MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

COURSE WORK and RESEARCH PROJECT MASTERS

in

BIOINFORMATICS and
COMPUTATIONAL MOLECULAR BIOLOGY

2012

DEPARTMENTS

of

BIOCHEMISTRY, MICROBIOLOGY & BIOTECHNOLOGY

CHEMISTRY, COMPUTER SCIENCE, MATHEMATICS and STATISTICS

QUALITY ASSURANCE MANUAL
DEPARTMENT OF BIOCHEMISTRY, MICROBIOLOGY AND BIOTECHNOLOGY

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

TABLE of CONTENTS:

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

ORIENTATION PROGRAMME

(3 February 2012, Friday, Biochemistry Tea Room)

10:00-10:30	Tea and welcoming
11:00 – 11:30	Introduction of students, lecturers
11:30 – 12:00	Introduction to the programme
12:00 – 12:30	Overview of bioinformatics by Ozlem

Lunch Break

14:30 – 15:30 Research talks from potential supervisors (Maximum 15 min each)

- Kevin Lobb
- Philip Machanick
- Earl Prinsloo
- Ozlem Tastan Bishop

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

PROPOSED PROGRAMME FOR 2012

Date	Module	Content
3 **Feb**, Fri	**ORIENTATION DAY**	
7 **Feb**, Tue	16h00-17h00: Departmental Meeting 2011 (Entire Department)[††] After DM : Welcoming Braai and Drinks (Entire Department)* ([††] Zoo Minor 01 Lecture Theatre, [*] Outside Biological Sciences Building)	
6 **Feb**, Mon - 10 **Feb**,Fri [20 contact hours]	**Introduction to Linux** Prof Philip Machanick	• Linux operating system and software installation • Use of Linux and Linux shell commands
13 **Feb**, Mon – 2 **Mar**, Fri [75 hours]	**Python for Bioinformatics** Mr Gustavo Adolfo Salazar Orejuela	• Introductory and advanced Python • Biopython
5 Mar, Mon	**Announcement of the research projects**	
5 **Mar**, Mon – 9 **Mar**, Fri [25 hours]	**Basic Genomics – Part 1** Dr Özlem Taştan Bishop	DNA and protein databases; database searching; sequence alignment; domain and motif searches; gene and promoter prediction

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

12 **Mar**, Mon – 16 **Mar**, Fri	**Self study – Python (assignment week)**	
19 **Mar**, Mon – 20 **Mar**, Tue	**Basic Genomics - Part 2** Prof Philip Machanick	
21 **Mar**, Wed – 27 **Mar**, Tue	• **Distribution of the projects** • **Self study for exams**	
28 **Mar**, Wed – 30 **Mar**, Fri	**EXAMINATIONS:** • **Basic genomics (Part 1 & 2)** (28 March) • **Python** (29 March) • **Linux** (30 March)	
2 **Apr,** Mon – 13 **Apr**, Fri	**Self study - (Project Proposals)** **(6 Apr - 9 Apr - Easter holiday)**	
16 **Apr,** Mon	**1st project proposal presentations – Literature review**	
17 **Apr**, Tue - 20 **Apr**, Fri [10 hours]	**Phylogenetics** Dr Michael Ludewig	Introduction to phylogenetics; programs for phylogenetics; applications
23 **Apr**, Mon - 27 **Apr**, Fri [25 hours]	**Introduction to Mathematics and Mathematical Programming** Dr Özlem Taştan Bishop & Prof Nigel Bishop	Basic Calculus; linear algebra The MATLAB and maxima computational environments, MATLAB scripts, matrix, graphical output, functions, systems of linear equations, eigenvalues, bioinformatics toolkit

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

30 **Apr**, Mon - 4 **May**, Fri [25 hours]	**Structural Bioinformatics - I** Dr Özlem Taştan Bishop	Protein visualization programs; structural biology techniques; template and non-template based protein structure prediction methods; homology modeling; Modeler.
7 **May**, Mon – 11 **May**, Fri [25 hours]	**Structural Bioinformatics - II** Dr Kevin Lobb	Molecular dynamics; protein-small molecule interactions; Autodock.
14 **May**, Mon – 18 **Mar**, Fri	**Study week**	
21 **May,** Mon– 25 **May,** Fri [25 hours]	**Statistics** Prof. Gunther Jäger & Mr. Jeremy Baxter	Introductory statistics; R: statistical software; hidden Markov models (HMMs)
28 **May,** Mon - 1 **Jun** Fri [25 hours]	**Advanced Genomics** Prof. Fourie Joubert	Genome annotation; genome assembly and analysis; comparative genomic
4 **Jun,** Mon - 6 **Jun** Wed [10 hours]	**Databases** Prof. Fourie Joubert	Mass storage of data, classic file handling techniques, modern databases and MySQL and Turbogears commands.
7 **Jun**, Thr – 8 **Jun**, Fri [10 hours]	**Structural Bioinformatics - II (continue)** Dr Kevin Lobb	NMR

