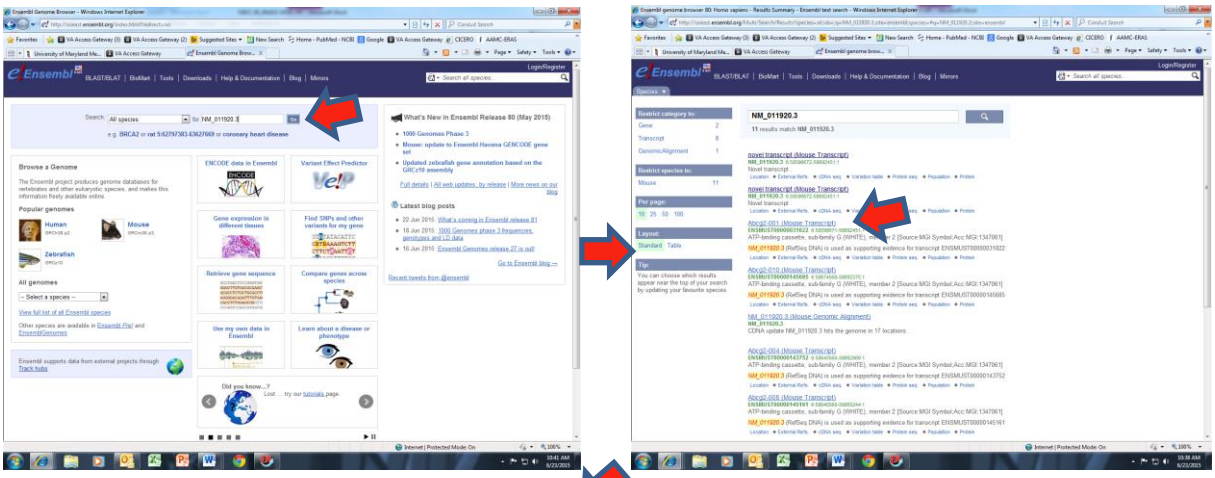
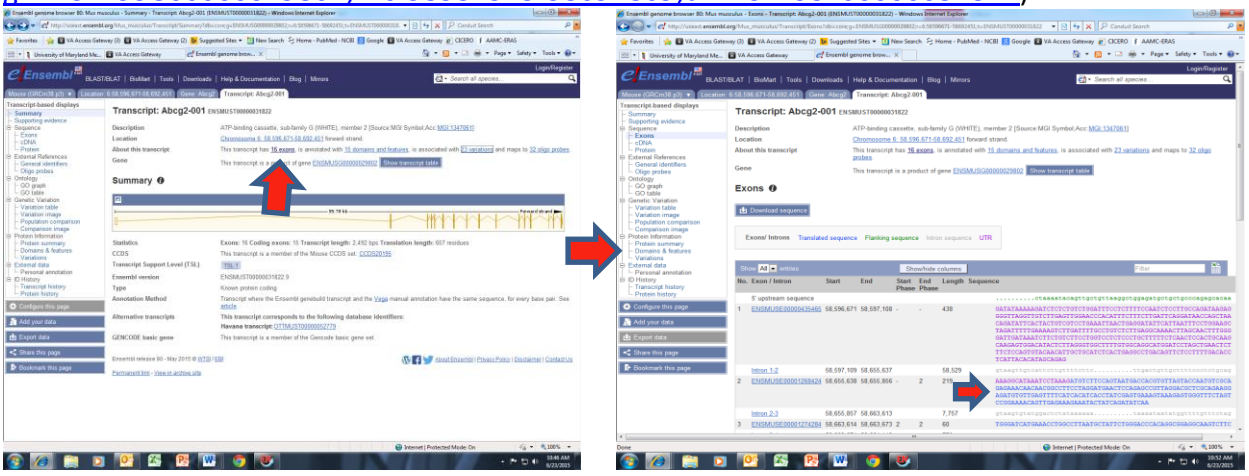
1. *In-silico* prediction of alternative first exons of Bcrp1 using BLAST analysis of EST database

Overview. This protocol describes how to search the mouse expressed sequence tag (EST) database for ESTs with sequence similarity to exon 2 of Bcrp1 and then how to align the matching EST sequences to genomic sequences to ascertain their location in the mouse Bcrp1 gene relative to the 5' end of Bcrp1 exon 2.

