MaGnET: Malaria Genome Exploration Tool

Free software tool for visualisation and exploration of integrated *Plasmodium falciparum* 3D7 functional genomics datasets

Home · Tutorial for MaGnET 2.0

Tutorial

For your convenience, consider printing out these pages before you begin or download a pdf-print here.

Starting MaGnET

Starting the Java Web Start version:

1. Open the downloaded JNLP file with Java Web Start Launcher.

2. You will be asked to choose a database mirror to connect to. Select the mirror closest to you from the drop-down menu.

3. A check box asks whether you would like to load the genomic data at start-up. By default, this check box is already ticked (keep it ticked for this tutorial).

*N.B. This is recommended in most cases, as it saves significant time later if you are going to be using the genome browser functionality of MaGnET. It is particularly recommended if you are accessing the database using a wireless connection and/or are not very close to the database location. It is **strongly** recommended that you try to access MaGnET over a fast broadband connection. Standard wireless broadband connections may not achieve satisfactory download speeds.*

4. Click on 'Connect'. The application will then proceed to connect to the database you have chosen and load the genomic data into memory.

5. When MaGnET starts up two windows open, the MaGnET Front Page with control panels and menus to open various sections within MaGnET and control its behaviour, and the MaGnET Window Control Panel.

MaGnET Window Control Panel

The MaGnET Window Control Panel is a pop-up panel visible at the front of the screen which contains thumbnail icons showing the application windows that are currently open. Clicking on a particular window icon in this panel will bring that window back to the front of the screen. If you close this panel it can be reopened at any time by clicking on one of the 'Window Control Panel' buttons in the MaGnET viewers or on the Front Page.

MaGnET Front Page

Load a group of genes from a file of gene identifiers:

1. Save this file to your local file system as a plain text file: antigen_genes.txt.

Important note regarding saving as text - recently updated versions of the common Web browsers (e.g. Firefox) may be saving this file as a "complete web page" unless instructed otherwise even if it is displayed as a text file in the browser. We recommend to either save the file directly from this page, holding down the ctrl-key so you can choose to save as text, or to copy-paste its contents from the browser to a plain text file (e.g. in Word, saving as .txt). If in doubt, view the file before you load it into MaGnET, to make sure it contains the gene list.

2. In the 'Groups' menu, click on 'Load from file into group A'.

Important note regarding the Java file-chooser dialog window - *in the earlier versions of Java 1.6 there was a bug causing file chooser windows to load and open folders very slowly.   This problem has been improved in the more recent releases of Java. Therefore, if you experience particularly a slow-running file chooser, try updating your Java version.*

Please be patient while loading the file, depending on your location and internet traffic this could take several minutes.
If the file were to not load within this time, and you are sure that it was saved as text file (you should see a list of genes when viewing it), consider using this (smaller) sample set as an alternative: antigen_genes_sample.txt.

Genome

To view the genomic location of the selected genes in group A:

1. In the light blue 'Genome' panel, select '*P. falciparum* 3D7 genome' and click 'Launch viewer'. The genome viewer should open displaying the 14 nuclear

chromosomes, the mitochondrial and the plastid (apicoplast) chromosomes. The location of the genes in group A are highlighted as orange lines. Lines on the right side of the chromosome are genes on the forward (w) strand and lines on the left side are on the reverse (c) strand.

2. Find the location of the gene PF3D7_1021900 (formerly known as PF10_0213) by typing its identifier into the search bar at the top of window, and pressing 'Find'. It will now be highlighted in purple on chromosome 10.

N.B. The MaGnET search function will recognise current gene IDs (such as PF3D7_1021900) as well as previous gene IDs (such as PF10_0213).

3. To view chromosome 10 in closer detail, double click on it.

The top panel displays the forward strand as the top bar (white) and the reverse strand as the lower bar. Protein coding genes are shown in medium grey.
The middle panel shows an overview of the chromosome (white) and the light grey region shows which part of the chromosome is being viewed in the top panel.

4. Scroll along the chromosome until reaching the purple-highlighted gene PF3D7_1021900. Click on the 'Clear' button at the top of the window. PF3D7_1021900 should revert to orange.

