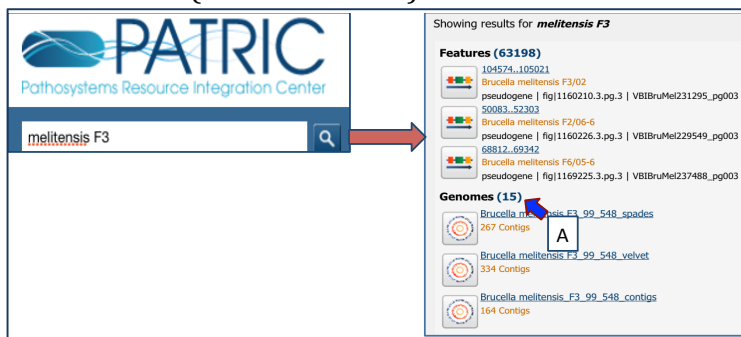


Protein Family Analysis: Creating genome groups

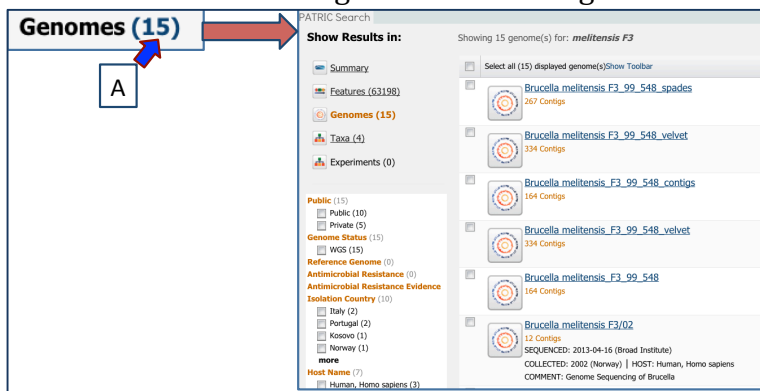
- I. Creating a group with a private genomes-Global Search
- II. Creating a genome group with private genomes-Taxon landing page

I. Creating a group with a private genome-Global Search

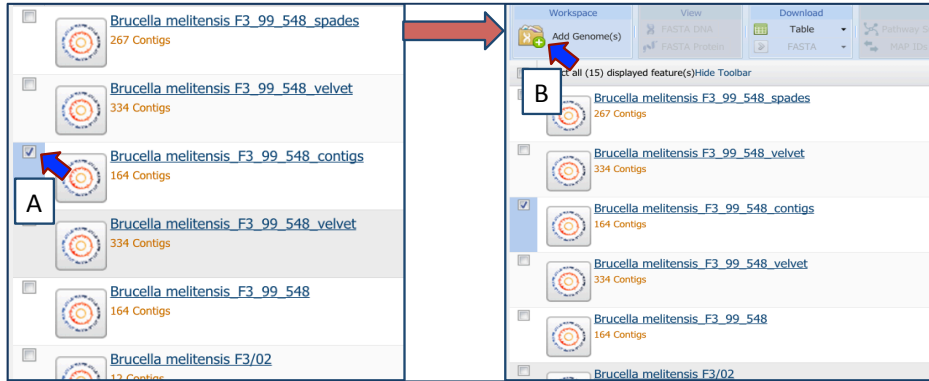
1. The first thing you will need to do is to find the genome that you annotated in PATRIC. The search function allows you to submit some key text terms to narrow the search for your genome. In the example below, to find a genome named *Brucella melitensis* F3 99 548, I enter “melitensis” and “F3” as seen in the screenshot below and hit return. This will open up the Search Results page that is divided into Features, Genomes, Taxonomy and Experiments. The results returned for Genomes are indicated (Arrow A below).