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

11 **Jun,** Mon – 15 **Jun,** Fri	**Study week**
18 **Jun**, Mon – 27 **Jun**, Wed	**EXAMINATIONS:** • **Phylogenetics** (18 June) • **Mathematics** (19 June) • **SB Part 1** (20 June) • **SB Part 2** (22 June) • **Statistics** (25 June) • **Advanced genomics** (26 June) • **Databases** (27 June)
28 **Jun,** Thr – 8 **Jul,** Sun	**BREAK**
18 **Jul,** Wed	**PROJECTS:** Hand-in Literature Review and Project Proposal to Supervisor and Co-supervisor – **Project starts!**
23 **Jul,** Mon	**PROJECTS:** Project Proposal Presentations
1 **Aug**, Wed – 16 **Nov**, Fri	**BIOINFORMATICS JOURNAL CLUB**
24 **Sep**, Mon	1st Presentation of Project Progress
22 **Oct**, Mon	2nd Presentation of Project Progress
26 **Nov**, Mon	Presentation of project results
10-14 **Dec**	**Thesis submission**

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

OVERALL TEACHING HOURS:

	Contact Hours	Lecturing Hours	Practicals and tutorials
Introduction to Linux	20	9	11
Python for Bioinformatics	75	34	41
Basic Genomics - Part 1	25	11	14
Basic Genomics - Part 2	10	5	5
Phylogenetics	10	5	5
Mathematics - Part 1	10	5	5
Mathematics - Part 2	15	7	8
Structural Bioinformatics - I	25	11	14
Structural Bioinformatics - II	35	16	19
Statistics - Part 1	15	7	8
Statistics - Part 2	10	5	5
Advanced Genomics	25	11	14
Databases	10	5	5
TOTAL	285	130	155

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

COURSE OUTCOMES

CRITICAL OUTCOMES ADDRESSED

1. Identify and solve problems and make decisions using critical and creative thinking
2. Work effectively with others as a team
3. Organise and manage time and activities effectively
4. Collect, analyse, organise, and critically evaluate information
5. Communicate effectively using written, electronic and language skills
6. Use science and technology effectively and critically showing responsibility towards the environment and others
7. Demonstrate an understanding of the world as a set of related systems

SPECIFIC OUTCOMES ADDRESSED:

1. Develop a broad understanding of what the field of Bioinformatics and Computational Molecular Biology comprises
2. Develop an in-depth knowledge of certain major areas of Bioinformatics and Computational Molecular Biology
3. Demonstrate the ability to conduct research by designing and carrying out a piece of research in Bioinformatics and Computational Molecular Biology, including design of computational experiments and collection and analysis of data
4. Demonstrate expertise in scientific writing, oral presentation and communication
5. Demonstrate an understanding of the relationship between Bioinformatics and Computational Molecular Biology, the community and the environment
6. Demonstrate the competence required for recognition as a professional Bioinformaticist or Computational Molecular Biologist in South Africa
7. Develop professional attitudes and values including scientific ethics and integrity

PARTICULAR SKILLS TO BE ACQUIRED:

1. Scientific communication and presentation skills including computer skills
2. Ability to use the scientific literature efficiently and effectively
3. Practical skills required for use and application of computers and software
4. Organisational skills required to acquire, manage and utilise data and information
5. Ability to analyse and evaluate scientific data
6. Good computer practice

GENERAL BACKGROUND & OUTCOMES

Bioinformatics and computational molecular biology is the systematic development and application of information technologies and data mining techniques for analysing biological data obtained by experiments, modelling, database searching and instrumentation to make novel observations and predictions about biological function. This course will be taught in an interdisciplinary manner and focussing on the interface between the computational sciences and the biological, physical and chemical sciences. Graduates who complete this course will

MSc PROGRAMME OUTLINE MSc in Bioinformatics and Computational Molecular Biology		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

be skilled in the assimilation of biological information through the use and development of computational tools for a range of applications including simple pattern recognition, molecular modelling for the prediction of structure and function, gene discovery and drug target discovery, the analysis of phylogenetic relationships, whole genome analysis and the comparison of genetic organization.

