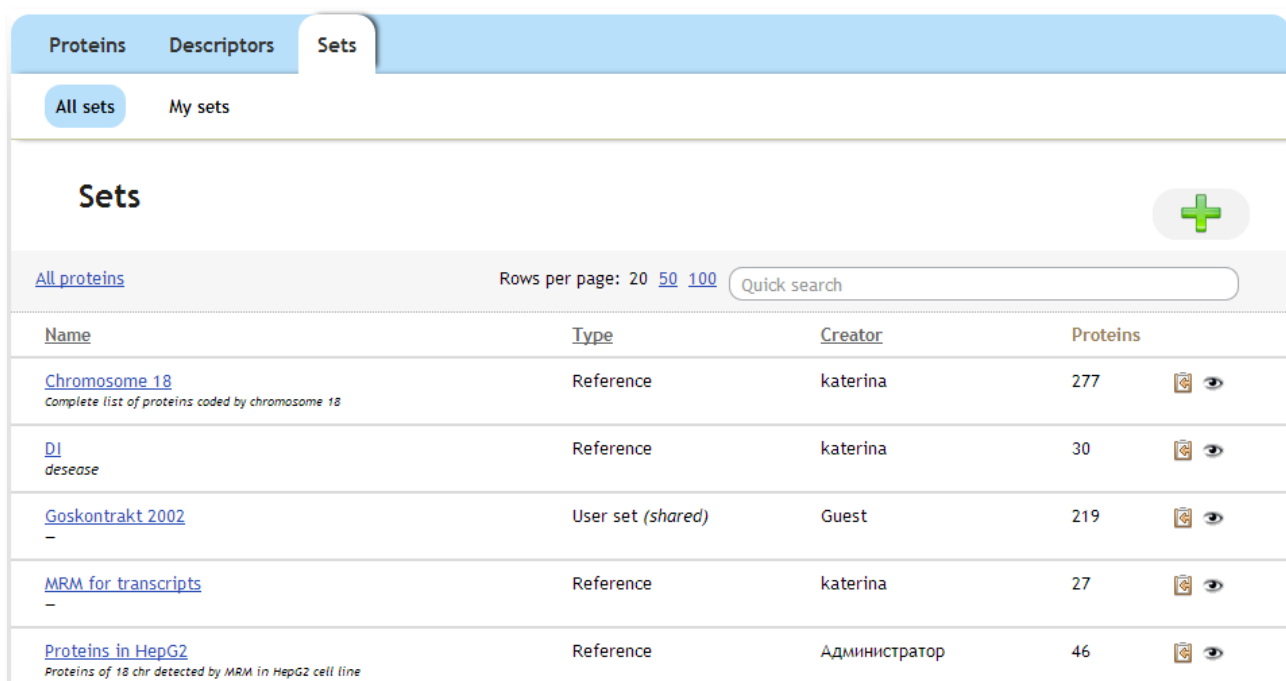
Login and Logout

After registration you can authorize in the system by using hyperlink "Login". Redirection to the workspace is a result of successful authentication. At the top right part of the Web-page the User Name is indicated and Logout hyperlink  is also provided. User name is a reference to your profile, where you can change the registration information, such as your password.

Sets


Selecting the Set of Genes

Select the “Sets” tab on the current Workspace page. The list of the gene sets will be provided (Figure 5). Each set has a short title and a description. Next column shows the set type: the reference sets are provided by Administrator, while User sets can be shared among the users of one GenoCMS server. The owner (creator) of the set is also indicated on the tab, followed by the number of entries (protein-coding genes) in the set. Select a particular set and click on its title to load it into the Workspace.




Name	Type	Creator	Proteins
Chromosome 18 <small>Complete list of proteins coded by chromosome 18</small>	Reference	katerina	277
D1 disease	Reference	katerina	30
Goskontrakt 2002 -	User set (shared)	Guest	219
MRM for transcripts -	Reference	katerina	27
Proteins in HepG2 <small>Proteins of 18 chr detected by MRM in HepG2 cell line</small>	Reference	Администратор	46


Figure 5 Selecting sets of genes/proteins from the “Sets” tab.

Selector on the Sets tab enables to switch between the whole list of available sets (“All sets”) and the sets that were created by the user (“My sets”). Use  button to create new set: you will be prompted to enter the name and description of the created set.

You will be able to save the set only if you are registered and authenticated in the system, otherwise the changes will be lost after GenoCMS Web-page is closed in the browser.

Special controls for the list of sets (more in Change Display Settings):

 to get the URL for broadcasting the set. The set will be accessible to the any recipient of the URL with the same tracks, as you have selected. Recipient of broadcasting URL should not pass the authentication procedure to view it. User will automatically be recognized as a Guest. Only if necessary, make the changes are saved in the configuration of set, the user will have to be registered and authenticated.

 to view the set with the same tracks and coloring, that has been recommended by the set owner.

Browsing the Gene Set

Matrix view.

The gene entries are displayed in several columns. The names of the listed genes can be provided in different notations depending the user preferences: Gene Symbol, Gene Name, Accession Code (more in section Style of Gene Naming).