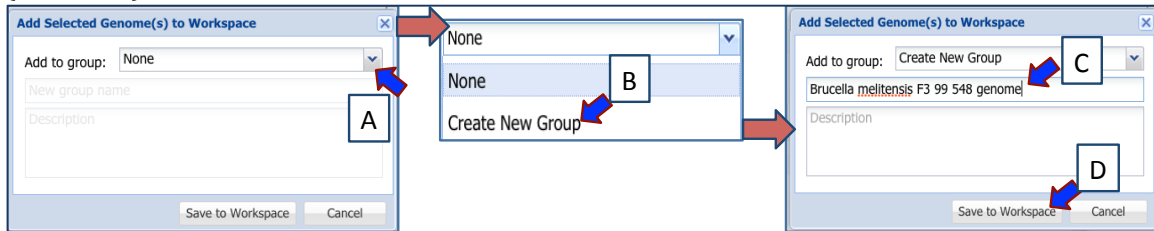
2. To see all the results, click on the number that follows “Genomes”, as indicated (Arrow A, below). This will open up a page that contains all the genomes that were returned by your search parameters. You will see a filter on the left side of the page, and a table with a list of genomes on the right side.



3. Find the genome that you want to save as a group and click on the box in front of it (Arrow A, below). You can click as many boxes as you want to specialize your group. This will open a header on the top of the table that allows you to save a group of genomes. To create a group that contains genomes this genome, click on the Add Genomes button at the top of the table (Arrow B).



4. This will open up a pop-up window. To create a group, you will need to click on the down arrow that follows the word “None” that you can see in the text box (Arrow A in the screenshot below). This will open up a drop down box that allows you to create a new group, or to add genome to an existing group. To create a new group, click on the line that says “Create New Group” (Arrow B). Then you must type in the name that you want to identify the group in the text box (Arrow C). To save that group, click on the Save to Workshop button at the bottom of the pop-up box (Arrow D).



II. Creating a group with public and private genomes-Taxon landing page

1. Let’s say that you wanted to create a group with your private genomes, and compare them with some of the genomes that are freely available in PATRIC. You can do that by using the global search, but sometimes its just as easy to go in by some of the PATRIC taxonomy pages and see the genomes that are available. In this example we will look at the species level. I want to see all the genomes available for *Brucella melitensis*, so I will enter “melitensis” in the global search box and return. You can do this for any taxon level, from strain all the way up the taxonomy to kingdom. This will open up the Search Results page that is divided into Features, Genomes, Taxonomy and Experiments. The results returned for Taxonomy is indicated (Arrow A below). As we are interested in *Brucella melitensis* we will need to select that by clicking on that name in the search page (Arrow B). You might have to scroll down the page to see Taxonomy.

Showing results for *melitensis*

Features (368693)

- 134462_134625
Brucella melitensis strain C-554
misc_binding
- 173596_173698
Brucella melitensis strain C-554
misc_binding
- 173781_173781
Brucella melitensis strain C-554
misc_binding

Genomes (60)

- Brucella melitensis bv. 1 str. Rev.1
22 Contigs
SEQUENCED: 2009-02-18T00:00:00Z (Broad Institute)
DISEASE: Brucellosis
COMMENT: Brucella melitensis bv. 1 str. Rev.1. Brucella melitensis bv. 1 str. Rev.1 is used as a live attenuated vaccine for immunization of livestock.
- Brucella melitensis bv. 3 str. Ether
11 Contigs
SEQUENCED: 2009-02-18T00:00:00Z (Broad Institute)
DISEASE: Brucellosis
COMMENT: Brucella melitensis bv. 3 str. Ether (ATCC 23458) is the reference strain for this biovar and will be used for comparative analysis with other Brucella species.
- Brucella melitensis bv. 1 str. M111
47 Contigs
SEQUENCED: 2011-04-26T00:00:00Z (Chinese Academy of Sciences)
DISEASE: Vaccine
COMMENT: Brucella melitensis bv.1 str. M111 is the vaccine strain using in China against brucellosis in animals. It was developed by Chinese Institute of Veterinary Drug Control.

