H3ABioNet
Pan African Bioinformatics Network for H3Africa

Issue 17: August 2016

H3ABioNet

In This Issue

Forward

August has been another busy month in H3ABioNet, particularly with the running of the Cloud Hackathon and final organisation of the DREAM challenges for the Malaria hackathon. I joined the first day of the Cloud hackathon, and was excited by the enthusiasm of the participants to contribute to the development of Docker containers for pipelines relevant to H3Africa project data. We had experts attending who made excellent contributions and I think all participants learnt something during the week. You can read more about the hackathon later in the newsletter.

The Introduction to Bioinformatics course is continuing to run smoothly and again I am encouraged by the participant's and teaching assistants' dedication. Online courses usually suffer from reduced participation as they proceed, but this course has held their interest and enthusiasm.

The research working group is continuing to do a great job running the webinar series, and this month we had interesting talks from Cloud and CWL experts Brian O'Connor and Michael Crusoe. Both speakers attended the Cloud hackathon. The User Support WG did not report in this issue as they are busy restructuring their membership and goals. We look forward to hearing about their plans in the next newsletter. The infrastructure WG is pleased with the appointment of a new data manager to work on the H3Africa archive, so we look forward to progress there and submission of the first H3Africa dataset to the EGA.

I invite you to continue reading to find out more about these activities, as well as the development of the eGenomics catalogue in this newsletter!

Education and Training

Education and Training
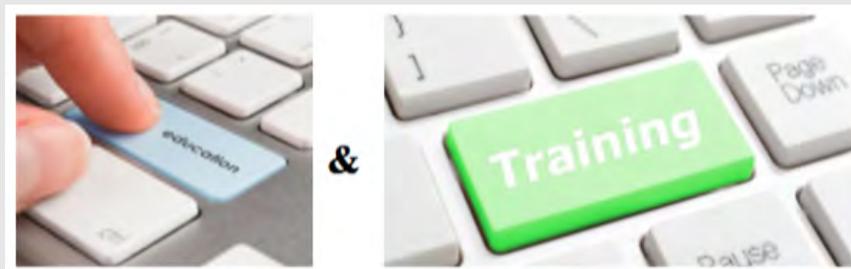
With the successful launch of the online Introduction to Bioinformatics course last month, the working group has turned its focus to the exciting workshops that have been planned for the upcoming months. In line with H3ABioNet moving away from running traditional training workshops, the upcoming series of training events are being organised as hackathons. It is wonderful that H3ABioNet's Infrastructure and Research working groups are getting involved in providing training through these hackathons (see below and in the relevant working group updates). For those of you unfamiliar with the term, a hackathon is an event where a group of experienced people, usually computer programmers and bioinformaticians, get together to develop and implement a solution to a particular problem.

The first workshop is a Cloud Computing Hackathon, which was hosted at the University of Pretoria from the 22nd - 26th August 2016. This was organised and run by H3ABioNet's Infrastructure working group. The hackathon was aimed at developing solutions to make important bioinformatics analysis pipelines available as *Docker* containers to be deployed on the cloud infrastructure. Have a look at the Infrastructure working group update in this month's issue of the H3ABioNet newsletter for more details.

H3ABioNet's Research working group, together with IBM Research Africa and the University of Notre Dame is presenting the second hackathon in September at the IBM Research Labs at the Tshimologong Precinct Technology Hub in Braamfontein, Johannesburg. The hackathon is aimed at addressing a research question that forms part of a DREAM challenge on malaria drug resistance. Read more about the hackathon in the Research working group update, in this month's issue of the H3ABioNet newsletter. The hackathon promises to be a great first step for H3ABioNet to participate in our first DREAM challenge.

The Education and Training working group has reviewed the milestones and deliverables for year 5 of the project and is currently assessing the need for hosting a metagenomics data analysis workshop with one of the West African nodes later in the year.

Finally, to assess the impact of H3ABioNet training on the larger H3Africa community, we have sent out our 6-month training impact survey to all individuals who have participated in an H3ABioNet training event to date. We have received a far amount of responses and will have some feedback on the results soon.

Nicky Mulder and Shaun Aron

Back to Contents

Infrastructure

**H3ABioNet**

Infrastructure

Good news for the ISWG is appointment of Ziyaad Parker as a data manager for H3ABioNet (see photograph to the left). In this position he will be primarily responsible for developing the database archive, which will be used as a repository for all submissions of data by H3Africa groups to EGA. With support of the Data Management Task Force, he will also help develop tools that allow users to search this data in order to plan future projects. Ziyaad is a graduate of UCT and brings valuable skills that will allow us to achieve this mission critical project. We welcome Ziyaad to H3ABioNet!