The matrix mode supports three types of tooltips, depending if the pointer is over the row name (gene name), column name (track label) or colored box (see below). The black background in the matrix view denotes that information is not available. You can highlight a gene by clicking at the checkbox near the gene name. Highlighting of genes is kept at switching the views. Sorting is available in the matrix view; use the column names to apply it (see below).

Clicking on the gene name results in the individual gene page, where all of the available features of the selected gene are displayed (Figure 6).

Cytosolic non-specific dipeptidase

Date created: 2 february 2012
Last updated: 14 december 2012

UniProt AC Q96KP4
 UniProt Id CNDP2_HUMAN
 Protein name Cytosolic non-specific dipeptidase
 Protein alternative names Glutamate carboxypeptidase-like protein 1; CNDP dipeptidase 2; Peptidase A
 Gene CNDP2
 Gene alternative names CPGL; PEPA; CN2
 Link for protein <http://kb18.ru/protein/matrix?assembly=131>
 Exists in UniProt? Yes
 Last checked in UniProt 28.11.2012 07:00

Descriptors

Descriptor's name	Descriptor's value	Tags
HepG2 iBAQ 11 cell line	7.1399	
Localization on CHR	256	18q22.3
MS data	1.67	
MS data from GPMDB	4	3177
MS data in Peptide Atlas	75	protein

Figure 6 Individual gene/protein page.

Table View

The matrix view can be switch to the tabular format selecting the "Table" hyperlink at the upper toolbar (Figure 7). The table view the numerical data color-coded by the matrix is available. Tooltips and sorting are also available (see below). You can highlight a gene by clicking at the checkbox near the gene name. Highlighting of genes is kept at switching the views.

The screenshot shows two views of the same data. On the left, the 'Matrix' view is active, showing a grid of colored cells (yellow and green) representing data points for various genes. On the right, the 'Table' view is active, showing the same data in a tabular format. The 'Table' view has columns for 'Gene', 'EL', and 'Dr'. The 'Matrix' view has a toolbar with 'Table' circled in red. The 'Table' view has a toolbar with 'Matrix' circled in red. Red arrows point from the 'Table' button in the Matrix view to the 'Matrix' button in the Table view, and vice versa.

Gene	EL	Dr
<input type="checkbox"/> ANKRD20A5P	1	—
<input type="checkbox"/> ABHD3	4	—
<input type="checkbox"/> MC2R	5	1

Figure 7 Switching between matrix and table view.

Tooltips

The matrix/table mode supports three types of tooltips, depending if the pointer is over the row name (gene name), column name (track label) or colored box/value (Figure 8). Row name tooltip provides hyperlinks to the gene information in external resources (e.g. in NeXtPot).

Column name tooltip provides hyperlink to individual track page and color-coding legend. Colored box tooltip store value, some comment and link to external resource, wherein explained as the value was obtained. Both column name and colored box tooltips include date of last update corresponding data. You could see it using [+]. Some additional information, such track description are also available by using [+] on the track label tooltip.

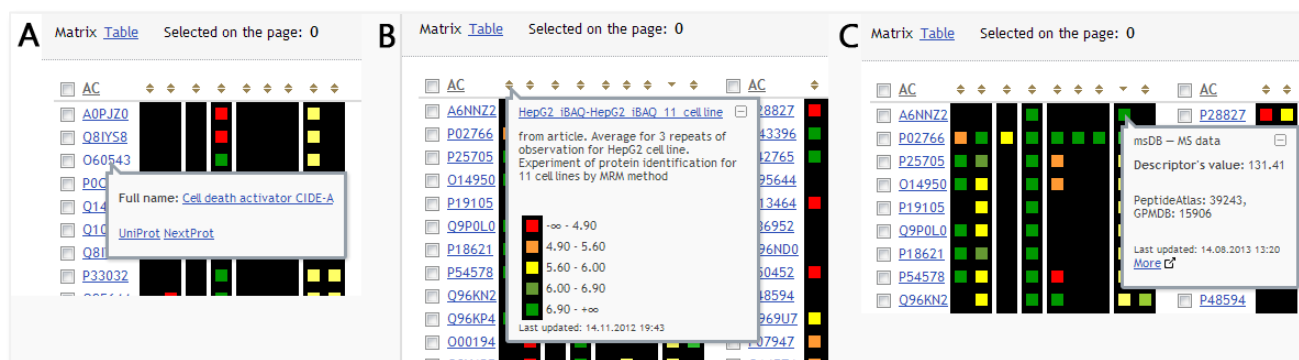


Figure 8 Browsing in the matrix mode – tooltips. A - row name (gene name), B - column name (track label), C - colored box.

Sorting

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

On switching to the table view the sorting will be lost. You can sort the table by using the column name. When sorting the track first time the values are

descending, second time - in ascending order. The null values in the table view are equal to the highest values, so they will be at top in case of descending sorting.

