# Table of Contents

1. Contacts .................................................................................................................3

2. Overview .................................................................................................................3

3. Unit objectives .........................................................................................................4

4. Method of Delivery & Communication .................................................................4

5. Assessment and semester activity ...........................................................................5
   5.1 Assignments ........................................................................................................6
   5.2 Examination .......................................................................................................6
   5.3 Exercises ............................................................................................................6
   5.4 Online participation ..........................................................................................6

6. Timetable ..................................................................................................................7

7. Textbooks .................................................................................................................7

8. Software ...................................................................................................................8

9. Immediately useful websites ..................................................................................8

10. Complaints policy .................................................................................................8
1. Contacts

Nicola Armstrong  Natalie Thorne

Mathematics and Statistics  Melbourne Genomics Health Alliance
Murdoch University  Walter and Eliza Hall Institute

Ph:  (08) 9360 2480  Ph:  (03) 9345 2747 (MTWTh)
Fax: (08) 9360 6346  Fax: (03) 9347 0852
Email: n.armstrong@murdoch.edu.au  Email: thorne@wehi.edu.au

This semester, Nicola is the unit co-ordinator. General questions about the unit content and organisation should be addressed to Nicola.

The BCA Executive Officer is:
Erica Jobling
Email: bca@ctc.usyd.edu.au
Phone: (02) 9562 5076
Questions about the BCA should be addressed to Erica.

2. Overview

Bioinformatics, the finding of patterns in very large sets of biological data, has grown very rapidly over the past few decades. In this unit the aim will be to provide an introduction to the main data sources, techniques and software that make up the area. The unit does not assume knowledge of biology, so begins by providing an introduction to the central biological ideas. It progresses in a natural way, stepping from DNA to evolution and phylogenetics and thence to RNA and microarrays, exploring key ideas at each stage; all work is directed at an understanding of biological systems. En route, web-based resources will be explored and used to provide data sets, and the statistical package R employed to perform analyses.
3. Unit Objectives

At the end of this unit you should be able to:

1. Explain the core dogma of molecular biology and the central ideas of population genetics
2. Access appropriate web based sources for data, and download the data in suitable format, when given a problem that requires genome or proteome data for its solution.
3. Understand and apply core bioinformatics techniques for the analysis of DNA and protein sequence data, such as global sequence alignment, BLAST, Hidden Markov Models, evolutionary models and phylogenetic tree fitting.
4. Process large quantities of data (such as the expression profiles of thousands of genes resulting from microarray experiments) using R, and communicate results in language suitable for presentation to both a bioinformatics journal and a lay audience.

The unit is divided into eight modules, with either one or two weeks devoted to each. These are listed below.

<table>
<thead>
<tr>
<th>Module</th>
<th>Weeks</th>
<th>Content</th>
<th>Lecturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Biology basics</td>
<td>Armstrong</td>
</tr>
<tr>
<td>2</td>
<td>2,3</td>
<td>Population genetics</td>
<td>Armstrong</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Web-based data sources, data retrieval and tools</td>
<td>Thorne</td>
</tr>
<tr>
<td>4</td>
<td>5,6</td>
<td>The analysis of single and multiple DNA or protein sequences</td>
<td>Thorne</td>
</tr>
<tr>
<td>5</td>
<td>7,8</td>
<td>Hidden Markov Models and their applications</td>
<td>Armstrong</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>Evolutionary models</td>
<td>Armstrong</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>Phylogenetic trees</td>
<td>Thorne</td>
</tr>
<tr>
<td>8</td>
<td>11,12</td>
<td>Analysis of microarrays</td>
<td>Thorne</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>Introduction to RNASEq</td>
<td>Thorne</td>
</tr>
</tbody>
</table>

4. Method of Delivery and Communication

Two instructors (Nicola Armstrong and Natalie Thorne) are jointly responsible for the course. In previous years, the course has been run by Graham Wood and Steven Brown from Macquarie University and they developed most of the material for this unit. This semester, Nicola will take primary responsibility for the course.

Questions about administrative aspects or course content can be emailed to Nicola, and when doing so please use “BIF:” in the subject line of your email. Both Nicola and Natalie will be available by email and phone to answer questions related to the module notes and exercises, and to address other issues requiring clarification. However, please note that instructors are not necessarily available every day of the week and you should expect a delay in response! In most instances, email will be the most effective form of communication for an initial response.

The unit is offered in distance mode. Our core means of communication will be via printed notes (which will be available on Blackboard), e-mail and Blackboard discussion.
The unit relies to an extent on the prescribed text by Durbin, Eddy, Krogh and Mitchison (see below). The study notes provide a guide to readings in this text, and often to other readings, which will be provided. They also provide additional explanation where this is needed. In the study notes for each module, exercises or tasks are given, referring to exercises in the textbook or provided readings.

The study notes for the first module will be mailed to you. These will also be posted on the Blackboard site. You will be alerted by email when this material has been mailed out, so if you do not receive these within a few days then please contact Nicola Armstrong (contact details given earlier).