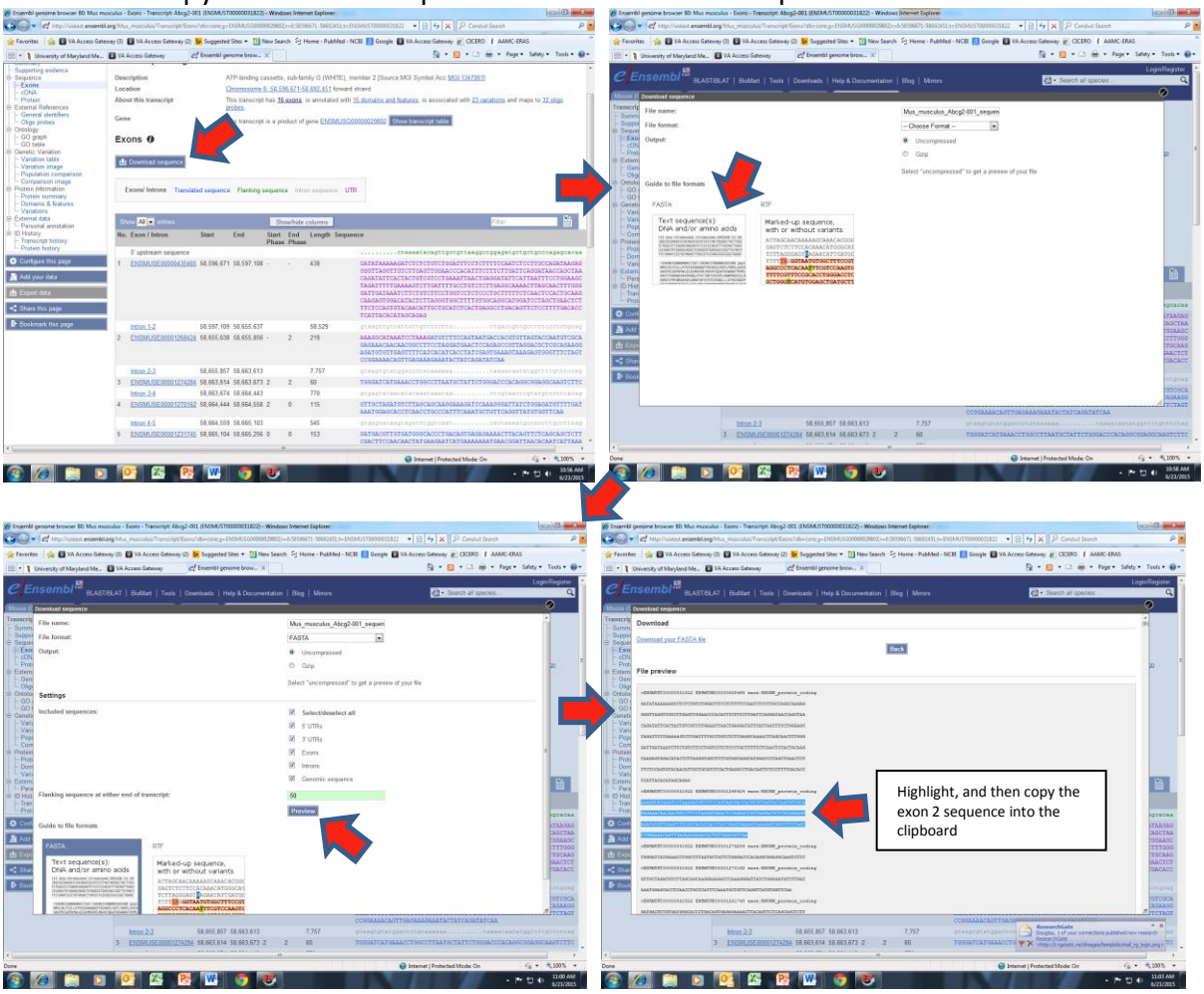
1.1 Obtain the sequence for mouse Bcrp1 exon 2 by inputting the mRNA reference sequence ID (NM_011920.3) into the search window of Ensembl (http://Ensembl.org), click on "GO," then select a full-length sequence (contains 16 exons).