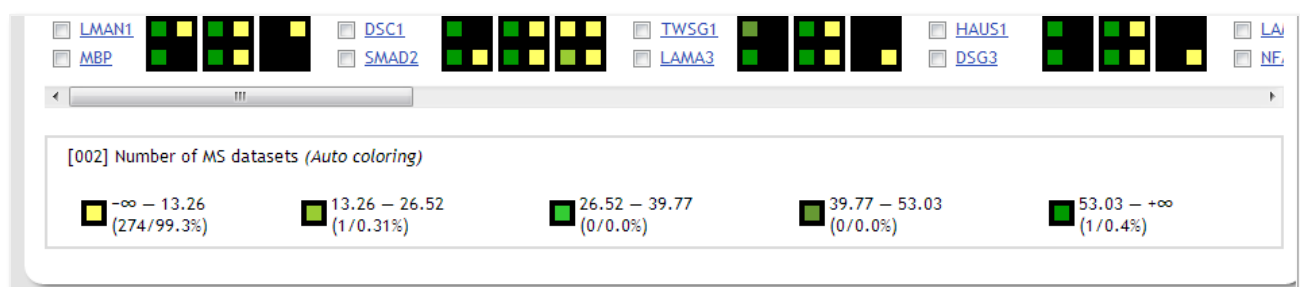


Figure 9 Statistics for the current track (used for sorting the data).

Controls

There is a control bar over the matrix/table pane, providing the features:



add selected genes into the set (section Selecting Genes into the New Set). Inactive if no genes were selected;

[Add from superset](#)

add genes to current set from superset (set of all genes). Available only for set creator;



remove selected genes from current set. Available only for set creator. Inactive if no genes were selected;



export current set as Excel spreadsheet;



change the layout settings (order of the tracks and separators between track groups, more in Change Display Settings).

Selecting Genes into the New Set

Select genes by toggling the checkboxes provided at the left to the gene name in matrix or table browsing modes. The selected genes will be highlighted. All of the genes in the matrix column can be selected by clicking the checkbox in the matrix

header. Number of the selected genes is prompted at the top of the matrix panel (Figure 10).

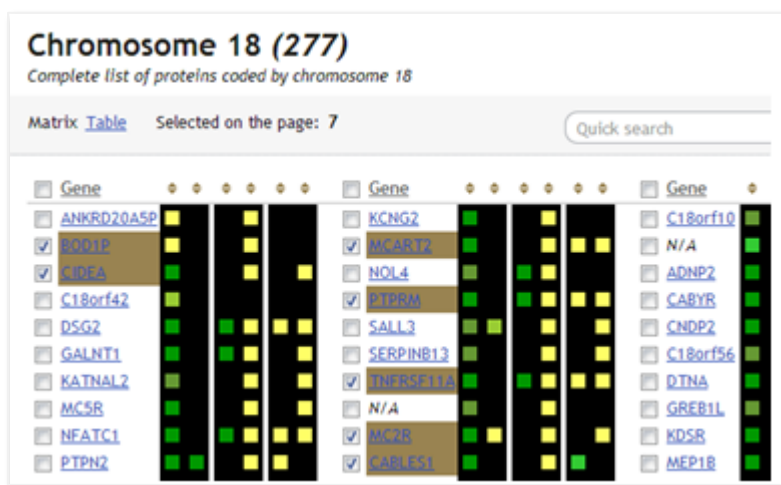



Figure 10 Selected genes highlighted in the matrix view.

After selecting the genes, you can either collect them to the new set, or add to the existing set using the  button in the "Proteins" tab. The dialog box will appear as shown in the figure below (Figure 11). The registered user can select one of the pre-existing sets to add new genes or create a new set. Guest user should check the "Add new set" option to proceed.

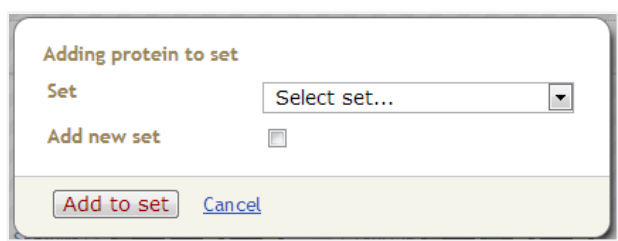





Figure 11 Dialog Box: Add genes to the set.

Sharing the Set



To share the set, select the "Sets" tab and use editing mode by clicking  in the row with your set. Toggle the checkbox "Shared" and save changes. In result the set will be available among the users of one GenoCMS server.