COURSE STRUCTURE, TEACHING METHODS & APPROACH

The Masters programme will be offered over 11 months and incorporate a number of course work modules and a research project running concurrently throughout the programme. The course work modules will involve an integration of formal lectures, self-learning computer-based tutorials and practicals. In addition, problem solving tutorials would be designed to guide the student through current information-based problems and involve the assimilation and reduction of biological information. A number of the tutorials and practical components will be assessed and contribute towards a course work year mark. The assessment of the course work component would be through assignments, tutorials, tests etc., and examinations. Each examination will have an external examiner, appointed by the lecturer's home Department (for lecturers from Rhodes), or by the Department of Biochemistry, Biotechnology and Microbiology (for external lecturers).

The research projects will be computer based. The projects will be assessed by seminar presentations of the proposed and final work, and by a written thesis. Each thesis will be examined by two external examiners.

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

ASSESSMENT

OVERALL

The course-work and the research work will each contribute 50% to an overall mark. An overall mark of at least 50% is required, subject to a subminimum of at least 40% for the course work, and at least 50% for the research work.

COURSE WORK

The course-work modules will be assessed by internal grading of tutorials, assignments, tests and practicals, etc. to give a class mark; and by internal and external grading of examinations. The calculation of the class mark for each module is given later in this manual under the detailed entry for the module. The examinations will be given during the period specified in the course programme earlier in this manual. For each module, the weighting between class mark and examination towards the module mark will be

- Class mark 40%
- Examinations 60%

The weightings of the various modules in the calculation of the overall course work mark will be proportional to the number of lectures given. For each module the weighting, and the duration of the examination, will be

Module	**Weighting**	**Duration (hours)**
Introduction to Linux	7.0%	2
Python for Bioinformatics	26.2%	4
Basic Genomics	12.3%	2.5
Phylogenetics	3.5%	1
Mathematical Programming	8.8%	2
Structural Bioinformatics – I	8.8%	2
Structural Bioinformatics – II	12.3%	2.5
Statistics	8.8%	2
Advanced Genomics	8.8%	2
Databases	3.5%	1

A student whose overall coursework mark is less than 40% will be deemed to have failed the programme. In addition, a student getting less than 40% in two or more modules, at the discretion of the Department, may also be deemed to have failed the programme.

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

PROJECT
The project will be graded internally and externally with the following weightings:

- Project proposal and presentation 10%
- Project results and presentations 30%
- Thesis 60%

PROPOSAL:

Guidelines
Preparation for the Research Project Proposal (written and oral) should be commenced as soon as the projects have been allocated.

Written
Style: Follow the style of any journal article on Bioinformatics
Length: Around 20 typed pages. Include sections on: Literature review (around 15 pgs); problem statement and hypothesis (1 pg); aims and objectives (1 page); outline of approach and methodology (1–2 pgs).
References: Follow the citation and listing style of the journal, (references may be single-spaced).

Oral
Length: 30 minutes; 25 minutes presentation and 5 minutes questions.

Dates
As specified in the programme earlier in this manual.

Marks Breakdown

- Proposal presentation: 50%
- Written proposal: 50%

PRESENTATION OF PROJECT RESULTS:

Guidelines
The Research Project Results presentation should include:

- **Introduction** - an explanation of the background to the project, the current status of the scientific field, a clear hypothesis statement, and the overall aims & objectives of the project.
- **Description of the approach**, the techniques and methodology, including reasons for why these computations were done.

- **Presentation and Explanation of Results.**
- **Critical discussion of results** including analysis of their implications, and any problem areas.
- **Conclusion** that includes the overall outcome of the project and where future research should be directed.

Dates
As specified in the programme earlier in this manual.

THESIS:

Structure
There is some flexibility in the choice of format for a thesis, but as a guide, it should contain the following sections in the order given:

Abstract
Table of Contents
Table of Figures
List of Tables
List of Abbreviations
Acknowledgments
Chapters 1 (Literature review)
Chapter 2, 3, etc
Conclusion
References

Each Chapter following Chapter 1 would normally contain
Introduction
Methods
Results and Discussion

Dates
As specified in the programme earlier in this manual.

ASSESSMENT CRITERIA & PROCEDURE
The thesis will be assessed by two external examiners. Preferably, at least one of the external examiners should be international.

NUMBER OF COPIES OF THE RESEARCH REPORT
You should prepare two copies of your thesis for external examiners. After corrections are done, one final copy should be prepared for RUBi.