5. Under the 'Genes' menu, click 'Introns'.

6. To display introns the scale must be equal to or less than 50 bp/pixel. Zoom in to 30 bp/pixel by pressing the 'Zoom in' button. The introns are displayed as pink regions within a gene.

To view the gene's product name:

7. Click once on the gene PF3D7_1021900. Its product name should appear at the top of the panel.

To find out more about a particular gene, view its factsheet:

8. Double click on gene PF3D7_1021900.

Gene Factsheet

1. Look down the first page ('Gene/Protein data'), you should see some genomic information, including a list of exons, near the top.

2. Below that is a dark grey bar, representing the gene's protein product or predicted product. The integer given on the right is the length of the protein.

3. In the following panels, any experimentally solved and comparatively modelled protein structures, respectively, are displayed as magenta bars below the corresponding region of the protein (see the section *'View 3D protein structures'* for further information).

4. At the bottom of the page the Gene Ontology annotation (if any) is displayed.

5. Click on the 'Additional annotation' tab at the top of the window. Scroll down the page to see the InterPro predicted sequence features: domains and motifs are orange, transmembrane regions are blue, low complexity regions are green, coiled coils are red and signal peptides are cyan.

6. Click on the 'Ortholog/paralog group' tab (you may have to click on the right-pointing arrow first, to see the additional tabs if they are hidden). This page shows any OrthoMCL-predicted paralogs within the genome and any orthologs within related organisms.

7. Click on the 'Interactions' tab. The table lists five proteins that are predicted to interact with this protein from a yeast-two hybrid screen. The times observed and number of searches information gives an indication of how reproducible the interaction was.

8. Click on the 'Links to other resources' tab. Here, linkouts to view this gene in other online resources are available by clicking on the respective buttons (opens a web browser).

View 3D protein structures: (You can skip steps 9-11 to save time if you wish)

9. Return to the 'Gene/Protein data' tab.

10. Scroll down to the section where '3D structure models' are displayed. Click once on one of the magenta bars. Some information about the model should appear at the bottom of the panel.

11. Double click on the same bar. The Jmol 3D structure viewer should open in a moment, with the model loaded and displayed as a cartoon representation of the protein backbone's secondary structure elements. For a Jmol tutorial click on the 'Help' menu in the top right corner.

Important note regarding the Jmol viewer - if you close the Jmol viewer the MaGnET program will exit also. This is because when you close one Java program all currently running Java applications will be closed together.

12. Return to the chromosome 10 window for the next part. (You may close the gene factsheet but it is recommended you minimise rather than close the Jmol viewer for now - see above note). Scroll back to the start of chromosome 10 (left side).

To view the genomic location of a gene's paralogs:

1. Under the 'Compare genomes' menu, select 'View ortholog/paralog group'.

2. Click once on gene PF3D7_1000100. It should now be highlighted in brown and a list of its orthologs will be displayed in the bottom panel. PfEMP1 has a large number of paralogs in the genome. Those on chromosome 10 will be highlighted in brown on the chromosome.

3. Return to the whole genome window. Paralogs of PF3D7_1000100 are highlighted in brown.

4. Return to the chromosome 10 window. Go to the 'Compare genomes' menu and untick 'View ortholog/paralog group'. Close the chromosome 10 window and return to the whole genome view window.

Protein-protein interactions

Examine the interaction network of the proteins in the selected group:

1. Under the 'Open' menu click on 'Protein-protein interactions'. A new window should open displaying all the protein-protein interactions in which the genes in group A are involved. A line joining two boxes represents an interaction between these two proteins.

2. Click the 'Length +' button to increase the length of the lines between the proteins. The line length has no significance, but longer lines give the network more space.

3. Stop the network from moving by pressing the 'Stop' button. Start it again by pressing the 'Start' button.

Search for a protein within the network:

4. Under the 'Display' menu, click 'Find'. Now, type the gene identifier PF3D7_1021900 into the text field at the bottom of the window and press the 'Display' button. The gene product of PF3D7_1021900 should now be highlighted in purple in the network.