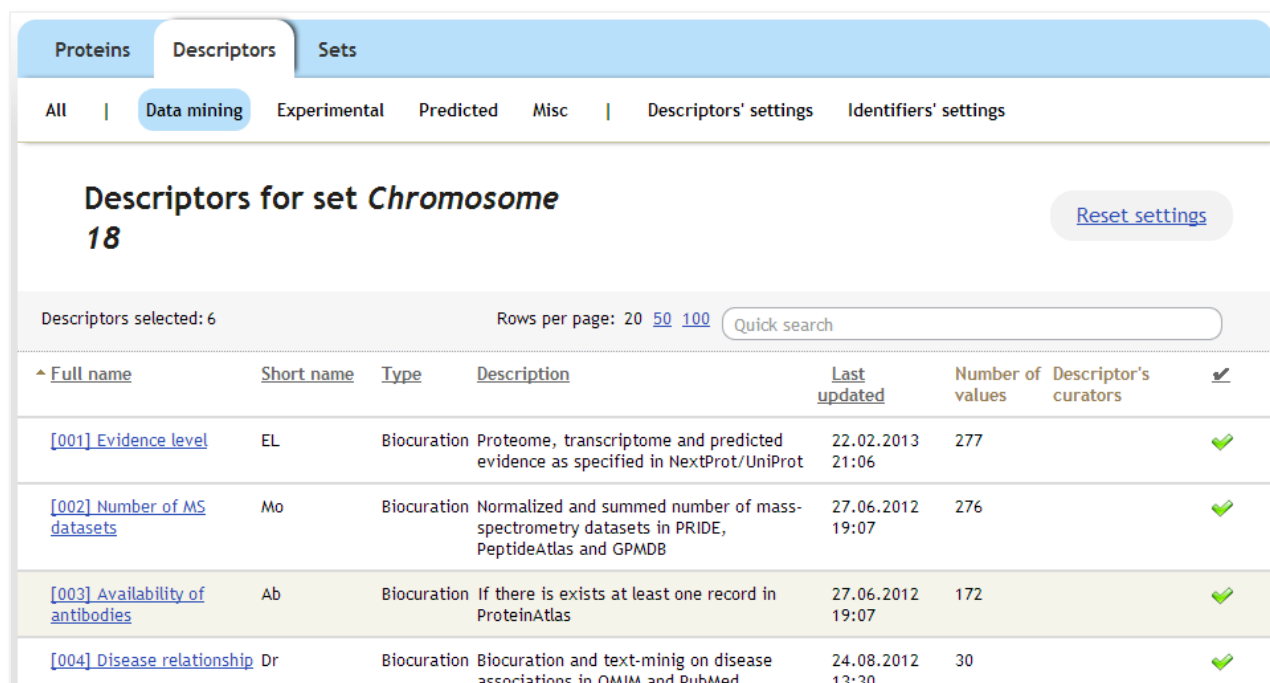
Each set has a Creator's settings (selected tracks, order for outputting the tracks, places for separators and selected gene names style), while these settings can be changed by any user (more in Change Display Settings). If user has modified the settings of the set, the initial setting by the Creator can be viewed by clicking to the "eye" icon . Otherwise, if the set should be broadcasted with the user settings, the appropriate hyperlink is generated by clicking the  icon (page with arrow).

Tracks

Choosing the Tracks

With some current gene set uploaded switch to the “Descriptors” tab. The list of track will be displayed (Figure 12). Every track is assigned to the short name and to the type. The “Biocuration” type indicates the manually updated tracks, while “Web-service” type corresponded to the automatically updatable tracks. For each track the number of the genes with available values is provided.

The currently visible tracks are marked by the green color of the check sign . Click the check sign to unselect the tracks and thus to exclude them from the matrix. Select other track by clicking the grey check sign  to include them into the matrix browsing. Number of currently selected tracks is indicated above the table. After selection of the track switch back to the “Proteins” tab in order to apply the changes.



The screenshot shows the 'Descriptors for set Chromosome' interface. It features a navigation bar with 'Proteins', 'Descriptors', and 'Sets' tabs. Below the navigation bar, there are filters for 'All', 'Data mining', 'Experimental', 'Predicted', and 'Misc', along with 'Descriptors' settings and 'Identifiers' settings. The main heading is 'Descriptors for set Chromosome' with a count of 18 and a 'Reset settings' button. Below this, there is a summary of 'Descriptors selected: 6' and 'Rows per page: 20 50 100' with a 'Quick search' input field. The main table lists tracks with columns: Full name, Short name, Type, Description, Last updated, Number of values, and Descriptor's curators. The first four tracks are highlighted in green, indicating they are selected.






Full name	Short name	Type	Description	Last updated	Number of values	Descriptor's curators	
[001] Evidence level	EL	Biocuration	Proteome, transcriptome and predicted evidence as specified in NextProt/UniProt	22.02.2013 21:06	277		
[002] Number of MS datasets	Mo	Biocuration	Normalized and summed number of mass-spectrometry datasets in PRIDE, PeptideAtlas and GPMDB	27.06.2012 19:07	276		
[003] Availability of antibodies	Ab	Biocuration	If there is exists at least one record in ProteinAtlas	27.06.2012 19:07	172		
[004] Disease relationship	Dr	Biocuration	Biocuration and text-minig on disease associations in OMIM and PubMed	24.08.2012 13:30	30		

Figure 12 List of tracks.

It is useful to sort the selected track by clicking on the header of the last column (underlined “check” sign ). The checked tracks will be gathered at the top of the list and can be observe on one screen. “Reset settings” hyperlink unselects all tracks with one click. That is useful when starting to arrange the tracks in accordance with your preferences.

The tracks are spread over several categories, which are displayed on the “Descriptors” tab. Categories foster the navigation to a particular group of tracks, e.g. tracks acquired from international resources (“Data mining”), from the pre-published or published data (“Experimental”) or other.

Creating the Track (Curators)

To create a track, you should send a request to administrator (support@kb18.org). Obligatorily include into the e-mail user name, e-mail address, which were specified at registration, the number of tracks necessary for you, their names and types (web service or biocuration).

