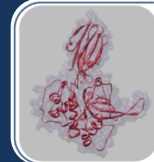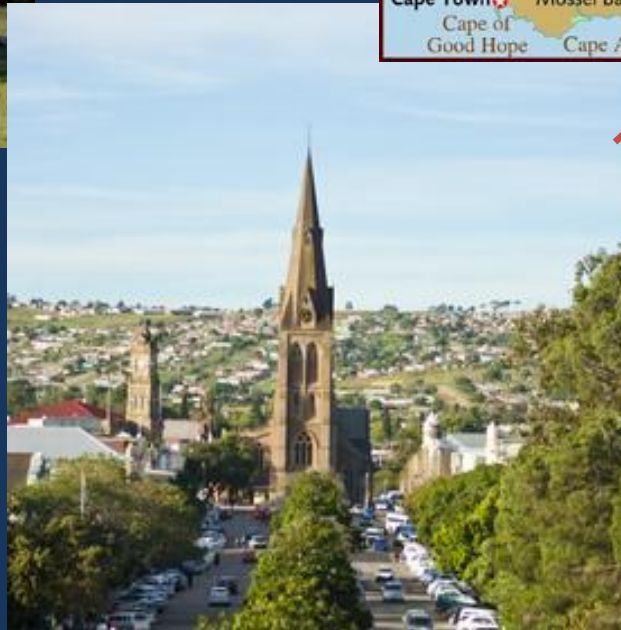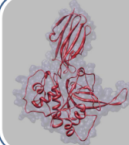# One-Year Coursework and Dissertation Bioinformatics MSc Programme at Rhodes University

Özlem Taştan Bishop
Research Unit in Bioinformatics (RUBi), Rhodes University
www.rubi.ru.ac.za
o.tastanbishop@ru.ac.za

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

November 2009 – February 2010

RHODES UNIVERSITY
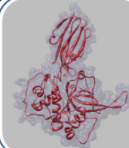Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

RUBi -2012

RUBi - 2014

- 6 PhD (+1 Biochem PhD)
- 8 MSc (one-year programme)
- 1 MSc (2 years research thesis)
- 2 Honours

Who is in RUBi

- **Head of the Unit**
  - OzlemTastan Bishop (Biochemistry & Microbiology)
- **Core members**
  - Philip Machanick (Computer Science, 2012)
  - Kevin Lobb (Chemistry, 2011)
- **Academic staff**
  - Jeremy Baxter (Statistics)
  - Nigel Bishop (Mathematics)
  - Mike Burton (Mathematics)
  - Adrienne Edkins (Biochemistry & Microbiology)
  - Mike Ludewig (Biochemistry & Microbiology, Postdoc)
  - Denis Pollney (Mathematics)
- **External lecturers**
  - Fourie Joubert (UP)
  - Oleg Reva (UP)
  - Gustavo Salazar (UCT, PhD student)
  - Jane Wright (Forensic company)

One-year Coursework and Dissertation MSc Programme:

In detail

Since bioinformatics is interdisciplinary, the programme accepts students from different backgrounds, and aims to bring them to an equal level of interdisciplinary knowledge.

The aim is to produce MSc graduates with a strong foundation.

The one-year MSc programme provides a bridging role; one end with multiple legs to get students from different disciplines, and the other end to transfer those suitable to a bioinformatics PhD degree.

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

Details of MSc Students

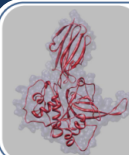| Year	Student number	Country	Gender	Background	Financial support	PhD continuation	Bioinformatics related employment
2011	3	1 South Africa 2 Kenya	1 Female 2 Male	1 Biochemistry 1 Biomedical Technology 1 Pharmacy	All University or government supported	1 at Rhodes University	1 in Kenya
2012	7	3 South Africa 3 Kenya 1 Nigeria	4 Female 3 Male	3 Biochemistry 1 Biochemistry/Microbiology 1 Medical Laboratory Science 1 Medical Microbiology 1 Zoology	All University or government supported	2 at Rhodes University 1 at University of Pretoria 2 at University of Western Cape 1 at University of the Witwatersrand	
2013	4	2 South Africa 1 Kenya 1 Zimbabwe	4 Male	2 Biochemistry 1 Biotechnology 1 Mathematics/Comp Science	All University or government supported	3 at Rhodes University	
2014	8	4 South Africa 1 Botswana 1 Lesotho 2 Zimbabwe	3 Female 5 Male	1 Biochemistry 2 Biochemistry/Microbiology 1 Biotechnology 1 Chemistry 1 Mathematics/Comp Science 1 Microbiology 1 Molecular and Cell Biology	6 University or government supported 2 self supported	N/A	N/A