Quickly view the name of the proteins:

5. Under the 'Click' menu, click on 'Name'. Now click on one of the proteins in the map. The name of the protein should appear in the text field at the bottom of the window.

Quickly view the cellular component of the proteins:

6. Under the 'Click' menu, click on 'Component'. Now click on one of the proteins in the map. The protein's GO component annotation should appear in the text field at the bottom of the window.

Quickly see which proteins interact with a particular protein:

7. Under the 'Click' menu, click on 'Clusters'. Now click on one of the proteins in the map. It and any proteins it interacts with should be highlighted in purple.

Expand the network to bring in other proteins that also interact with the proteins in the existing map (indirect interactions with group A proteins):

8. Under the 'Click' menu, click on 'Extend'. The labels of the proteins should now display a number after the gene identifier - the total number of proteins which the protein interacts with (you may need to press 'Start' at this point).

9. Click on a few proteins and notice that new proteins are added to the map.

Remove some of the detail and examine part of the network:

10. Click on the 'Size' button and then the 'Stop' button. The protein labels should now become small squares.

11. Under the 'Click' menu, click on 'Expand'. Click on one of the orange proteins. The labels of this protein and any it directly interacts with should now increase in size.

12. Click the 'Size' button again to return all the labels to their normal size.

Add proteins to a new group:

13. Locate the protein 'PF3D7_0422200'. Select all the proteins that interact with this protein by using the 'Clusters' option from the 'Click' menu (as described

earlier).

14. Under the 'Groups' menu, click on 'Replace group B with selected'. The purple-highlighted genes should now be coloured blue.

Examine the expression profiles for the genes in group B:

15. Under the 'Open' menu click on 'Expression Data', the Expression Data viewer should open in a new window.

Expression Data

1. Under the 'Expression' menu, and the 'Datasets' submenu, select the dataset 'P.fal 3D7 LeRoch et al. (2003) (mRNA) asexual stages treated with sorbitol + gametocytes & sporozoites'.

2. In the text field in the middle of the page type the letter 'B'. Click the 'Display time-series graph' button. The expression profiles for all genes in group B will be displayed. A message will first appear telling you if any of the genes in the group do not have an expression profile in this dataset.

3. Zoom in on the upper part of the graph by clicking close to the top of the graph and dragging downwards while holding down the left mouse button.

4. Zoom out by clicking anywhere on the graph and dragging the mouse upwards.

5. Close the graph window.

View expression profile for a single gene of interest:

6. Under the 'Expression' menu click 'Clear current selection' to clear the selected dataset.

7. Under the 'Expression' menu, and the 'Datasets' submenu, select the dataset 'P.fal 3D7 Otto et al. (2010) (RNASeq) - IDC 0-48 hrs'.

8. In the text field type the gene identifier PF3D7_0422200 and click the 'Display time-series graph' button.

9. Close the graph window.

Query the expression data for genes matching certain criteria:

10. Under the 'Query builder' menu, click on 'Search the expression data'. A new window should open.

Query builder

Search for genes with mRNA decay half-life greater than 60 minutes at the ring life cyle stage :

1. In the fourth search panel ('Search for genes with a particular mRNA decay rate'), set the statement to read 'Search for genes with mRNA decay half life *greater than 60.0* minutes over 3D7 IDC timepoints *ring hr 10-14*'. Click 'Search'. A table of results should be returned.

2. Select the gene 'PF3D7_0422400' and click on 'View time-series graph'. A new window should open that shows the mRNA decay profile of this gene against its recorded expression level at particular stages of the IDC in a comparable dataset from the same research group. Note that the mRNA decay data is in some cases recorded for multiple oligos per gene and there may be some timepoints where particular oligos are excluded because the signal did not meet quality controls. You can also access this graph from the Expression Data window, by selecting 'View mRNA decay profile for IDC' under the 'Expression' menu and entering the gene id PF3D7_0422400 in the text field.

3. Close the graph window.

Search for genes with expression level increasing during the early intraerythrocytic-development cycle (IDC):

4. In the third search panel ('Search for genes whose expression changes between certain timepoints'), set the statement to read 'Search for genes which *increase > 3 fold* between *hour0* and *hour16* in experiment *Otto RNASeq 3D7 IDC 0-48 hrs (2010)*'. Click 'Search'. A table of results should be returned.