Taxonomy (97)

- Brucella melitensis biovar
0 Genome(s)
- Brucella melitensis (species)
59 Genome(s)

2. This will take you to the landing page for that species. Across the top of the page, you will see a number of tabs. You will need to click on the Genome List tab (Blue Arrow A).

Bacteria > Proteobacteria > Alphaproteobacteria > Rhizobiales > Brucellaceae > Brucella > Brucella melitensis (3)

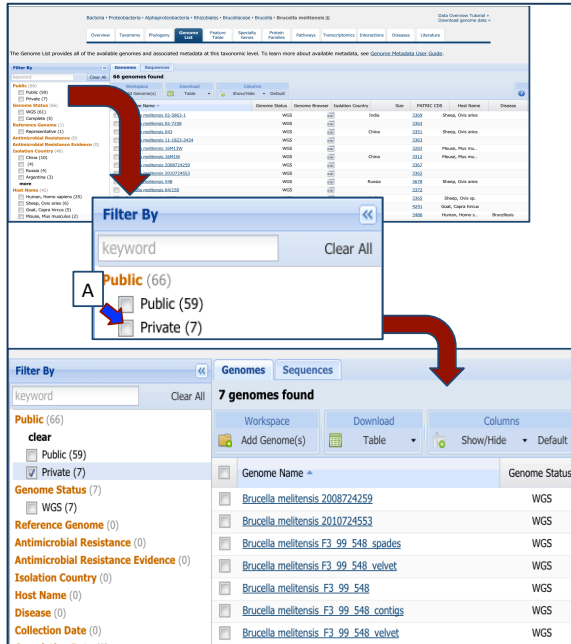
Overview Taxonomy Phylogeny Genome List

Search Tools: Genome Finder, Feature Finder, Comparative Pathway Tool, Protein Family Sorter

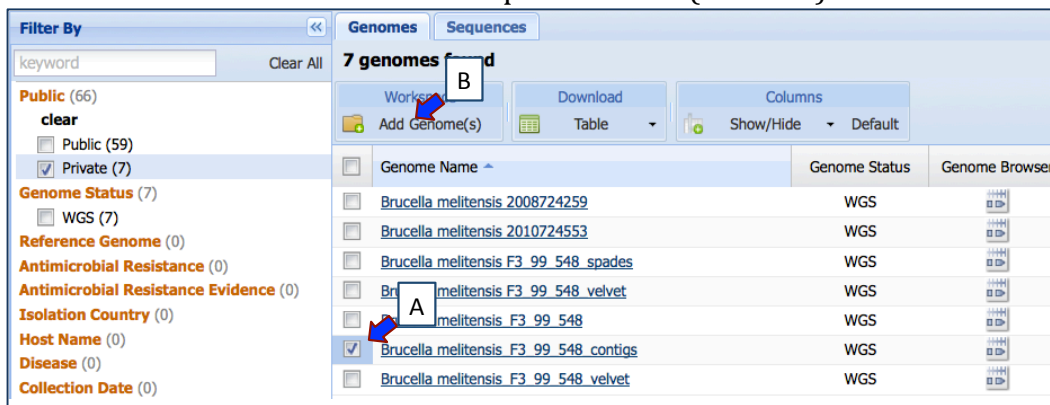
Experiment Summary: Transcriptomics from GED (30), Transcriptomics from ArrayExpress (28), Proteomics from PRIDE (0), Structure from NCBI (38), Protein-Protein Interaction from IntAct (0)

Taxonomy Summary: Taxonomy ID (2959), Lineage (cellular organisms > Bacteria > Proteobacteria > Alphaproteobacteria > Rhizobiales > Brucellaceae > Brucella > Brucella melitensis), External Links (Genome Ecoloop Database and Analysis Resource), Summary Terms (Public (59), Genome Status (WGS (54), Complete (1)), Reference Genome (Representative (1)), Antimicrobial Resistance (Antimicrobial Resistance), Evidence (Isolation Country (China (10), (5), show all 46 genomes), Host Name (Human, Homo sapiens (23), Sheep, Ovis aries (5), show all 42 genomes), Disease (Brucellosis (8), Vaccine strain (2))