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

History

- Ran in 2003 and 2004
- 6 months NBN central training
- NBN closed down in 2008
- Re-initiated in 2011

Overview of the Programme

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	_____Coursework_____					B ____Project & Thesis_____					
	Project proposal										
			E		E	PP		P	P		FP

B - Break
E – Exams
PP – Project proposal presentation
P – Progress report presentation
FP – Final presentation

RHODES UNIVERSITY
Grabamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

Overview of the assessment

- The coursework and the research project each contribute 50% to an overall mark.

- Successful completion of the degree is subject to a final mark of at least 50%, provided that a candidate obtains at least 50% for the coursework, with a sub-minimum of at least 40% from each module, and at least 50% for the project thesis.
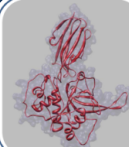
Overview of the assessment

- The coursework modules are assessed by internal grading of tutorials, assignments, tests and practicals, etc. to give a class mark, and by internal and external grading of examinations.

- For each module, the weighting for the class mark is 40%, and for the examination is 60%. The weightings of the various modules in the calculation of the overall coursework mark is proportional to the number of lectures given.

RUBi
Research Unit in Bioinformatics

Overview of the assessment

- The project is graded internally by evaluating
  - project proposal and presentation (10%)
  - project results and presentations (30%)

- The thesis (60%) is graded externally by two external examiners.

- The university prefers that at least one of the external examiners be international.

RUBi
Research Unit in Bioinformatics

Coursework

Core modules	Content	Duration-contact hours
Introduction to Linux	• Linux operating system and software installation • Use of Linux and Linux shell commands • Application to bioinformatics problems	20
Introduction to Programming	• Basics for (Python) programming	10
Python for Bioinformatics	• Introductory and advanced Python • Biopython	75 & 1 week for assignment
Basic Mathematics	• Review of basic calculus • Review of linear algebra	15
Mathematical Programming	• The MATLAB computational environment, MATLAB scripts, graphical output, functions, systems of linear and non-linear equations, differential equations • Use of the Bioinformatics Toolbox	20
Statistics	• Introductory statistics • R: statistical software	25
Basic Genomics – Part I	• Genome sequencing techniques • DNA and protein databases; database searching • Databases and API • Sequence alignment	25
Basic Genomics – Part II	• Discovering features of interest in DNA including transcription factor binding sites • Using genome browsers to obtain data • Using web services and the command line to performance genome-wide and specific sequence analyses	25
Comparative Genomics	• Introduction to pairwise and multiple complete genome alignment • Phylogenomics • Genome evolution and horizontal gene transfer • New approaches, techniques and challenges	25
Structural Bioinformatics – Part I	• Protein visualization programs; • Structural biology techniques • Template and non-template based protein structure prediction methods • Homology modeling in detail	25
Structural Bioinformatics – Part II	• NMR • Docking (Autodock) • Molecular dynamics	25

Coursework

Supplementary modules	Content	Duration- contact hours
Databases	• Introduction to databases • Introduction to web frameworks • MySQL; Django	25
Neural Network	Origins of artificial neural networks, perceptrons: their construction and deployment, convergence of perceptrons, gradient descent for optimisation, general feed-forward networks with differentiable transfer functions, backpropagation, training, assessing performance, construction and deployment of feed-forward neural networks for prediction and pattern recognition, various applications, problems.	25
Phylogenetics	Introductory phylogenetics covering neighbor-joining in detail and the principles of maximum likelihood and Bayesian inference. Bootstrap analysis, evolutionary models and comparison of tree topologies.	25
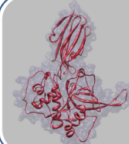
Project component

- Time dedicated purely to research is very limited (July – December)

- Techniques used to increase research efficiency and critical thinking ability of students:
  - Freedom in the project
  - Project presentations
  - Weekly research meetings
  - Thesis update – every two weeks
  - Journal club meetings
  - Public bioinformatics talks and private group meetings to improve presentation skills