5. Under the 'Groups' menu, click on 'Clear groups'. All previously saved groups will be cleared.

6. Highlight all genes in the table. Under the 'Groups' menu, click on 'Replace group A with selected'.

7. Close the Query builder and Expression Data windows. Return to the Genome window.

Genome

View expression data overlaid onto genomic location:

1. Under the 'Expression' menu, click on 'Show expression data for groups only'.

2. Under the 'Expression' menu, click on 'Select a new mRNA or protein expression dataset'.

3. In the option dialog choose 'RNASeq - P.falciparum 3D7 - Otto et al. (2010) - IDC 0-48 hrs'. Please be patient as the dataset may take a while to load. The expression data for genes in group A will be displayed as coloured lines on the choromosomes. The colour scale represents a gene's expression level at the current timepoint coloured by its rank in the interquartile range of the gene's expression over the entire time-series. The genes in group A are represented by the orange lines to the sides of the chromosomes.

4. Move through the time-series by pressing the '+' button or dragging the slider along in the lower panel.

5. Open one of the chromosomes by double clicking on it. Notice how the genes in group A are coloured orange in the upper quarter of the gene.

View changes in expression data:

7. Under the 'Expression' menu, click on 'Colour according to each gene's change in expression from the previous timepoint'. The colour scale now represents the change in a gene's expression from the previous timepoint (if there's been a signifcant change greater than two-fold). Note: at the first timepoint it is not possible to calculate a change so the genes are coloured dark grey.

See a summary of information about the genes stored in the groups:

1. In the 'Open' menu, click 'Data Analysis' to open the query facility.

[Data Analysis](#)

2. Under the 'Groups' menu, click on 'Get group information'. All genes in the group are now listed with their curated and Gene Ontology (GO) annotation. Each GO annotation is displayed as a separate row in the table, with the result that an individual gene can span several rows.

Search for genes with a function of interest and add them to a second group. To search for all genes with the GO annotations "proteolysis" or "protease":

1. In the 'Advanced Search' panel (the second panel from the top), select 'Gene Ontology terms' under the menu, and under 'Search options' select 'OR' and 'Keywords'.

2. In the search field enter the phrase "proteolysis,protease" (without quotes). Multiple terms can be searched by separating with a comma.

3. Click on the 'Submit search' button. The results of the search should be returned as a table.

Add the returned genes to a group:

4. Highlight all the genes in the table (click on the table and press *Ctrl* and *A* on Windows or drag the mouse down the table).

5. In the 'Groups' menu, click on 'Replace group B with selected'. The table's gene_id column should now be coloured blue.

Search for further proteases that do not have a GO annotation for proteolysis:

6. In the 'Quick search' panel (top), under 'Search options' select 'OR' and 'Keywords', and enter the terms "proteolysis,protease" in the search field.

7. Click on the 'Submit search' button. The results of the search should be returned as a table. Notice that further genes are returned which do not yet have a GO annotation for 'proteolysis' or 'protease'.

Add the returned genes to a group that already contains genes:

8. As 4 above. In the 'Groups' menu, click on 'Append selected to group B'. Now the results of the two searches are joined together in one group.

9. Under the 'Groups' menu, click on 'Get group information'.

Finally, save the genes in group B to a file on your computer (optional):

1. Return to the MaGnET front page.

MaGnET Front Page

Assign a name to the genes in the group (This is optional and can be done at any time):

2. Under the 'Groups' menu, click on 'Group B title' and enter a name of your choice, e.g. 'protease'.

Save the genes in group B to a file:

3. Under the 'Groups' menu, click on 'Save group B to file'. Find a location on your computer to store the file and click 'Save'.

4. You can use the load options under the same menu to load the saved genes into MaGnET next time you wish to explore them.

Close the MaGnET Front Page window to exit the MaGnET application.

Navigation

- [About](#)
- [Software](#)
- [Tutorial](#)
- [Datasets](#)
- [Contact](#)
- [Links](#)

Design by [Minimalistic Design](#)