

Unit title: **Bioinformatics**

Unit code: **M/601/0312**

Level: **4**

Credit value: **15**

Aim

This unit enables learners to explore bioinformatics to develop a structured approach to biological data and a detailed understanding of the software tools needed for analysis.

Unit abstract

This unit will enable learners to appreciate that, over the past few decades, advances in genomic technology, coupled with major advances in the field of molecular biology, have led to an explosive growth in the biological information generated in the scientific community.

This expansion of genomic information has led, in turn, to the need for biological databases to store, organise and index the data, and for specialised tools to view and analyse the data.

In this unit learners will review key methods, data sources and the aims of bioinformatics, including its use in medical diagnosis and the treatment and prevention of genetic diseases.

Learners will gain an understanding of computational biology, data collection and the computer procedures required by biologists. They will gain a knowledge and understanding of how to manage biological databases and how to analyse data using appropriate software tools.

Learning outcomes

On successful completion of this unit a learner will:

- 1 Understand the aims, methods, data sources and applications of bioinformatics
- 2 Understand the processes of computational biology
- 3 Understand key features involved in building biological databases
- 4 Be able to carry out data analysis within the field of bioinformatics.

Unit content

1 Understand the aims, methods, data sources and applications of bioinformatics

Bioinformatics aims: aims eg finding out how living things work, modelling biological systems, making advances in understanding basic biological processes, making advances in diagnosis, treatment and prevention of genetic diseases

Bioinformatics applications: applications eg hypothesising about functions of newly discovered genes, predicting how the protein folds, modelling protein and metabolite working together to make the cell function, developing predictive methods to model function and phenotype of an organism based on its genome sequence

Methods: types eg representation, storage, organisation, manipulation, distribution and maintenance of data, design of data formats and databases, creation of tools to query databases, development of tools to ask complex questions about the data

Sources of biological data: technologies and research initiatives that are making increased amounts of data available eg nucleotide sequences, amino acid sequences, protein domains, protein structures

Impact: on biological sciences eg from a purely laboratory-based science to an information science, greater global perspective of 'informed' design of genomic experiments, future developments for biomedical scientists and medicine

2 Understand the processes of computational biology

Computational biology: computer procedures required by biologists; the process of analysing and interpreting genomic data; use of specialised tools to view and analyse data

Data collection: new approaches to data collection; quality standards for new data sets; human genome and other genome projects; parallel computing; micro array technology; using search engines to find biological information; the key role of the worldwide web

Data warehousing: data from many sources gathered together in a consistent and useful way

Data analysis: analysis of biological data; use of software techniques for finding patterns and regularities in data sets; extracting relevant information for a specific (biological) question

3 Understand key features involved in building biological databases

Design of biological databases: design eg data capture, file formats, typical records within files, ease of access to stored data, complex interfaces to access existing data and submit new or revised data, software for database building

Management of biological databases: management eg submitting new or revised data, management of various types of information, database management strategies, database management systems, flat file and relational databases

Records within a file: types eg contact name, input sequence, type of molecule, scientific name of the source organism, literature citations

Types of biological database: types eg sequence, structure and specialised databases, public biological databases

Database resources: resources eg familiarity with the Protein Data Bank and GenBank, depositing data into public databases

4 **Be able to carry out data analysis within the field of bioinformatics**

Mathematical and statistical modelling: applying mathematical and statistical methods to analyse biological data

Data analysis: methods eg spreadsheet programmes (such as Excel) to create simple plots, sequence and gene expression analysis, accuracy of predictions

Software tools: pattern detection eg using Practical Extraction and Reporting Language (PERL) for efficient pattern detection, automated data analysis, familiarity with an assortment of available software, writing own programmes

Present information: format eg communication, organisation of biological information within a multi-disciplinary field, type

Learning outcomes and assessment criteria

| Learning outcomes On successful completion of this unit a learner will: | Assessment criteria for pass The learner can: |
|---|--|
| LO1 Understand the aims, methods, data sources and applications of bioinformatics | 1.1 explain the aims and applications of bioinformatics 1.2 discuss methods employed in informatics 1.3 discuss the impact of bioinformatics in the biological sciences |
| LO2 Understand the processes of computational biology | 2.1 discuss computational processes for the collection and manipulation of data 2.2 discuss the need to extract only the information relevant to a specific biological question |
| LO3 Understand key features involved in building biological databases | 3.1 summarise the design and management of biological databases 3.2 identify a range of records within a file 3.3 discuss the need for biological databases to store, organise and index basic biological processes 3.4 discuss the nature of new data available and the types of database and resources that might be used |
| LO4 Be able to carry out data analysis within the field of bioinformatics | 4.1 apply mathematical and statistical methods to analyse biological data 4.2 use software tools for efficient pattern detection 4.3 present information clearly. |

Guidance

Links

This unit has particular links with the following units within this qualification:

- *Unit reference number J/601/0221: Project for Applied Science*
- *Unit reference number J/601/0297: Statistics for Experimental Design*

Essential requirements

Delivery

The development of appropriate computer skills, especially in the correct use of software, must be encouraged whenever possible. Where appropriate, analysis of experimentally derived data must be attempted.

A case-study approach could be used to unify a number of themes, for example searching for a nucleotide string contained in a block of sequence data.

Assessment

Evidence for this unit could be provided by reporting on experimental work, such as DNA sequences or a protein structure. Use of experimentally derived data in case studies and interpretation exercises, literature searches and reviews would also be appropriate.

Resources

Learners will need access to computer facilities, appropriate software and library resources. The use of computer teaching programmes and CD ROM simulations of experiments must be encouraged.

Employer engagement and vocational contexts

Learners will benefit from visiting speakers engaged in bioinformatics research or processes. This would enable learners to appreciate the application of bioinformatics within a particular context and the required equipment and personnel.