We ran a very successful Cloud Hackathon at the University of Pretoria from 22 to 26 August 2016 (see select photos on the next page). The hackathon was attended by participants from several H3ABioNet groups, as well as some US collaborators. The goal of the hackathon was to design four pipelines that H3Africa groups can use in Bioinformatics analyses. We explored the use of workflow languages/systems (CWL and Nextflow) and Docker so as to make the pipelines scalable and portable, and showed how these pipelines could run on range of different systems and cloud architectures. The four pipelines will deal with Genome-Wide Association Studies, Imputation, NGS and Metagenomics. All groups have a prototype pipeline ready and have a set of actions detailing what is needed to complete their pipelines. We plan to write a paper based on our experiences.

Issue 17: August 2016

Infrastructure
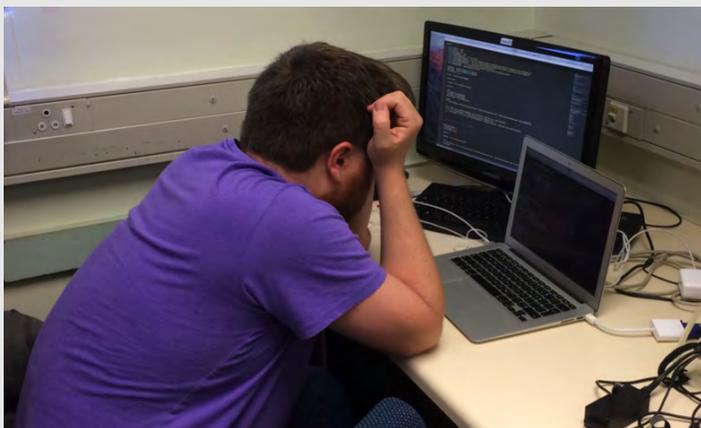
Hard at work at the cloud hackathon, University of Pretoria, South Africa (22-26 August 2016)

The deployment of eBiokits continues. Working with our collaborators at the Swedish University of Agricultural Sciences (SLU), new images and databases have been made available and we are starting their deployment. Our Globus Online and Netmap projects continue and three additional nodes have now been added to the list of participating sites.

Scott Hazelhurst and Suresh Maslamoney

Issue 17: August 2016

H3ABioNet

Research

Research

The H3ABionet webinars series covered two topics during the months of July and August: "Cloud computing and reproducible science" and "Microbial Bioinformatics" respectively. For the "Microbial Bioinformatics" webinars series, the first speaker was Dr. Brian O'Connor, the Technical Director of the UCSC Genomics Institute Analysis core. He gave an excellent talk on "Large-scale, cloud-based analysis of cancer genomes: lessons learned from the PCAWG project". The second speaker was Mr. Michael R. Crusoe who presented on "Portable and reproducible data analysis with the Common Workflow Language". Both speakers contributed to H3ABioNet's first hackathon on cloud computing, organised by the H3ABioNet Infrastructure working group, described above.
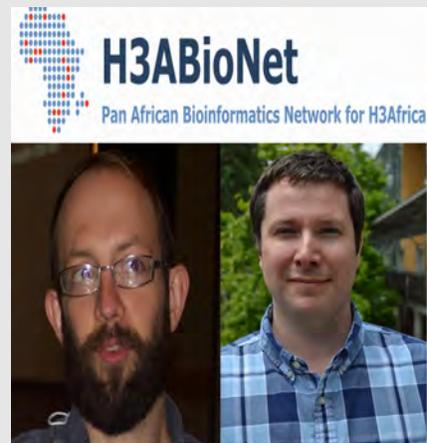
#H3ABionet webinar about #cloudcomputing
Twitter Advertisement for August Webinar posted through H3ABioNet social media channels

The H3ABioNet Research Proposals Sub-committee was established and had their first online meeting. The Sub-committee is now ready to receive research proposals. The objective of this sub-committee is to improve funding possibilities for graduate students and researchers as their proposals will go through internal peer review, within H3ABioNet, before they are submitted. These proposals will be sent to the relevant expert volunteers who signed up as reviewers.