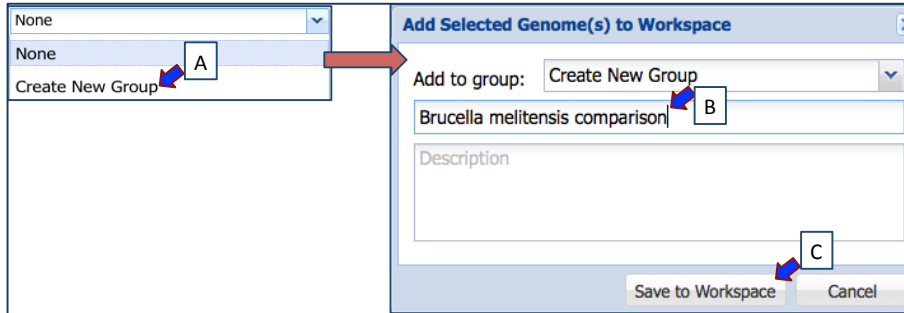
3. This takes you to the page that contains a list of all the genomes within the species *B. melitensis*, and a filter on the left that you can use to narrow the search. At the top of the filter, you will see that you can choose to limit your search to genomes that you have annotated in PATRIC (denoted as Private), and those that are freely available in PATRIC (denoted as Public). If you wanted to select one or more of the genomes that you annotated in PATRIC, you could do this by clicking on the box in front of the word “Private” (Blue Arrow A, below). This will change the table to only show you the genomes you have annotated that fall under the *B. melitensis* taxonomy designation.



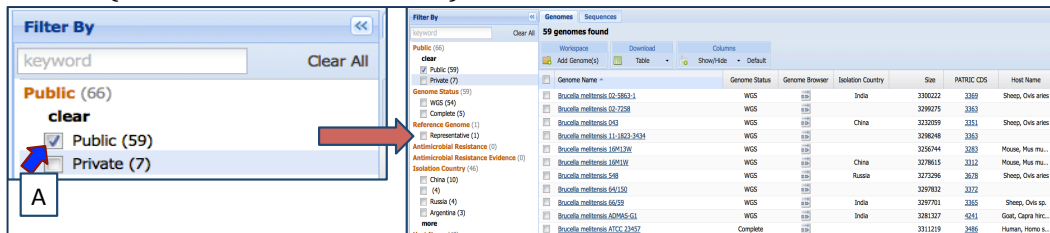
4. You can select one or more of the genomes by clicking on the box in front of the name (Arrow A, below). To save your choice(s) to a group, you need to click on the Add Genomes button at the top of the table (Arrow B).



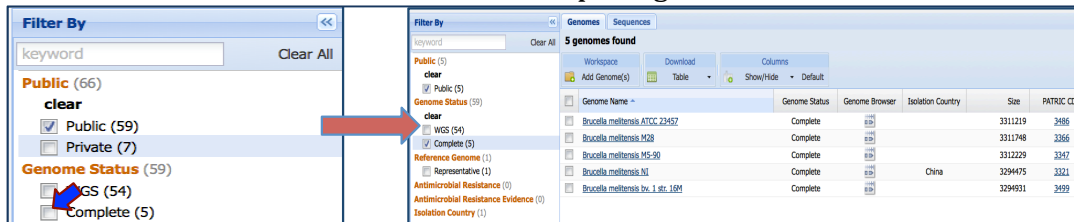
5. This will open up a pop-up window. To create a group, you will need to click on the down arrow that follows the word "None" that you can see in the text box. This will open up a drop down box that allows you to create a new group, or to add genome to an existing group. To create a new group, click on the line that says "Create New Group" (Arrow A in the screenshot below). Then you must type in the name that you want to identify the group in the text box (Arrow B). To save that group, click on the Save to Workshop button at the bottom of the pop-up box (Arrow C).