Administrator will make you Curator and give to you hyperlinks to manage of your tracks and instructions for curators.



Settings

Change Display Settings

There are three types of settings to change:

- visibility of the tracks (section Choosing the Tracks);
- order of the tracks and separators between track groups (section Order and Separators);
- visibility and order of the genes names (section Style of Gene Naming).



GenoCMS stores the set's display settings for each user. The owner's display settings of the set are copied to the user while visit the set first time. So:

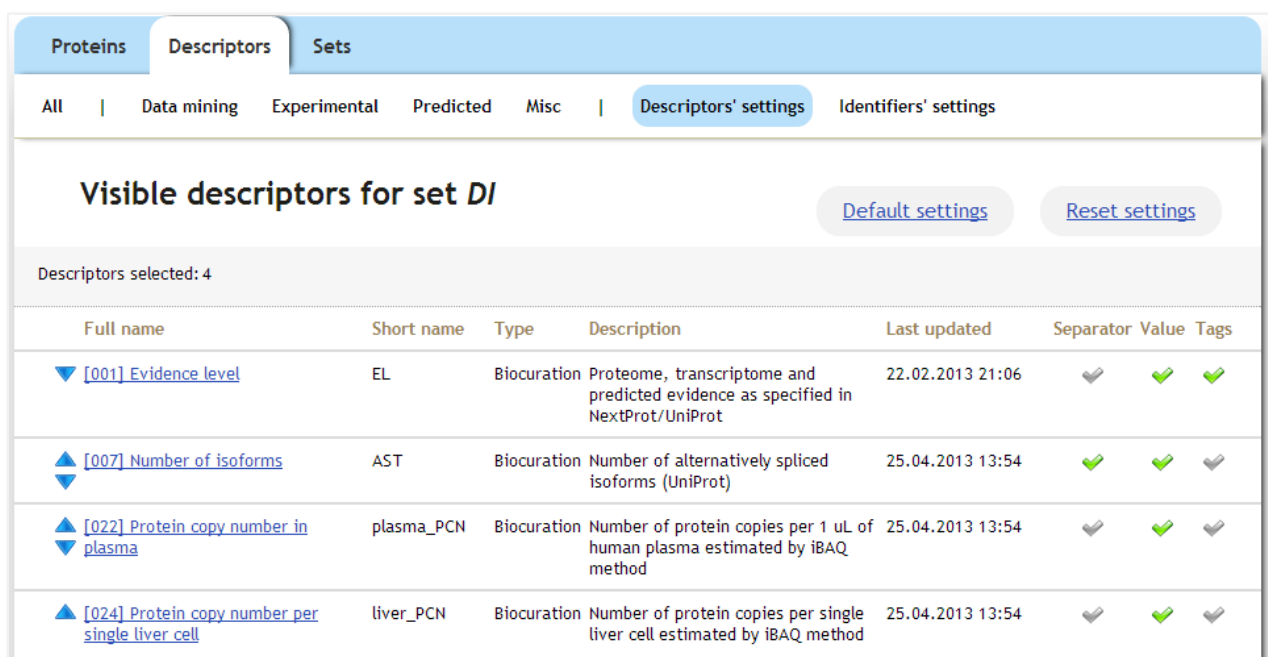
- To access to the set with display settings of the current user select tab "Sets" and click on its title.
- To get the URL for broadcasting the set click  in the row of corresponding set (tab Sets). The set will be accessible to the any recipient of the URL with the same display settings, as you have selected.
- To view the set with display settings of owner (creator) select tab "Sets" and click  in the row of corresponding set. The set will be loaded in Workspace.

Order and Separators

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

- Separator (show separator after track column, white line on matrix, gray background color on table view);
- Value (show value on table view);
- Tags (show tags on table view).

If both Value and Tags unselected, track will be invisible for current set. Check sign  matches switched-on toggle,  - switched-off.



Proteins | **Descriptors** | Sets

All | Data mining | Experimental | Predicted | Misc | **Descriptors' settings** | Identifiers' settings

Visible descriptors for set DI [Default settings](#) [Reset settings](#)

Descriptors selected: 4




















Full name	Short name	Type	Description	Last updated	Separator	Value	Tags
 [001] Evidence level	EL	Biocuration	Proteome, transcriptome and predicted evidence as specified in NextProt/UniProt	22.02.2013 21:06			
 [007] Number of isoforms 	AST	Biocuration	Number of alternatively spliced isoforms (UniProt)	25.04.2013 13:54			
 [022] Protein copy number in plasma 	plasma_PCN	Biocuration	Number of protein copies per 1 uL of human plasma estimated by iBAQ method	25.04.2013 13:54			
 [024] Protein copy number per single liver cell	liver_PCN	Biocuration	Number of protein copies per single liver cell estimated by iBAQ method	25.04.2013 13:54			

Figure 13 "Descriptors' settings"

The order of the tracks can be changed by swapping the neighboring tracks in the list. Use  button to swap the tracks for arranging them into the desired order.

"Default settings" hyperlink returns to the selection of the tracks, recommended by the administrator. "Reset settings" hyperlink unselects all tracks with one click.