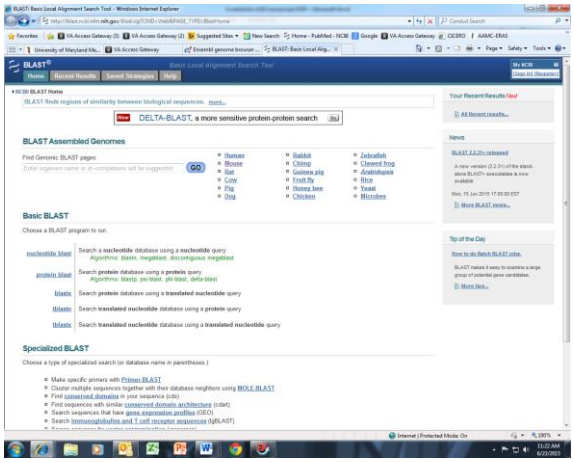
1.1.1 In the "Transcript-based Display" window, select "16 Exons." Exon 2 of Bcrp1 should read "5'-AAAGGC…TATCAA-3'." The results obtained should be similar to those shown at the following url: http://www.ensembl.org/Mus_musculus/Transcript/Exons?db=core;g=ENSMUSG00000029802;r=6:58584523-58692869;t=ENSMUST00000031822).



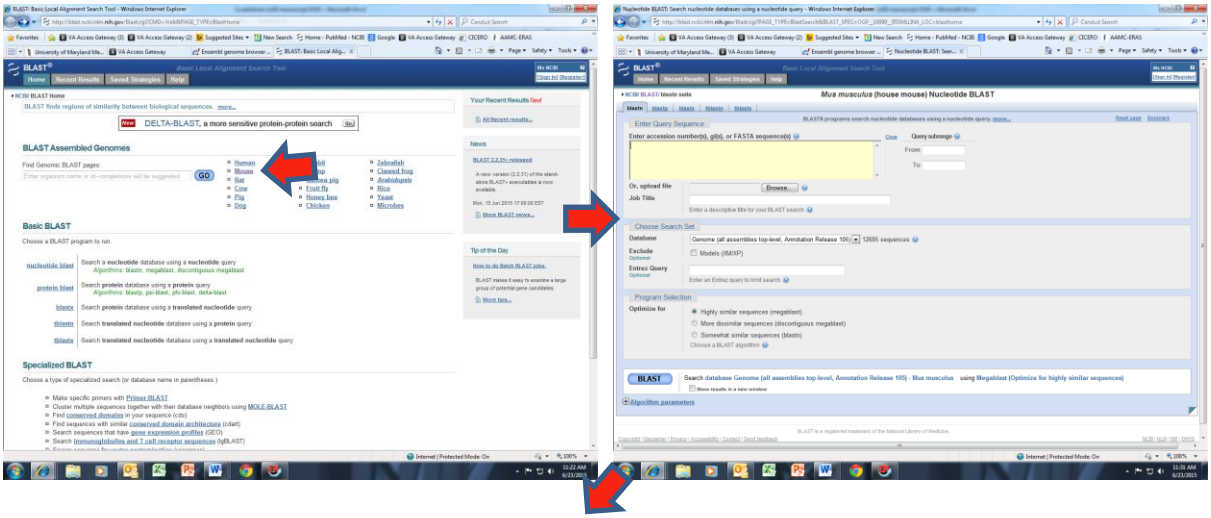
1.1.2 Using the “Download sequence” option, choose the FASTA format, then click on “Preview.” Copy the Preview sequence of exon 2 into the clipboard.



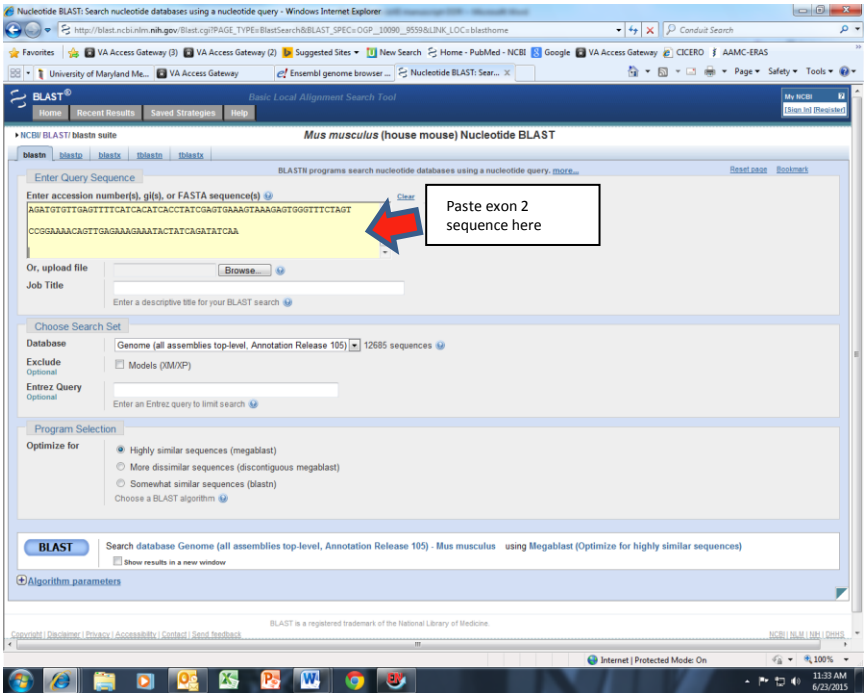
1.2. Navigate to the BLAST homepage (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome) on the NCBI website.