RUBi
Research Unit in Bioinformatics

Research Projects

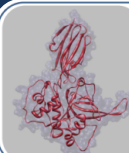| Year	Thesis title	Methodology used – Novel findings
2011	*In silico* Characterisation of the Four Canonical *Plasmodium falciparum* 70 kDa Heat Shock Proteins	Data retrieval, sequence alignment, homology modeling, protein-protein interaction analysis * A novel modeling method was suggested
	Structural Analysis of Prodomain Inhibition of Cysteine Proteases in Plasmodium Species	Data retrieval, sequence alignment, phylogenetic tree analysis, homology modeling, interprotein bonding analysis *Few short peptides as potential inhibitors were designed*
	Structural Analysis of Effects of Mutations on HIV-1 Subtype C Protease Active Site	Data retrieval, sequence alignment, homology modeling, large scale ligand docking * *South African HIV-positive infants were analysed for drug related mutations*
2012	Falcipains as Malarial Drug Targets	Data retrieval, sequence alignment, phylogenetic tree analysis, homology modeling, ligand docking, protein-ligand interaction analysis * *A South African natural product was identified as a potential malarial inhibitor*
	Sequence and Structural Investigation of the Nonribosomal Peptide Synthetases of *Bacillus atrophaeus* UCMB 5137(63Z)	Genome annotation for NRPS modules, phylogenetic studies, homology modeling and structural analysis, motif analysis * *New modules were identified and linker regions were analysed*
	In silico Analysis of *Plasmodium falciparum* Hop Protein and Its Interactions with Hsp70 and Hsp90	Data retrieval, sequence alignment, phylogenetic tree analysis, motif analysis, homology modeling, protein-protein interface analysis, alanine scanning * *New Hop-Hsp90 binding region was modelled and differences between human and parasite proteins were identified*
	Structural Bioinformatics Analysis of Plasmodium DNAJ Proteins and the Interactions with *Plasmodium falciparum* Hsp70s	Data retrieval, large scale sequence alignment, grouping, homology modeling, docking, protein-protein interaction analysis * *New Plasmodium DNAJ proteins identified*
	Structural Bioinformatics Analysis of the HSP40 and HSP70 Molecular Chaperones from Humans	Data retrieval, large scale sequence alignment, grouping, homology modeling, docking, protein-protein interaction analysis * *Detailed analysis between Hsp40 and Hsp70 was done*
	Influence of Non-Synonymous Sequence Mutations on the Architecture of HIV-1 Clade C Protease Receptor Site: Docking and Molecular Dynamics Studies	Data retrieval, sequence alignment, homology modeling, large scale ligand docking, protein-ligand interaction analysis , preliminary MD calculations * *Analysis on differences on drug resistance between Clade B and Clade C showed interesting results*
	A central enrichment-based comparison of two alternative methods of generating transcription factor binding motifs from protein binding microarray data	Data retrieval, motif enrichment analysis * *Differences in motif quality across two competing databases measured*
2013	A Step Forward in Defining Hsp90s as Potential Drug Targets for Human Parasitic Diseases	Data retrieval, analysis of physicochemical properties (in large scale), phylogenetic tree calculations, motif analysis (script based analysis), SCA analysis (co-evolution studies) * *Some important differences between the parasite and human Hsp90 proteins were identified*
	Large Scale Bioinformatics Analysis of Auxiliary Activity Family 9 Enzymes	Data retrieval, sequence alignment, phylogenetic tree analysis, motif analysis, homology modeling * A novel sub-type group is identified
	Transcription factor binding: Investigating the role of distance between transcription factor binding site and transcription start site	Data retrieved, statistical analysis * *Differential locality of transcription finding with respect to transcription start site measured*
	Analysis of transcription factor binding specificity using ChiP-seq data	Data retrieval, differential motif enrichment analysis * *Differential motif enrichment across cell lines measured*

Outcomes

- Research-productive graduate students
  - Over 70% (10 out of 14) of completing students have continued in PhD programmes in various South African universities
  - Only one student (7%) has, to date, moved on to bioinformatics-related employment (Biosciences Eastern and Central Africa, Kenya)
  - We have no information on the remaining 3 students.