Style of Gene Naming

To change the style of gene naming adjust the settings at "Descriptors" >> "Identifiers' setting" Web-page of GenoCMS (Figure 14).

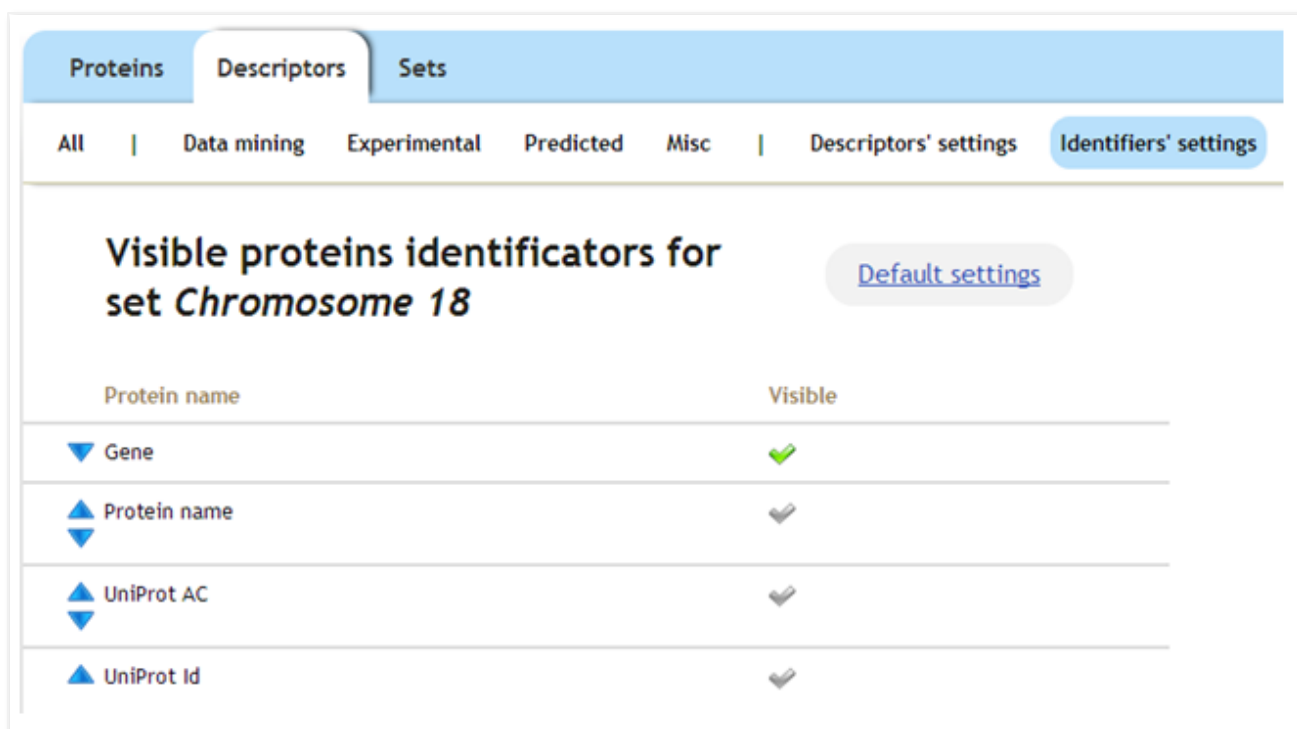


Figure 14 "Identifiers' settings"

Any number of gene names can be selected to show by click ☐. To unselect use ✓.

The order of the gene names can be changed by swapping the neighboring row in the list. Use ▲ button to swap the gene name for arranging them into the desired order.

Search





Quick and Extended Search on Matrix/Table View

The Quick Search uses the keyword co-occurrence within the current set.

The Extended search sends the query first to UniProt (<http://www.uniprot.org/>), retrieves the Accession Codes for found UniProt entries and filters the matching proteins from the current set.

Highlighting of genes will be lost after searching.




Search the Set or the Track

Searching the sets or the track foster the navigation. To find the set or the track, you can enter some part of their title or description. Specials controls, like    , work from search results.

Sample Scenario

User and Creator Setting of the Set

For understanding how it works, follow the scenario:

1. Login with Guest access
2. Go to tab "Sets"
3. Select the set *Chromosome 18* (Figure 15)
4. Go to tab "Descriptors"
5. Click "Reset Settings" hyperlink
6. Select some tracks using  (Figure 16)
7. Go to tab "Protein" (Figure 17)
8. Go to tab "Sets"
9. Click  for the set *Chromosome 18* (Figure 18)
10. Go to tab "Sets"
11. Click  for the set *Chromosome 18*
12. Copy the link, and use it in web-browser (Figure 19)