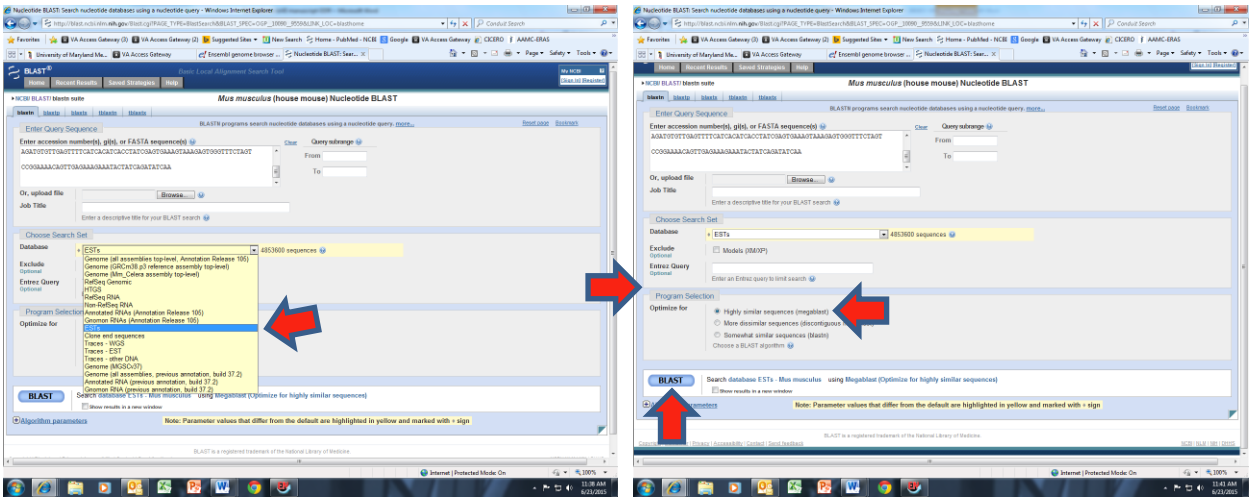
1.2.1. Select “Mouse” genome.



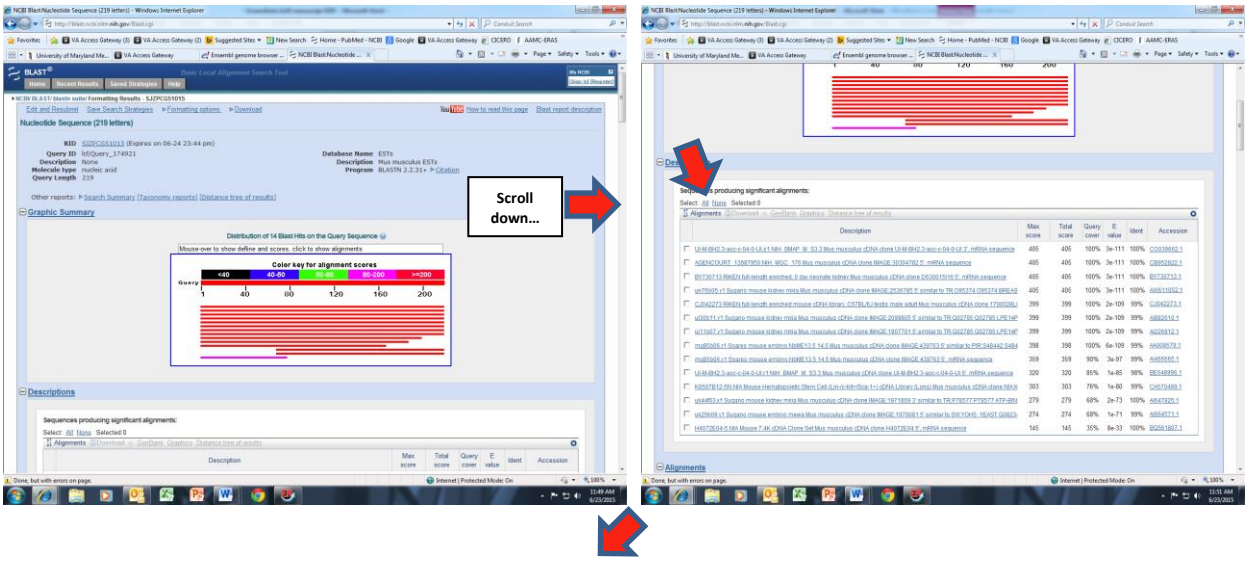
1.2.2. Paste the exon 2 sequence from the clipboard into the query box.