**Blackboard - our web forum**

We will be using Blackboard for online discussions, posting of course notes, solutions and data sets, and submission of assignments. The link is: [http://elearning.sydney.edu.au/](http://elearning.sydney.edu.au/) which you should bookmark.

As Blackboard is the primary medium for communication in this unit, we expect that students will access the site at least every second day during semester.

**Student “Profiles”**

In order to introduce ourselves to each other, we encourage you to create a profile on Blackboard. Just go to “Discussions” then “Introductions”, then “Create Thread” and away you go. You are encouraged to include at least the following:

- your preferred name;
- the degree program you are studying;
- your current job role, and the organisation for which you work;
- what you hope to gain from the unit;
- and a little about yourself!

You may include your e-mail address if you want to be contactable directly by other students. It would be good if you could do this by the end of the first week of semester.

**Feedback**

Your constructive comments are always valued and guide us in the continuous improvement of the unit.

**5. Assessment and semester activity**

Assessment for the unit consists of: three assignments and a written examination. The weighting is as follows:

<table>
<thead>
<tr>
<th>Assignment Type</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three assignments (20% each)</td>
<td>60%</td>
</tr>
<tr>
<td>Examination</td>
<td>40%</td>
</tr>
</tbody>
</table>
5.1 Assignments

Rules for assignment submission are given in the document “BCA Assessment Guide”, which is included with these notes and is also on the BCA Blackboard site. Please read this document carefully. All students are required to sign an “own work” declaration before submission of assignments is allowed via Blackboard. Please see the “Assessment Guide” for more information.

Late submission: Requests for an extension of the due date for an assignment must be made in advance of the due date for that assignment. These requests must be made directly to the unit coordinator by email. The unit coordinator will reply by email with the decision as to whether an extension has been granted and the new due date.

Late penalty: If no extension has been given, 5% of the earned mark for an assignment will be deducted for each day that an assignment is late, up to a maximum of 50%.

NOTE: It is not the intention of this late penalty policy to cause a student to fail the unit when otherwise they would have passed. If deductions for late assignments result in the final unit mark for a student being less than 50, when otherwise it would have been 50 or greater, the student's final mark will be exactly 50.

5.2 Examination

The examination will be released on Blackboard at 5pm on Friday 6 November, and is due at 9am on Monday 9 November. You should allocate about two free days to complete the examination. You should consider yourself under examination conditions, i.e. there is to be no communication between yourself and other students or any other person, in connection with the examination.

5.3 Exercises

There are exercises in the notes. Exercises should be attempted. Online discussions around each module's exercises can be conducted on Blackboard. You are encouraged to make discussion contributions, on any of the questions, including (but we hope not limited to) those designated for submission.

5.4 Online participation

You are strongly encouraged to participate in online discussions. We are a dispersed group and this forum can bring a little unity!

Instructors will generally let discussion flow between the students, except where key points need resolution. The instructors will not usually make comments on the exercises required for submission, at least until after the deadline.
Any general Discussion items, in particular on the module notes, can be posted to the other Discussion areas, e.g. R, Module 1, Module 2, etc. If you want the attention of the lecturers you are very strongly advised to be in contact directly, using email. This is monitored more regularly than the discussion area on Blackboard.

6. Timetable

<table>
<thead>
<tr>
<th>Week</th>
<th>Starting date</th>
<th>Module</th>
<th>Assessment due</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27 July</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3 August</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>10 August</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>17 August</td>
<td>3</td>
<td>Assignment 1</td>
</tr>
<tr>
<td>5</td>
<td>24 August</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>31 August</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>7 September</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>14 September</td>
<td>5</td>
<td>Assignment 2</td>
</tr>
<tr>
<td>9</td>
<td>21 September</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mid semester break</td>
</tr>
<tr>
<td>10</td>
<td>5 October</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12 October</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>19 October</td>
<td>8</td>
<td>Assignment 3</td>
</tr>
<tr>
<td>13</td>
<td>26 October</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Assignments are due on the Monday of the indicated weeks. Assignments will be available on Blackboard at least two weeks before the due date. The examination (see above) will be held on the weekend of 7/8 November.

7. Textbooks

The prescribed text is:


Copies are available from the Co-op bookshop (can be ordered online). You will start to need the text in Module 4, so it is a good idea to order it early.

There are numerous texts on bioinformatics. The following may be useful:

2. Warren J. Ewens and Gregory R. Grant (2005) *Statistical methods in bioinformatics - an introduction*, Springer-Verlag. (Excerpts from this texts will be provided.)
4. Anna Tramontano (2007) *Introduction to bioinformatics*, Chapman and Hall. (Good on data sources and sequence analysis; strong on protein analysis, not in this course.)

8. **Software**

We will be using the statistical package R. You can download and install a version of this reliable freeware from the internet at [http://www.r-project.org/](http://www.r-project.org/).

9. **Immediately useful websites**

- BCA homepage: [http://www.bca.edu.au](http://www.bca.edu.au)
- R homepage: [http://www.r-project.org/](http://www.r-project.org/)

10. **Complaints policy**

Please see the BCA complaints policy in the Assessment Guide