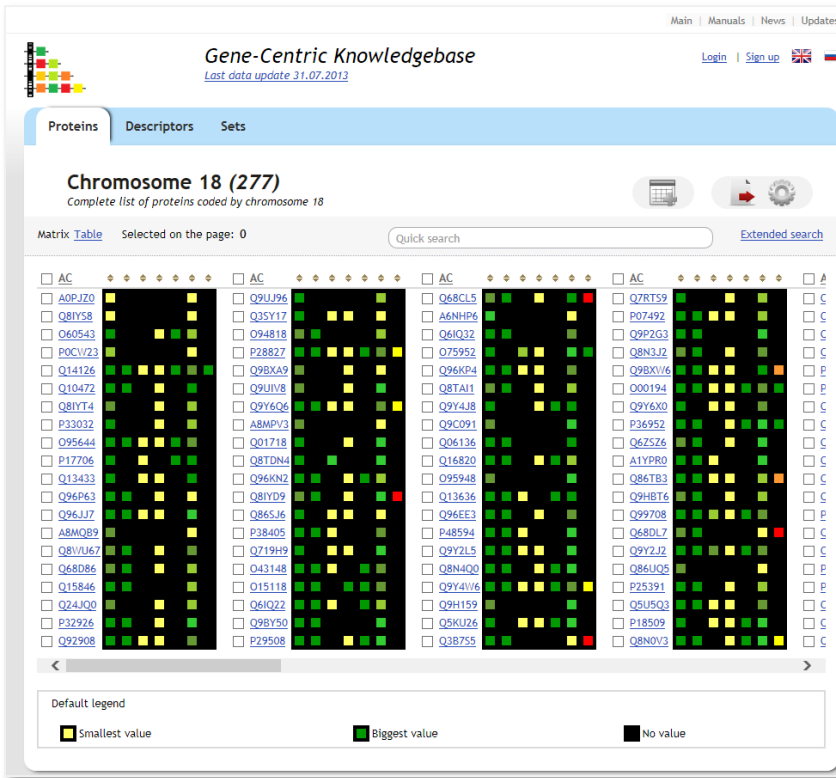


Figure 15

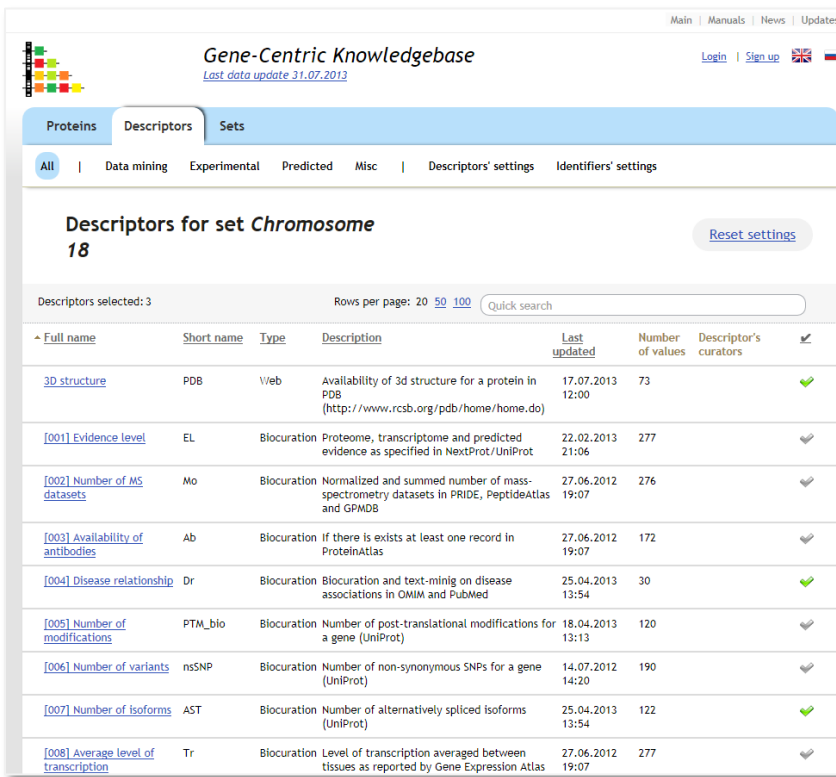


Figure 16

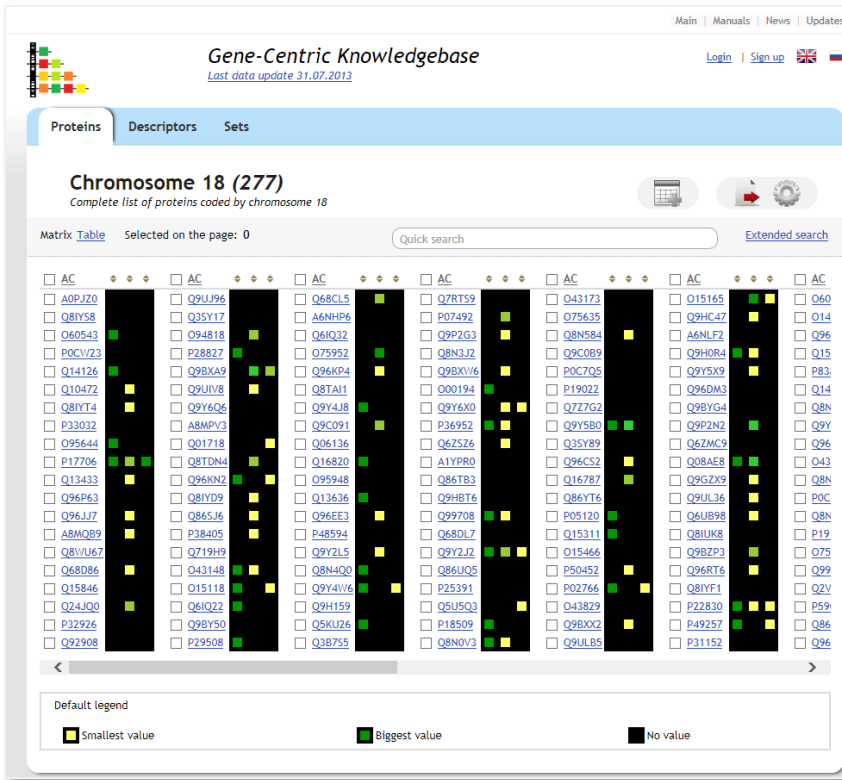