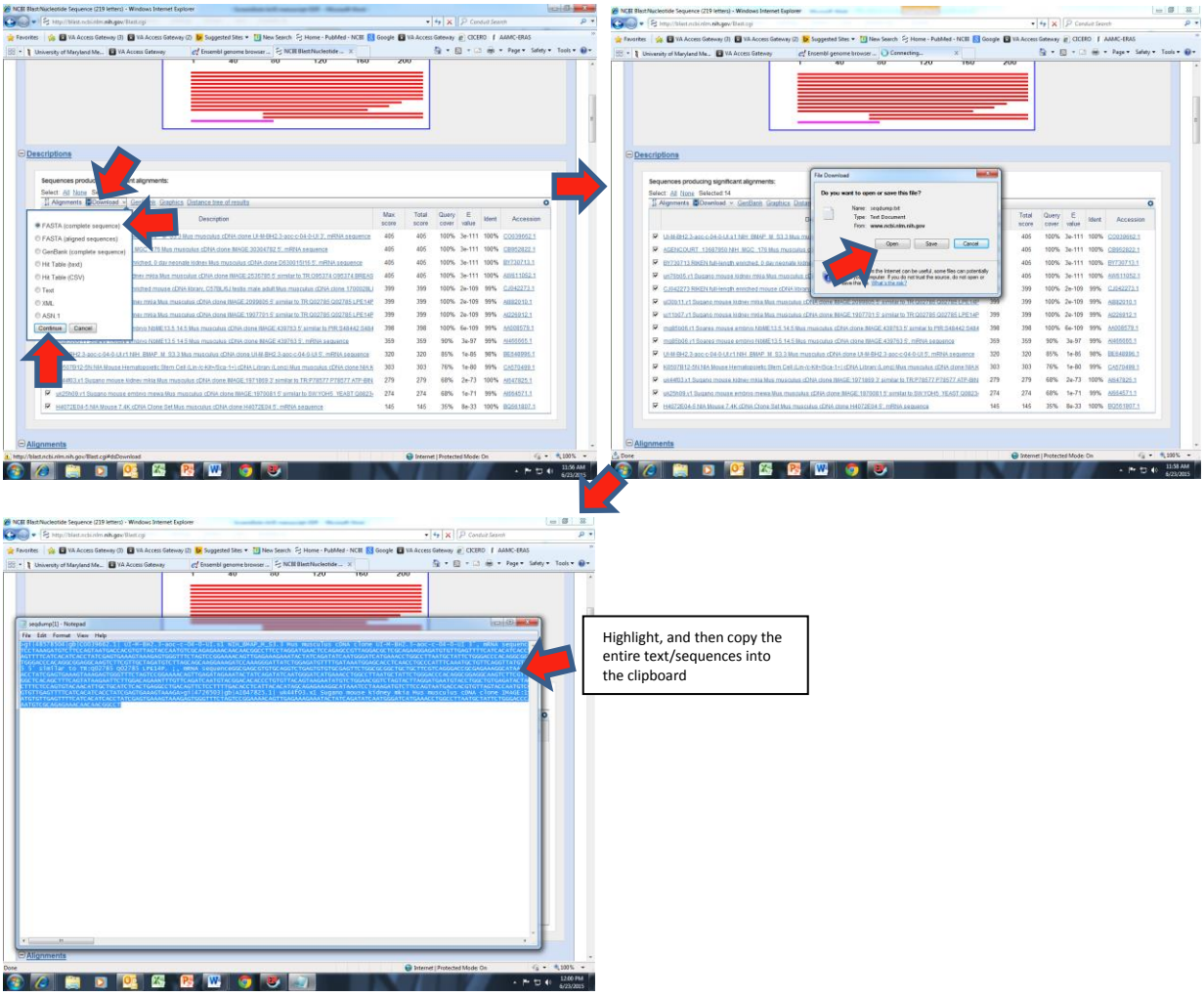


1.2.3. Select “ESTs” from the database dropdown, optimize for “highly similar sequences,” then select “BLAST.” Run time will take a few minutes.



1.2.4. When the BLAST run is complete, the “Results” page will appear. To analyze the EST sequence 5’ to exon 2, under the “Descriptions” subheading of the “Results” page, select “ALL,” then “Download,” select “FASTA (complete sequence)” in the dropdown, then select “Continue.” A .txt file should appear; open and then copy the entire file into the clipboard. The .txt file contains the sequences