

# 2005/06 Taught Postgraduate Module Catalogue

## **BIOL5020M**

Bioinformatics of Protein Sequence, Structure and Function

**15 credits**

**Module manager** Dr D R Westhead

**Taught** Semester 1 [View Timetable](#)

**Year running** 2005/06

### **Pre-requisites**

BIOL5110M or equivalent knowledge of biochemistry/molecular biology.

**This module is not approved as an Elective**

### **Objectives**

On completion of the module the students should have attained the following objectives:

1. to be aware of the most important primary and secondary bioinformatics databases;
2. to understand the nature of the data, from high quality protein sequence and structure databases to ESTs;
3. to be aware of the most important WWW sites hosting databases and software;
4. to understand how bio-molecular sequences and structures evolve;
5. to understand the problem of sequence alignment;
6. to understand dynamic programming alignment algorithms;
7. to understand the derivation of common amino-acid substitution matrices;
8. to understand the similarity detection methods used in database searching;
9. to understand statistical issues associated with database searching;
10. to understand issues of specificity and sensitivity for database searches;
11. to understand multiple sequence alignment and associated algorithms;
12. to understand the concept of a protein domain family;
13. to understand consensus sequences and sequence profiles;
14. to be aware of the main secondary databases associated with protein families;
15. to understand iterative database searches and PSI-BLAST;
16. to understand the principles and problems of phylogenetic tree construction;
17. to be aware of databases associated with bio-molecular structures;
18. to be able to use software for display of such structures;
19. to be aware of tools for structure analysis and verification;
20. to understand algorithms for structural alignment;
21. to understand protein structure classifications;

22. to understand how structures reveal distant evolutionary relationships;
23. to understand the principles of protein secondary structure prediction;
24. to understand protein structure prediction by comparative modelling;
25. to understand some methods of protein fold recognition;
26. to be aware of ab initio protein folding algorithms;
27. to be aware of whole genome sequencing projects and associated databases;
28. to understand the use of bioinformatics in annotating newly sequenced genomes;
29. to understand the problem of gene prediction;
30. to be able to select appropriate methods for a bioinformatics investigation;
31. to be able to interpret the results correctly;
32. and, to be able to describe the investigation and results in good scientific English.

### **Syllabus**

Bioinformatics databases, software and WWW sites; molecular evolution; sequence alignment; dynamic programming and substitution matrices; sequence database searching, methods and statistics; multiple sequence alignment, protein domain families, profiles, patterns and hidden Markov models; iterative searching and PSI-BLAST; phylogenetic tree estimation; protein structure analysis; structural alignment and classification; protein structure prediction, secondary structure prediction, comparative modelling, fold recognition, ab initio; whole genome databases and analyses; gene prediction.

### **Teaching methods**

Lectures: 16 x 1 hour;  
Practical classes: 3 x 3 hours;  
Problem based learning: 2 x 3 hours.

### **Private study**

2.5 hours reading per lecture: 40 hours;  
Problem-based learning activities: 34 hours;  
Coursework: 45 hours.

### **Progress monitoring**

Progress will be monitored in the practical classes and problem-based learning sessions.

### **Methods of assessment**

1. Problem-based learning group work: 30%;
2. an essay in the format of a scientific journal paper (with length typical of the journal *Bioinformatics*) detailing the analysis of a biomolecular sequence or structure using the methods of the course: 70%

### **Reading list**

The [reading list](#) is available from the Library website