Figure 17

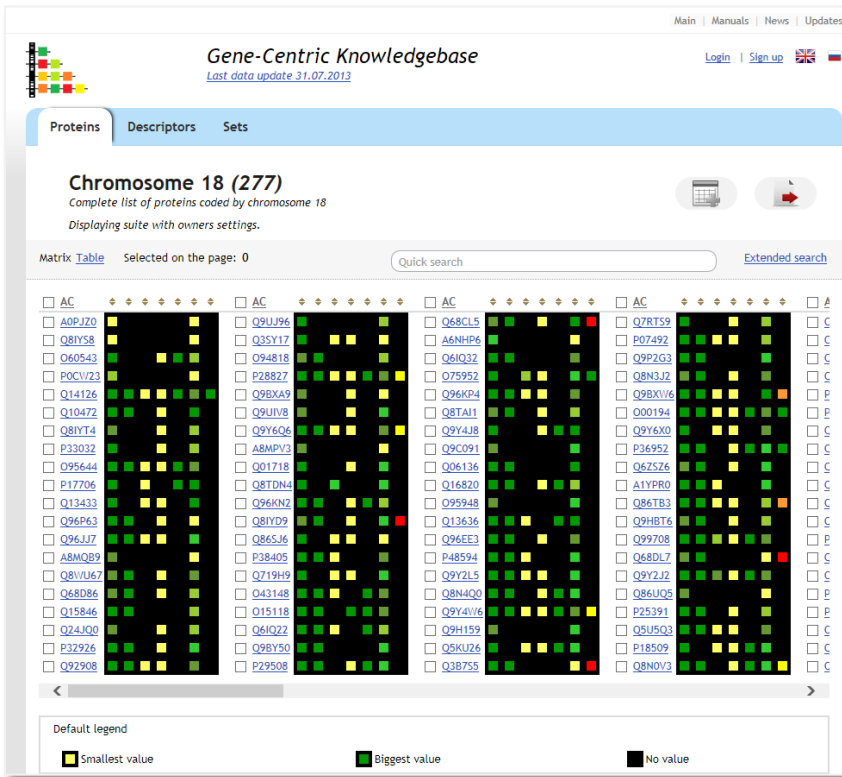


Figure 18

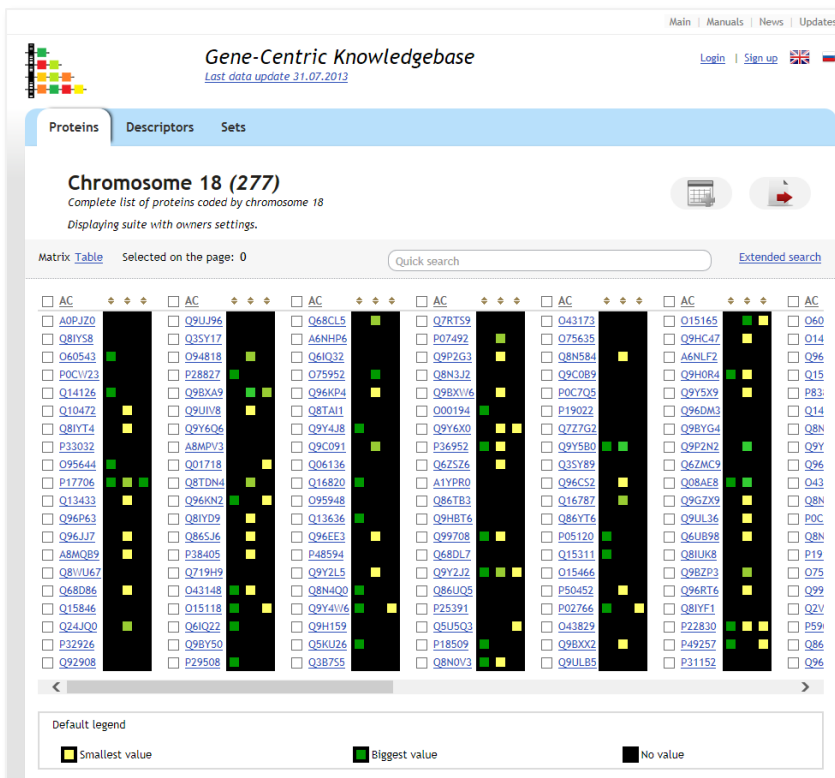


Figure 19