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

DESCRIPTION OF THE MAJOR SECTIONS OF THE THESIS

1. ABSTRACT

An abstract has to stand alone and should: (i) state the principal objectives and scope of the investigation; (ii) state the methodology used; (iii) summarize the results; (iv) state the principal conclusions. It should not exceed a page.

2. CHAPTER 1

Literature review

This should be a concise summary that describes the current status of the research field. It should be current and comprehensive.

Project aims, objectives and motivation

A clear statement of the aims & objectives of the project and motivation for these should be given. Knowledge gap should be explained.

3. FURTHER CHAPTERS

Introduction

This should be a concise summary that describes the current status of the literature related to the chapter.

Methodology

This should give a logical account of the methodology. It should be precise and complete.

Results and Conclusion

This section should give a description of the results of the experiments together with an explanation of why they were done. It should include critical analysis of the data and interpretation of the implications of the results.

5. CONCLUSION

Should be a concise and relevant summary, including the contribution the research makes to the current status of the field. A statement of the direction of future research arising from the project should be given.

6. REFERENCES

Current research articles should be used and cited in the text of the thesis using the style of a bioinformatics journal.

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

MSc Proposal Presentation Evaluation Form

Criterion	**Weight**	**Mark**
Concise, accurate & up-to-date literature review	**20**	
Knowledge gap and/or problem clearly identified and stated	**20**	
Clear research hypothesis & objectives; Concise description of approach and methods	**20**	
Research objectives, approach & methods. Realistic? Feasible?	**15**	
Time management, visual media and speaker – audience contact	**10**	
Ability of speaker to answer questions in a clear & meaningful manner.	**15**	
Total	**100**	

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

MSc Written Proposal Evaluation Form

Criterion	**Weight**	**Mark**
Concise, accurate & up-to-date literature review	**30**	
Knowledge gap and/or problem clearly identified and stated	**20**	
Clear research hypothesis & objectives; Concise description of approach and methods	**20**	
Research objectives, approach & methods. Realistic? Feasible?	**15**	
Quality of scientific writing	**15**	
Total	**100**	

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

MSc Project Presentation Evaluation Form

Criterion	**Weight**	**Mark**
Concise, accurate & up-to-date literature review	**15**	
Knowledge gap and/or problem clearly identified and stated	**15**	
Clear research hypothesis & objectives; Concise description of approach and methods	**15**	
Results and discussion: interpretation of results and critical analysis of their meaning and impact	**25**	
Summary of findings and future plans	**5**	
Time management, visual media and speaker – audience contact	**10**	
Ability of speaker to answer questions in a clear & meaningful manner.	**15**	
Total	**100**	

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

COURSE WORK MODULES

INTRODUCTION TO LINUX
Lecturer: Prof Philip Machanick
Dates: 6 February to 10 February 2012
Contact hours: 20

SPECIFIC OUTCOMES ADDRESSED
1. To be able to install a Linux operating system
2. To be able to install various programs
3. Log in and out of a Linux system
4. Work with directories and files and change file permissions
5. Master several shell commands
6. Redirect input and output and print documents

BACKGROUND KNOWLEDGE REQUIRED
Basic computer literacy: proficiency with word-processing, spreadsheets and graphics programmes, exposure to standard bench-top computational tools and the web

TEACHING METHODS/APPROACH
The lectures will be complemented by tutorials and self-study.

BOOKS & OTHER SOURCES USED
Introduction to Linux – A Hand on Guide by Machtelt Garrels (tldp.org/LDP/intro-linux/intro-linux.pdf)

COURSE CONTENT
1. What is Linux?
2. How to install an operating system
3. Quick start
4. About files and file systems
5. Processes
6. I/O redirection
7. Text editors
8. Home
9. Printers and printing
10. Fundamental backup systems
11. Networking
12. Installation of various programs

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Test 1: 40%
2. Test 2: 60%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

PYTHON AND BIOPYTHON
Lecturer: Gustavo Salazar
Dates: 13 February to 2 March 2012
Contact hours: 75

SPECIFIC OUTCOMES ADDRESSED

1. To be able to write short Python program to manipulate data
2. To understand the differences between numbers, strings, lists and arrays
3. To master the use of various control structres and functions within Python program
4. To understand the concepts of the Object Oriented paradigm and how to use it in python
5. To retrieve and manipulate data from databases and files
6. To use the most common procedures in Biopython

BACKGROUND KNOWLEDGE REQUIRED
Basic computer literacy: proficiency with word-processing, spreadsheets and graphics programmes, exposure to standard bench-top computational tools and the web