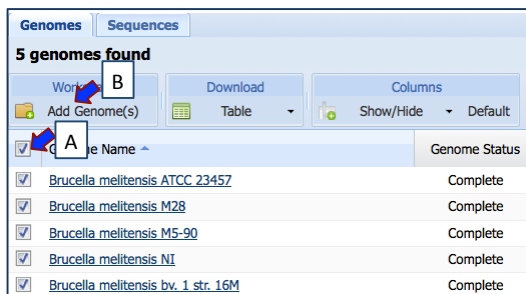
6. Once saved, you will return to the table view where you can choose more genomes. To select from the genomes that are freely available in PATRIC, click on the box in front of the word “Public” (Blue Arrow A, below). This will resort the table to show all the genomes that meet that criterion under the taxon level you have chosen (in this case, *B. melitensis*).



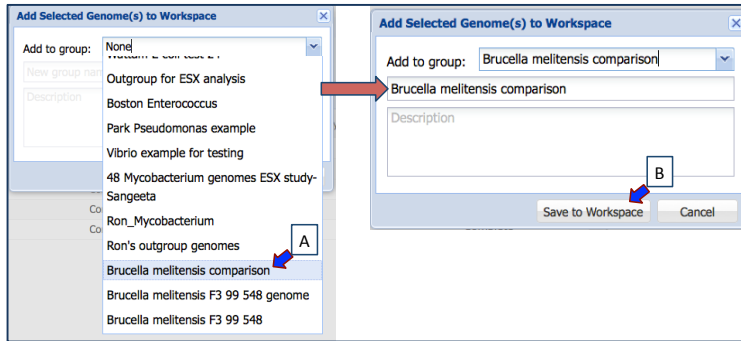
7. Let’s say you were interested in complete genomes. You could find them by clicking on the box in front of the word “Complete” (Blue Arrow, below) and this would resort the table to show all the complete genomes.



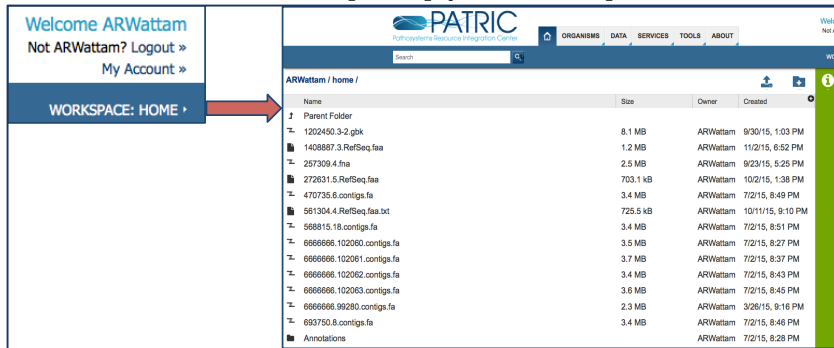
8. Now we are going to combine these genomes with the group we previously created. To do this, click on the box in front of “Genome Name” (Arrow A below). This will select all the genomes in the table. To save those genomes to a group, you need to click on the Add Genomes button at the top of the table (Arrow B).



9. To add these genomes to a group you have previously created, you will once again have to click on the down arrow that follows the word “None” that you can see in the text box. This time, scroll down the lists of names until you find the group you created. If you have just a few genome groups, this will be easy. If you have a lot like I do, keep scrolling! When you find the name, click on it (Blue Arrow A below) and this will auto-fill the text box with the name of the genome group. To add the new genomes to this group, click on the Save to Workshop button at the bottom of the pop-up box (Arrow B).



10. You can check your group to make sure that it has all the genomes you wanted. To do this you will need to find the Workspace icon at the top of any PATRIC page. Click on that, and it will open up your workspace.