- As the research time is very limited, the projects are small and do not lead to publications immediately. When students stay on for a PhD degree, there is more opportunity to complete the project and write a papers

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

Details of MSc Students

Year	Student number	Country	Gender	Background	Financial support	PhD continuation	Bioinformatics related employment
2011	3	1 South Africa 2 Kenya	1 Female 2 Male	1 Biochemistry 1 Biomedical Technology 1 Pharmacy	All University or government supported	1 at Rhodes University	1 in Kenya
2012	7	3 South Africa 3 Kenya 1 Nigeria	4 Female 3 Male	3 Biochemistry 1 Biochemistry/Microbiology 1 Medical Laboratory Science 1 Medical Microbiology 1 Zoology	All University or government supported	2 at Rhodes University 1 at University of Pretoria 2 at University of Western Cape 1at University of the Witwatersrand	
2013	4	2 South Africa 1 Kenya 1 Zimbabwe	4 Male	2 Biochemistry 1 Biotechnology 1 Mathematics/Comp Science	All University or government supported	3 at Rhodes University	
2014	8	4 South Africa 1 Botswana 1 Lesotho 2 Zimbabwe	3 Female 5 Male	1 Biochemistry 2 Biochemistry/Microbiology 1 Biotechnology 1 Chemistry 1 Mathematics/Comp Science 1 Microbiology 1 Molecular and Cell Biology	6 University or government supported 2 self supported	N/A	N/A

RHODES UNIVERSITY
Grabamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics
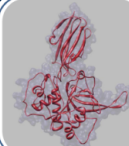
Challenges and our solutions

Required infrastructure and administrative support:

- **Lack of a dedicated Linux computer laboratory:**
  - a seminar room was converted into a lecture room
  - laptops are provided to the students (excellent mobility)
  - All the lectures and practicals are held in the same place
  - For large computational runs, server access is provided

- **Lack of adequate system administrative support:**
  - extra load for to the course coordinator/faculty
  - Linux module teaches students how to install Linux and how to handle basic system administration
 - This approach adds to the skills acquired by the students

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

Challenges and our solutions

Range of topics:

- The programme is unable to cover every bioinformatics topic:
    - limited staff members
   - time limitations

    Most of the specialized bioinformatics modules are designed according to the research interests of the supervisors, and further learning in these topics is supported by research projects. Although students are not exposed to a wide range of bioinformatics topics, they gain strong background in core bioinformatics modules, and they learn how to do independent research on new bioinformatics topics.
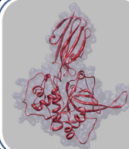
Challenges and our solutions

_Collaborative effort from different departments:_

- **The challenge : to familiarize lecturers from other research fields with bioinformatics applications:**
  - For instance, a lecturer might be excellent in teaching mathematical programming and Matlab, but he or she might not be familiar with the bioinformatics tool kit.
  - In these cases, a bioinformatics lecturer also participates to help the Matlab lecturer bridge the two fields.
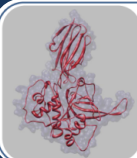
Challenges and our solutions

Diversity of students:

- It is challenging to take students from a diverse range of backgrounds and levels, and teach them bioinformatics in depth in a short time.

  - However, this diversity is used as an advantage in the programme by encouraging students with different backgrounds to work together.

  - Furthermore, having students with different backgrounds has important impacts in the classroom, as they analyse a topic from different angles.

RHODES UNIVERSITY
Grabamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics
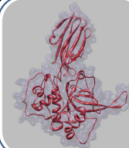
Challenges and our solutions

Finding examiners:

- One of the biggest difficulties that we have is to find external examiners to review all the research theses at once.

- Consistent marking process!

- Each of the research theses is about 100 pages!

RHODES UNIVERSITY
Grabamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics

Conclusion

- Modules aim to bridge gaps in the diverse backgrounds of students who range from biologists with no prior computing exposure to computer scientists with no biology background.

- The programme is evenly split between coursework and research, with diverse modules from a range of departments covering mathematics, statistics, computer science and biology, with emphasis on application to bioinformatics research.

- The early focus on research helps bring students up to speed with working as a researcher.

- We measure success of the programme by the high rate of subsequent entry to PhD study: 10 out of 14 students who completed in the years 2011-2013.

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

RUBi
Research Unit in Bioinformatics