Staden Course Exercises

version 3.2 (public)

Licence

This manual is © 2007-8, Simon Andrews.

Exercise 1: Running Pregap

- ♦ You should have a directory containing a batch of example sequencing files

- ♦ Start Pregap by opening the Software repository and double clicking on the Pregap4 icon

- ♦ Use the Add Files button, and the file selector to add in all the SCF files in the Staden directory

- ♦ Make the following changes to the default module configuration
  - ♦ In Quality Clip increase the "Average Confidence" cutoff to 20

- ♦ Switch to the Textual Output window and run Pregap

- ♦ Look at the log which is generated to see if everything worked OK

Exercise 2: Importing Sequences into Gap

- ♦ Start Gap by selecting double clicking the Gap4 icon in the Molbio Repository

- ♦ Create a new database called "assembly" in the Staden directory

- ♦ Use the "Normal Shotgun Assembly" to add in all the files listed in the pregap.passed file into this new database. Use the default assembly parameters

- ♦ Check how many contigs are created by looking at the Contig Selector

- ♦ Visualise how the traces are distributed across the contigs by generating a template display for all contigs

Exercise 3: Using the Contig Editor

- ♦ Open the longest contig in the contig editor

- ♦ Adjust the settings so that any differences are highlighted with a different background colour. Do not make difference highlighting case sensitive

- ♦ Save your view options so they will be restored in future

- ♦ Scroll through the contig and identify where there are problems. Look at the underlying traces in the problem regions to see how they might be resolved

Exercise 4: Correcting Mistakes

♦ Use the editing techniques you are shown to correct any mistakes in the assembly. Always use the lowest edit mode you can

♦ Save the changes you have made to your contig

Exercise 5: Joining Conitgs

♦ Use the Find Internal Joins tool to look for potential overlaps between your contigs

♦ Use the join editor to join contigs together where the overlaps are solid. Edit the individual contigs as necessary to get the joins to work

♦ You should be able to get your assembly down to two contigs. You may find a small overlap between them but do not use this to make a final join.

Exercise 6: Adding More Data

♦ Add the "rawseq.txt" file as a single file to your existing assembly

♦ Use the find internal joins tool and the join editor to assemble your sequences to a single contig.

♦ Correct the contig so you are happy with the consensus it provides and save these changes.

Exercise 7: Exporting Data

♦ Export the consensus sequence for your contig into a FastA format file

♦ Dump the whole alignment of your assembly project into a text file and import it into a Word document.