TEACHING METHODS/APPROACH
Lectures: utilizing self-study tutorials and demonstration programmes
Numerous small exercises to build up experience and skills progressively

BOOKS & OTHER SOURCES USED
Python documentation: http://docs.python.org/index.html
Biopython http://biopython.org/wiki/Biopython

COURSE CONTENT

1. Introduction to Python (Thinking, writing and running)
2. Flow Control
3. Data Structures
4. Strings in Depth
5. Functions
6. Importing Standard Modules
7. Files for Input and Output
8. Regular Expressions
9. Basic Parsing
10. Exceptions and error handling
11. Recursion
12. Classes and Objects
13. Database Theory and Relational Databases
14. Biopython
15. Graphical User Interfaces

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Test: 30%
2. Mini-project: 30%
3. Assignments: 40%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

BASIC GENOMICS – PART 1
Lecturer: Dr Özlem Taştan Bishop
Dates: 5 March to 9 March 2012
Contact hours: 25

SPECIFIC OUTCOMES ADDRESSED

1. Ability to retrive data from databases and analyse the data
2. To be able to align homologous sequences in DNA or protein format and understand the advantages and disadvantages of the two approaches
3. Understand various alignment algorithms

BACKGROUND KNOWLEDGE REQUIRED

Basic biochemistry and genetics knowledge.

TEACHING METHODS/APPROACH

The lectures will be complemented by tutorials, self study and article discussions.

BOOKS & OTHER SOURCES USED

1. Essential Bioinformatics by Jin Xiong
2. Introduction to bioinformatics by Anna Tramontano
3. Bioinformatics – A practical guide to the analysis of genes and proteins by Andreas Baxevanis and Francis Ouellette
4. Research articles and other bioinformatics books in the library

COURSE CONTENT

1. Biological databases
2. Sequence alignment
    a. Pairwise sequence alignment
    b. Database similarity search
    c. Multiple sequence alignment
    d. Profiles and HMMs
    e. Protein motifs and domain predictions

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Assignment 1: 25%
2. Assignment 2: 25%
3. Test: 50%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

BASIC GENOMICS – PART 2
Lecturer: Prof Philip Machanick
Dates: 19-20 March 2012
Contact hours: 10

SPECIFIC OUTCOMES ADDRESSED
Understanding how DNA is computationally analysed for features of interest, including but not limited to transcription factor binding sites.

BACKGROUND KNOWLEDGE REQUIRED
Role of DNA in genetics, basic understanding of developmental biology.

TEACHING METHODS/APPROACH
Lecturing, demonstrating techniques and problem-solving.

BOOKS & OTHER SOURCES USED
Web searches and academic literature.

COURSE CONTENT
1. Transcription factors and DNA sequence analysis
2. Coding and non-coding DNA
3. Conservation and its role in sequence analysis
4. Use of the UCSC genome browser
5. Other web-based tools for sequence analysis

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS
1. Assignment 1: 20%
2. Assignment 2: 30%
3. Write short paper: 50%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

PHYLOGENETIC ANALYSIS
Lecturer: Dr Michael Ludewig
Dates: 16 to 20 April 2012
Contact hours: 10

SPECIFIC OUTCOMES ADDRESSED

1. Introduction to the concepts of molecular phylogenetics
2. Understand phylogenetic trees
3. Generation and interpretation of phylogenetic trees using computer programmes
4. Development of an understanding of the potential applications of phylogenetics

BACKGROUND KNOWLEDGE REQUIRED

1. Generation of multiple sequence alignment
2. Familiarity with genome databases, protein and DNA sequence data

TEACHING METHODS/APPROACH

Lectures will be held in the Department of Biochemistry, Microbiology and Biotechnology. There will be a component of self study with theory and practical assignments. The practical component will involve the generation and interpretation of phylogenetic data.

BOOKS & OTHER SOURCES USED

1. Bioinformatics and Molecular Evolution, Higgs, P.G. and Attwood, T.K., Blackwell Publishing.
2. Understanding Bioinformatics, Zvelebil, M. and Baum, J.O. Garland Science, Taylor & Francis Group.
3. The Phylogenetic Handbook: A Practical Approach to DNA and Protein Phylogeny, Edited by Salemi, M. and Vandamme, A-M. Cambridge University Press.
4. Phylogenetics: the Theory and Practice of Phylogenetic Systematics. E.O. Wiley, John Wiley and Sons.

COURSE CONTENT

This course is designed to be an introduction to phylogenetics and associated computer software.

