Worked example: Viewing the structure of the amyloid fibrils

1. Go to EBI at http://www.ebi.ac.uk/
2. Enter app in the EB-eye search box with All Databases selected in the drop down menu and click the Go button
3. Expand the Genomes databases, and select the Ensembl link
4. Scan down the retrieved records and select the gene record for human APP
5. Scroll down to the first transcript and to the Similarity Matches section
6. Identify the PDB (structure) record ids
7. Click on the linked PDB id 2BEG
8. Confirm you have retrieved the structure file for 2BEG
9. (Note that we could have gone directly to PDB at http://www.rcsb.org/pdb and searched for amyloid fibril to retrieve this record)
10. We will visualize the structure using a couple of the Java applet viewers: Jmol and QuickPDB.
11. Underneath the image on the right, there are several display options.
12. Click on the linked Jmol first and wait for the Java applet to start up
13. You are viewing the NMR structure of five molecules of beta amyloid arranged in a sheet
14. Note that each 44 amino acid molecule forms a U-shape and that the molecules stack upon one another in a non-twisting fashion
15. Click on the structure and move the mouse to rotate the structure to get a better view
16. The arrow cartoon indicates a beta sheet—along with alpha helices, a common secondary structure in proteins
17. Click the back button on the browser
18. Click on the QuickPDB link under the image on the right
19. Here we are viewing just the peptide backbone, and no cartoon structures have been drawn
20. Rotate the structure to a better view by clicking and dragging
21. Note when we mouse over one of the peptide backbones the residue number appears
22. Click on several different residues and note the corresponding residues in the sequence viewer panel are highlighted
23. In the lower left control panel, check one of the chain Ids and then select hydrophobic in the Residue Properties dropdown menu
24. We see that beta amyloid peptide is composed primarily of hydrophobic (Greek for "water fearing") residues, which would could explain why it self-assembles into fibrils in the water-rich extra-cellular environment.

Worked example: Extracting and aligning human and mouse APP upstream regions

1. First we will extract DNA sequence upstream of human and mouse APP genes
2. Go to NCBI at http://www.ncbi.nlm.nih.gov/
3. Query for APP[Gene Name] AND (human[Organism] OR mouse[Organism])
4. Click on the Gene results
5. Click on the human APP gene record
6. Click on the gene's chromosome sequence, and select FASTA format
7. Note that we are viewing the sequence on the minus strand of DNA
8. We will extract sequence 10kb upstream of the transcription start site (TSS)
9. Copy the coordinate in the "Range to" text box, which is our working TSS
10. Paste it into the "Range from" text box
11. Add 10000 to the number in the "Range to" box (remember we on the minus strand)
12. Click the Refresh button
13. Select File in the "Send to" dropdown menu
14. Confirm that there is a file called "sequences.fasta" on the desktop
15. Click the browser's go back button until you are at the Gene database results
16. Click on the mouse APP gene record, and repeat steps 5-12.
17. Confirm that there is a second file on the desktop called "sequences(2).fasta"
18. Now we will confirm that we extracted the correct sequences
19. Open a new browser window and go to UCSC Genome Browser at http://genome.ucsc.edu/
20. Click on the BLAT link on the left sidebar
21. Make sure the genome dropdown menu is set to human
22. Click the Browse… button below the large text area and select the human "sequences.fasta" file on the desktop
23. Click the submit button
24. Click on the browser link on the first hit in the list of results
25. Your query sequence is at the top of the genome browser viewer as the black box and arrows indicate direction of your sequence
26. Click on the zoom out 3X button and verify that your sequence is 10kb in length and directly upstream of the APP TSS
27. Click the browser's back button until you are at the BLAT Search Genome page
28. Select the mouse genome and browse for the mouse "sequences(2).fasta", and then repeat steps 23-26
29. Now we will do an alignment using some comparative genomics tools
30. Go to DCODE.org at http://www.dcode.org/
31. Click on the zPicture link on the left sidebar
32. We will enter the human DNA as sequence 1 and mouse as sequence 2
33. In the Sequence 1 box, click on the Browse… button, go to the Desktop and select "sequences.fasta"
34. In the Sequence 2 box, do the same but select "sequences(2).fasta"
35. Click the SUBMIT button at the bottom of the page, and you'll get a queue message
36. Click on the "click here" to refresh link, and you'll eventually see a results page
37. Click on the Dynamic visualization icon

38. The graphic on this page shows a comparison between our human and mouse DNA sequences as percent identity and red shaded segments are segments showing high identity.
39. Click on one of the red-shaded sections and you'll retrieve a regional alignment and the sequence segments.
40. Click on the browser's go back button
41. Click on the Dot-plot icon
42. A dot-plot is the traditional way to view pairwise sequence comparisons, particularly whole genome-genome, since it is excellent for viewing duplications and inversions. One sequence is on the horizontal axis and the other on the vertical axis. A dot indicates a high degree of sequence identity.
43. Click on the browser's go back button
44. Click on the rVista icon
45. Make sure that TRANSFAC professional library is selected and click the SUBMIT button
46. Since we are interested in the MEF2 transcription factor, scroll down to the M section and check the box next to MEF2 and click the SUBMIT button
47. Click on the CHECK button
48. Click on the Dynamic Visualization icon
49. Notice that this is a similar display to what we have already seen, except there is an extra track above the image
50. Check smooth plot in the Picture section
51. Check conserved and all in the Show section
52. Check flip above the SUBMIT button so that we see the sequence as we extracted it from NCBI (zPicture flipped our sequence to the positive strand for some reason)
53. Click SUBMIT
54. The ALL track shows all MEF2 sites detected in the human sequence, the CONSERVED track shows MEF2 sites that were found in both sequences in a conserved region
55. Click on the browser's go back button
56. Click on the linked Highlight in the Alignment section
57. Click the SUBMIT button
58. Scroll down the alignment of the conserved segments and note the single conserved MEF2 site in blue.