11. Scroll down to find Genome Groups. Once you have located it, double click on the icon in front of it (Blue Arrow A below) and this will open up a page with all your genome groups.

Name	Size	Owner	Created
Parent Folder			
2 Ochro 40 Brucella	45 genomes	ARWattam	6/5/15, 8:42 PM
2011 E. coli genomes	313 genomes	ARWattam	9/14/15, 4:09 PM
300 Pseudomonas genomes	297 genomes	ARWattam	9/9/15, 6:05 PM
40 genomes Brucella	40 genomes	ARWattam	1/12/12, 2:56 PM
42 complete genomes-Mycobacterium	20 genomes	ARWattam	6/1/12, 2:13 PM
48 Mycobacterium genomes ESX study-Sangeeta	48 genomes	ARWattam	10/9/15, 6:47 PM
Acinetobacter AYE and SDF	2 genomes	ARWattam	5/19/15, 6:51 PM
Boston Enterococcus	12 genomes	ARWattam	9/22/15, 2:59 PM
Boyle Polyamine	60 genomes	ARWattam	3/26/15, 4:02 PM
Brucella	40 genomes	ARWattam	7/30/15, 2:12 PM
Brucella melitensis F3 99 548	4 genomes	ARWattam	11/23/15, 2:41 PM
Brucella melitensis F3 99 548 genome	1 genomes	ARWattam	11/23/15, 9:16 PM
Brucella melitensis F3_99_548 comparison	4 genomes	ARWattam	11/18/15, 10:04 PM
Brucella melitensis comparison	6 genomes	ARWattam	11/24/15, 1:40 PM

12. At the top right of the Genome Groups page is a column head named Created. To easily find the group you just named, you can double click on that word. The first click sorts the table from the oldest group to the most recent, and the second click reverses this to put the most recent group at the top. Incidentally, all the heads of columns are sortable and clickable.

Name	Size	Owner	Created
Parent Folder			
40 genomes Brucella	40 genomes	ARWattam	1/12/12, 2:56 PM
42 complete genomes-Mycobacterium	20 genomes	ARWattam	6/1/12, 2:13 PM
M. haemophilum comparison	16 genomes	ARWattam	8/31/12, 11:39 AM

Name	Size	Owner	Created
Brucella melitensis comparison	6 genomes	ARWattam	11/24/15, 1:40 PM
Brucella melitensis F3 99 548 genome	1 genomes	ARWattam	11/23/15, 9:16 PM
Brucella melitensis F3 99 548	4 genomes	ARWattam	11/23/15, 2:41 PM
Brucella melitensis F3_99_548 comparison	4 genomes	ARWattam	11/18/15, 10:04 PM
Ron's outgroup genomes	2 genomes	ARWattam	11/6/15, 6:03 PM
Ron_Mycobacterium	100 genomes	ARWattam	11/5/15, 4:33 PM
canis group	9 genomes	ARWattam	11/2/15, 6:23 PM
Dark Pseudomonas example	7 genomes	ARWattam	10/28/15, 10:58 AM

13. The group you have just created should appear in the top row. To see what genomes are included in that group, double click on the icon that is in front of the name you gave that group (Blue Arrow A below). This will open a page that shows you all the genomes in that group.

Name	Size	Owner
Brucella melitensis comparison	6 genomes	ARWattam

Genome Name	Genome Status	PATRIC CDS	Isolation Country	Host Name	Disease	Collection Date	Completion Date
Brucella melitensis NI	Complete	3321	China	Cattle, Bos taurus			10/5/11
Brucella melitensis bv. 1 str. 16M	Complete	3499		Human, Homo sapiens	Brucellosis		1/4/02
Brucella melitensis ATCC 23457	Complete	3486		Human, Homo sapiens	Brucellosis		3/24/09
Brucella melitensis MS-90	Complete	3347			Brucellosis		3/21/11
Brucella melitensis M28	Complete	3366					3/17/11
Brucella melitensis_F3_99_548_contigs	WGS	3365					