of all ESTs with high sequence similarity with the mouse Bcrp1 exon2 but not their position in relation to exon2 in the Bcrp1 gene. An analysis performed on April 15, 2015 identified 14 murine ESTs that aligned with Bcrp1 exon 2. These are listed in Table 1.



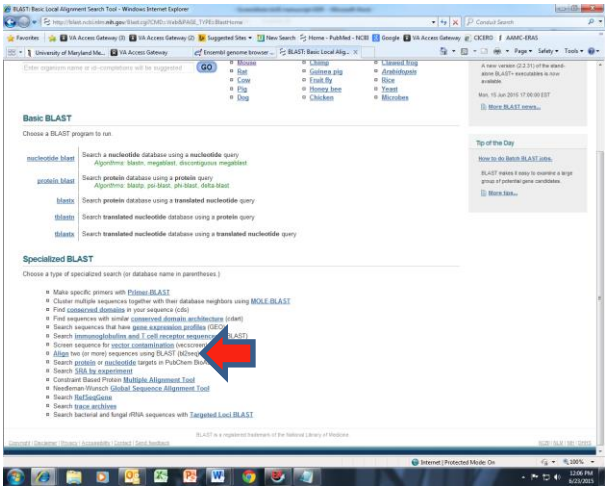


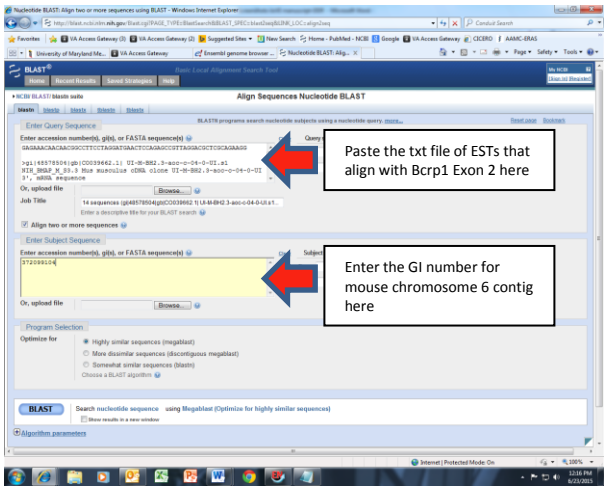
Highlight, and then copy the entire text/sequences into the clipboard

1.3. Identify the location of the EST sequence that is 5’ to exon 2 in the Bcrp1 gene. The mouse Bcrp1 gene is located in the chromosome 6 contig NC_000072.6 (GI:372099104).

