

MCB 5472 Assignment #7: Multiple sequence alignment

March 5, 2014

This week's assignment will be a brief comparison between some of the multiple sequence alignments that we discussed in the lecture on Monday. Remember that last week's assignment will be due alongside this week's work before the midterm exam next Wednesday March 12, 2014.

Question: (10 marks) How do multiple sequence alignments created using MUSCLE, MAFFT and PRANK compare to each other and to a reference alignment derived from 3D structural alignments? You can find one of the BALiBASE3 sequence datasets on the website (BALiBASE3.faa) to use as your input, along with their alignment based on structural features (BALiBASE3.msf). The syntax to run each program is quite simple:

MUSCLE:

```
>muscle -in [input_file] -out [output_file]
```

MAFFT:

```
>mafft --auto [input_file] > [output_file]
```

PRANK:

```
>prank -d=[input_file] -o=[output_file]
```

(Note: PRANK produces 2 similar output files for reasons I cannot find in its manual. Either will work for your purposes.)

To look at these alignments you will need to download a program designed for alignment viewing. Some options include

MEGA6 (Windows, Mac): <http://www.megasoftware.net/>

Jalview (all platforms): http://www.jalview.org/Web_Installers/install.htm

Seaview (all platforms): <http://doua.prabi.fr/software/seaview>

Before next week: Email Jonathan (jonathan.klassen@uconn) a short paragraph comparing the three alignment programs with each other and the reference alignment. As discussed in class, there are many ways to make such comparisons, including both the alignments themselves and computational efficiency. Also remember that similarities can be as interesting as differences.