1. Introduction to phylogenetics concepts
2. Methods of phylogenetic inference
3. The applications of phylogenetics data
4. Practical usage of phylogenetics programmes

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

a. Tutorial 50% b.Practical assignment 50%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

MATHEMATICS – PART 1: (INTRODUCTION TO MATHEMATICS WITH BIOLOGICAL APPLICATIONS)
Lecturer: Dr Özlem Taştan Bishop
Dates: 23 April and 25 April 2012
Contact hours: 10

SPECIFIC OUTCOMES ADDRESSED
1. Describe biological/bioinformatics problems using mathematics.
2. Solve these problems using calculus, linear algebra.
3. Get required background for Matlab and Statistics courses

BACKGROUND KNOWLEDGE REQUIRED
Basic mathematics

TEACHING METHODS/APPROACH
The lectures will be complemented by self-study and tutorials.

BOOKS & OTHER SOURCES USED
Lecture notes
Any Calculus, Linear Algebra books

COURSE CONTENT
1. Calculus (Differentiation and integration)
2. Linear Algebra (Matrices, eigenvalue / eigenvector problems)

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS
1. Assignment 1: 25%
2. Assignment 2: 25%
3. Test: 50%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

MATHEMATICS – PART 2: (MATHEMATICAL PROGRAMMING WITH MATLAB)
Lecturer: Prof Nigel T Bishop
Dates: 24, 26 and 27 April 2012
Contact hours: 15

SPECIFIC OUTCOMES ADDRESSED

1. Introduction to mathematical programming with MATLABand maxima.
2. Solve problems using mathematical programming.
3. Introduction to the MATLAB Bioinformatics toolbox

BACKGROUND KNOWLEDGE REQUIRED
Computer programming in any other language.

TEACHING METHODS/APPROACH
Lectures will be mainly in the form of demonstrations of MATLAB and maxima features, with discussion. Relevant notions from various aspects of mathematics will be discussed as necessary. At each lecture a set of exercises will be presented, which students should complete and submit by the next lecture.

BOOKS & OTHER SOURCES USED
Essential MATLAB for Scientists and Engineers, B Hahn, Pearson, 3 rd edition ISBN 1 868 91143 82

COURSE CONTENT
The purpose of the course is to enable the student to construct a computational environment with MATLAB in which to model, study and simulate real-world processes. It is intended that the student learn this skill by hands-on experience with the computer. The lectures are meant to provide an overview and a forum for discussion. The exercises are there to provide practical experience. Most of the real learning will be accomplished by doing the exercises.

1. Introduction to the MATLAB and maxima environments; programming in MATLAB: statements, data structures, input / output, flow control, functions, graphics
2. Linear algebra with MATLAB and maxima: systems of equations, over-determined systems and linear regression, eigenvalues and eigenvectors
3. Other applications of MATLAB and maxima: differentiation, integration, solving nonlinear equations, cubic splines
4. The MATLAB Bioinformatics toolbox Importing data from a databank, graphical viewing and manipulation of structures, genomics, phylogenetics, alignment

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Assignment 1: 20%
2. Assignment 2: 20%
3. Project: 60%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

STRUCTURAL BIOINFORMATICS – I
Lecturer: Dr Özlem Taştan Bishop
Dates: 30 April to 4 May 2012
Contact hours: 25

SPECIFIC OUTCOMES ADDRESSED

1. To understand structural biology terminology, especially X-ray crystallography, and to be able to follow the literature
2. To learn how to use different protein visualization programs
3. To understand various secondary and tertiary structure prediction algorithms
4. To understand the range, applications and limitations of modeling methods
5. To learn modeling by using Modeller

BACKGROUND KNOWLEDGE REQUIRED

1. Knowledge on biochemical properties of amino acids
2. Basic understanding of the primary, secondary, tertiary and quaternary structure of proteins.
3. Knowledge on non-covalent bond formations

TEACHING METHODS/APPROACH
The lectures will be complemented by tutorials, self study and article discussions.

BOOKS & OTHER SOURCES USED

1. Essential Bioinformatics by Jin Xiong
2. Introduction to bioinformatics by Anna Tramontano
3. Bioinformatics – A practical guide to the analysis of genes and proteins by Andreas Baxevanis and Francis Ouellette
4. Manuals and tutorials of various modeling and visualization programs

COURSE CONTENT

1. Structural biology techniques
2. Protein visualization programs
3. Protein secondary structure prediction
4. Protein tertiary structure prediction
5. Homology modeling; Modeller

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Assignment: 20%
2. Short project: 40%
3. Test: 40%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

STRUCTURAL BIOINFORMATICS – II
Lecturer: Dr. Kevin A. Lobb
Dates: 7 May to 11 May 2012 and 7-8 June 2012
Contact hours: 35

SPECIFIC OUTCOMES ADDRESSED
This course introduces the theory and practice of molecular modelling as used in chemistry and medicinal chemistry. Although competence in the use of several software packages is a critical component, emphasis will be on the understanding of the methods and on strategies in their application to a wide variety of problems.