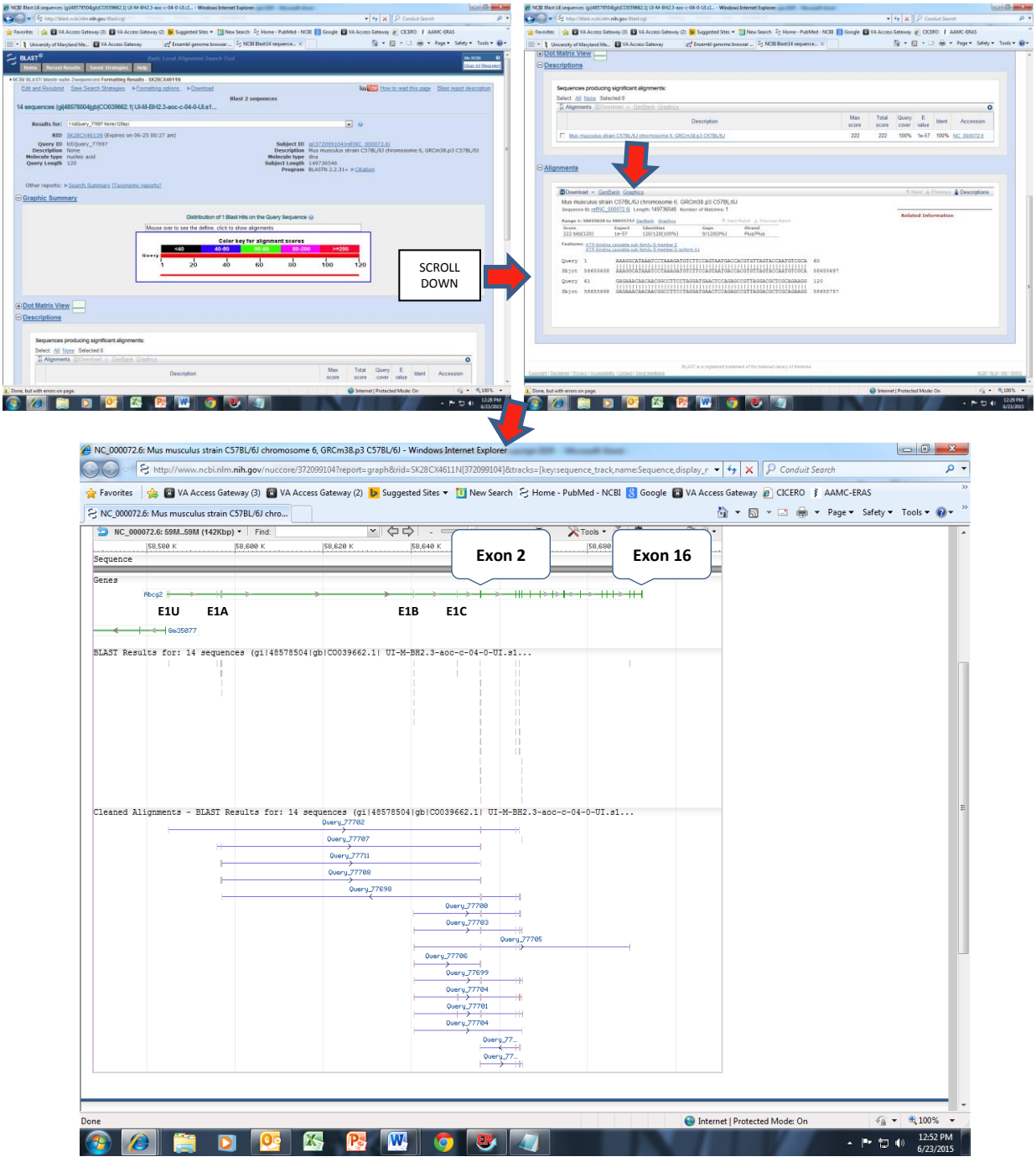
1.3.1. On the BLAST homepage under the “Specialized Blast” subheading, select “Align two (or more) sequences using BLAST (bl2seq).”

1.3.2. Paste the text file from the clipboard into the query box and enter 372099104 in the subject sequence box. Optimize for “highly similar sequences” under program selection, and run BLAST.

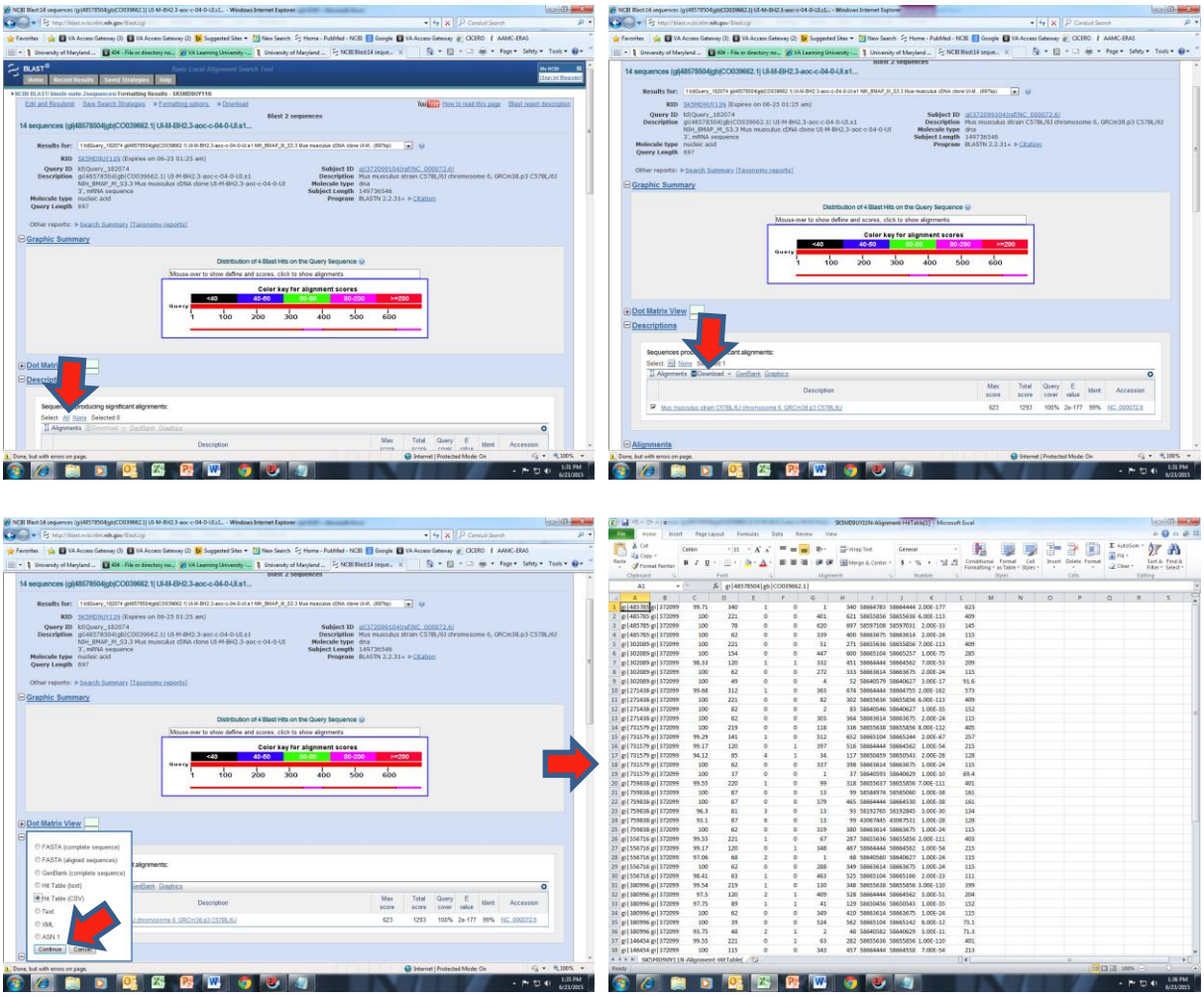




1.3.3. Once the results window appears, view the alignments graphically by clicking on “Graphics” in the “Alignments” window. Use the right and left arrows and zoom to focus on Bcrp1/Abcg2 and the sequence alignments.



1.3.4. Save the sequence alignments: select “ALL” in the “Descriptions” box, then under the “download” dropdown select “Hit table (CSV),” then click on “continue.” This file contains the sequence alignment of Bcrp1 exon 2 and the alignment of all the EST sequences with sequence similarity to Bcrp1 Exon2 relative to the numbering of the nucleotides in the mouse chromosome 6 contig. Each complete EST sequence might generate multiple alignments spanning regions 5’ and 3’ to exon 2 including the sequences overlapping with Bcrp1 exon 2.



1.3.5. The position of the 5’ end of exon 2 will correspond to nucleotide 58,655,638 in the chromosome 6 contig. Designate this as +1, and then calculate the position of the partial sequences of each EST 5’ to 58,655,638. The results for the 14 ESTs are given in Table 1.

Note: Be careful to analyze EST sequences that are 5’ to to +1 (i.e., have a negative nucleotide value) as potential first exons. For example, in two of the ESTs that aligned with Bcrp1 exon 2 shown in Table 1 (AI647825 and AI664571) the remaining sequence was 3’ to exon 2.