BACKGROUND KNOWLEDGE REQUIRED
Little background knowledge is required, beyond that of basic chemistry. However it is essential that you are comfortable with chemical structures and that you can quickly identify whether they are correct or incorrect in terms of positioning and the valency of atoms. Familiarity with the any following concepts would be helpful, though not essential as we will deal with what is necessary during the course. Conformational analysis (e.g. boat and chair cyclohexane); orbitals, HOMO, LUMO, bonding and antibonding, excited state; Infrared Spectroscopy; transition state; activation energy; enthalpy, entropy and free energy.

TEACHING METHODS/APPROACH
The teaching will be split equally between lectures and practicals.

BOOKS & OTHER SOURCES USED
User manuals and background from the programs Materials studio, Gaussian, CHARMM, GAMESS, VASP, Autodock, Vega ZZ, CPMD, Sparky and relevant supplied journal articles.

COURSE CONTENT
Theories used in calculations, molecular mechanics, semi-empirical, Hartree-Fock, configuration interaction and density functional theory. Correlation energy. Basis sets. Strategies for dealing with extremely large systems. Combined methods QM/MM, ONIOM, discrete and continuum solvation. Exploring the potential energy surface and vibrational analysis. Conformational searches. Calculable properties. Excited states. Calculations in vacuo, periodic boundary conditions. Molecular dynamics (MM, Born-Oppenheimer and Car-Parrinello). Interaction between systems – basis set superposition error, protein-small molecule interactions and docking. NMR – relaxation, coupling and relevant experiments used in biomolecular NMR. Principles of structure assignment. Protein-ligand interactions by NMR.

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS
There will be an assignment which will make up 100% of the mark for this course.

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

INTRODUCTORY STATISTICS
Lecturer: Mr. Jeremy Baxter
Dates: 21 - 23 May 2012
Contact hours: 15

Lecturer: Prof. Gunther Jäger
Dates: 24 -25 May 2012
Contact hours: 10

SPECIFIC OUTCOMES ADDRESSED
The aims of this course are:

1. To provide students with an introduction to statistics in order to solve problems of bioinformatics.
2. To provide students with the basics of probability theory (probability, probability axioms, conditional probability, probability density function, cumulative distribution function, expectation, variance, discrete random variable, continuous random variable) and statistical background, concepts and techniques (statistical experiment, descriptive statistics, inference statistics) that are most useful to Bioinformaticians.
3. To provide students with the theory and application of Hidden Markov Models.

On completion of the course students should, inter alia, be able to:

1. Explain the differences between a population and a sample.
2. Collect, summarise and describe data using suitable numerical and graphical techniques.
3. Explain the concepts of probability, interpret probabilities and use suitable theory to calculate simple and conditional probabilities.
4. Identify discrete and continuous probability distributions.
5. Demonstrate the use of the binomial, Poisson, normal, Student t, chi-square and F distributions.
6. Calculate point and interval estimates, one- and two-sample, for the population mean(s), proportion(s) and variance(s) and interpret the meaning of each.
7. Perform suitable hypothesis tests (parametric and or non-parametric procedure) for one- and two-sample analyses and draw meaningful conclusions and decisions for the population mean(s), proportion(s) and variance(s).
8. Estimate, interpret and make predictions using linear models. Perform suitable statistical inference and model diagnostics for linear models.
9. Understand and apply the theory of Hidden Markov Models.

BACKGROUND KNOWLEDGE REQUIRED

1. Basic Calculus: Differentiation and integration
2. Linear algebra: Matrices, vectors

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

3. Matlab literacy, specifically matrix operations.
4. Basic programming experience, in python or perl

TEACHING METHODS/APPROACH

This course will be taught using formal lectures, typically in the morning, and self-study tutorials and practicals. Use of hand-outs, notes, text books, board-work and overheads. Relevant notions from linear algebra and statistics will be discussed and the student will then be required to read portions of prescribed texts on his/her own. At each lecture a set of exercises will be presented and completed ready for assessment by the next lecture.

BOOKS & OTHER SOURCES USED

1. J Baxter, Introductory Statistics for Bioinformaticians using R (course notes/slides).
2. G Jäger, Markov chains and Hidden Markov models (course notes).
3. Wim P. Krijnen, 2009, Applied Statistics for Bioinformatics using R, CRAN

COURSE CONTENT

1. A brief introduction to R.
2. Descriptive statistics (Graphical and numerical summaries of univariate, bivariate and multivariate data).
3. An introduction to statistical distributions.
4. Estimation and inference for one/ two random samples (Parametric and non parametric methods.)
5. An introduction to correlation, linear regression and linear models: (One and Two Way ANOVA)
6. Markov chains: basic properties, concepts and examples.
7. Hidden Markov models: backward and forward algorithm, Viterbi algorithm, Baum-Welch algorithm, applications.

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

1. Daily assignments/exercises: 40%
2. Tests: 60%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

ADVANCED GENOMICS
Lecturer: Prof Fourie Joubert
Dates: 28 May to 1 June 2012
Contact hours: 25

SPECIFIC OUTCOMES ADDRESSED

1. Genome assembly
2. Next-generation sequencing methods
3. Next generation sequencing data analysis
4. Genome annotation
5. Comparative genomics

BACKGROUND KNOWLEDGE REQUIRED

Students need to have the necessary biological knowledge as related to prokaryotic and eukaryotic genome features and genome structures. They also need to know the basic techniques used in the analysis of sequence features.

TEACHING METHODS/APPROACH

1. The course will be taught over five days from 8:30 to 16:30
2. It will include theory and practical sessions.

BOOKS & OTHER SOURCES USED

1. Baxevanis & Ouellette 3rd Edition.
2. Free internet sources.

COURSE CONTENT

The course aims to educate students regarding the assembly and analysis genomic data. It addresses concepts and techniques related to this analysis.

This includes:

1. Genome assembly
2. Next-generation sequencing methods
3. Next generation sequencing data analysis
4. Genome annotation
5. Comparative genomics

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS

Completion of assignment: 100%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

DATABASES
Lecturer: Prof Fourie Joubert
Dates: 4 – 5 June 2012
Contact hours: 10

SPECIFIC OUTCOMES ADDRESSED

1. Understanding the need for relational databases
2. Understanding the structure of relational databases
3. Understanding Structured Query Language

BACKGROUND KNOWLEDGE REQUIRED

1. Basic computer knowledge
2. Python programming knowledge.

TEACHING METHODS/APPROACH

3. The course will be taught over two days from 8:30 to 16:30
4. It will include theory and practical sessions.

BOOKS & OTHER SOURCES USED

1. Database Systems by Connolly & Begg
2. Free Internet sources

COURSE CONTENT
The course is aimed at introducing students to relational databases. It will explain the need for structured databases systems, together with their advantages and disadvantages. It will then explore relational databases together with the necessary language aspects for use of these systems.

1. Introduction to databases
2. Databases and database management systems
3. Database design
4. Database relations
5. Structured Query Language

ASSESSMENT ACTIVITIES AND THEIR WEIGHTS
Completion of assignment: 100%

MSc PROGRAMME OUTLINE **MSc in Bioinformatics and Computational Molecular Biology**		
Course Work / Project Masters	**Year: 2012**	**Coordinator: Dr Özlem TAŞTAN BISHOP**

CONTACT DETAILS OF LECTURERS

Jeremy Baxter
Department of Statistics, Rhodes University
E-mail: J.Baxter@ru.ac.za

Nigel Bishop
Department of Mathematics, Rhodes University
E-mail: N.Bishop@ru.ac.za

Gunther Jäger
Department of Statistics, Rhodes University
E-mail: G.Jager@ru.ac.za

Fourie Joubert
Bioinformatics and Computational Biology Unit, University of Pretoria
E-mail: Fourie.Joubert@up.ac.za

Kevin A. Lobb
Department of Chemistry, Rhodes University
E-mail: K.Lobb@ru.ac.za

Michael H. Ludewig
Biomedical and Biotechnology Research Unit (BioBRU)
Department of Biochemistry, Biotechnology and Microbiology, Rhodes University
E-mail: M.Ludewig@ru.ac.za

Philip Machanick
Department of Computer Science
E-mail: P.Machanick@ru.ac.za

Gustavo Adolfo Salazar Orejuela
University of Cape Town
E-mail: gsalazar@cs.uct.ac.za

Özlem Taştan Bishop
Rhodes University Bioinformatics (RUBi)
Department of Biochemistry, Microbiology and Biotechnology, Rhodes University
E-mail: O.TastanBishop@ru.ac.za