H3ABioNet

Research

The Research working group is working on drafting and finalizing the deliverables, milestones and list of products for the fifth and last year of the current H3ABioNet grant. This has been discussed with working group members during the last meeting. The working group is also currently working on identifying a set of scoring metrics and performance indicators for comprehensive evaluation of the H3ABioNet webinars. The chairs of the working group and the H3ABioNet webinars coordinating team are planning on writing a full manuscript entitled "H3ABioNet Webinar Series: Challenges, Experiences and Lessons learned" that could be submitted for publication early next year. The paper will by then cover two years since we started the webinars series (January 2015 to December 2016) and include a list of metrics and performance indicators that should support us in assessing the seminars and speakers performance, especially for graduate students who presented their on-going research during this period. Both ex-moderators of the seminar series for 2015 will also be involved. The H3ABioNet Webinar Series manuscript will be opened to other members of the H3ABioNet consortium for whoever wants to contribute to its development.

We are organizing a joint H3ABioNet, University of Notre Dame and IBM Research Africa Hackathon on Malaria Drug Resistance for a *DREAM Challenge* during the period 6th-10th September 2016 in Johannesburg, South Africa. Ten participants have been selected from the different H3ABioNet nodes and will be working in teams to prepare preliminary datasets for a future open collaborative challenge on drug combination prediction. Representatives from Notre Dame University and IBM research Africa will be working with the different teams. Further information is available *here*.

Faisal Fadlemola and Amel Ghouila

Back to Contents

H3ABioNet

Internships at University of Illinois

Internships at University of Illinois, two productive summers

Every summer over the last four years, the H3ABioNet node at the University of Illinois at Urbana-Champaign (UIUC) has hosted interns from other nodes: Ibrahim Maman Laminou (CERMES, Niger), Mohamed Alibi (IPT, Tunisia), Deborah Fasesan (NABDA, Nigeria) and Azza Ahmed (U of Khartoum, Sudan). The hosting units at UIUC have been the Carl R. Woese Institute for Genomic Biology (IGB) and the National Center for Supercomputing Applications (NCSA). We require that scientists applying for an internship with us define in some detail what the scientific problem is that they want to work on, what data they propose to analyze, and what skills they hope to bring back to their home institution. For the last couple of years, we have also required that the internship should contribute directly to the ability of their home node to participate in a Node Accreditation Exercise.

Debbie Fasesan came to UIUC to work on the microbiome of leg ulcers from sickle cell anemia patients. Because of poor circulation, these patients often develop painful, and difficult to treat, leg ulcers, which can get infected by environmental pathogens. The goal of the study was to examine the microbial populations present in different parts of the ulcers (skin and deep tissue), and the effects of the compression therapy commonly used to treat them. DNA extracted from samples had been sent to UIUC ahead of Debbie's arrival. She spent her first two weeks in the lab, looking on as the samples were PCR amplified using multiple primer pairs in a Fluidigm Access Array, sequencing libraries were prepared, and sequencing was performed on an Illumina MiSeq machine. She spent the rest of her time analyzing the sequence data: QC, demultiplexing, gathering basic statistics, making multiple sequence alignments, determining OTUs in the samples, using these to document alpha (within sample) and beta (between samples) diversity values, and relating the results to biologically relevant phenomena. Most of the work was performed using the *QIIME* (Quantitative Insights Into Microbial Ecology) software. Debbie and the NABDA group are working on a paper describing the results, and on installing QIIME so that NABDA can take the 16S rDNA analysis Assessment Exercise.

Azza Ahmed has a background in computer science, and came to UIUC to gain a deep understanding of genomic variant calling workflows. Her primary project was to explore how changes in the parameters used to call crucial steps in the best practices workflow - such as: the initial alignment of reads, the realignment / recalibration step, the haplotype calling, and the joint genotyping of multiple individuals - influenced the ability to detect variants that are rare but significant in large populations. This question is of great interest for groups that look for rare causative variants in large GWAS studies that rely on exome or whole genome sequencing. During her first week at UIUC, she attended the *Computational Genomics course* organized with the Mayo Clinic. She then worked with members of the HPCBio group at IGB, *trainees from Fisk University*, and students from the *SPIN program* at NCSA. Azza produced a comprehensive map of all of the software components of a standard variant calling workflow and their respective adjustable parameters, and measured the effects of varying their values on the most crucial components. She is now in a perfect position to lead an effort in Khartoum where her node will take the Human Variant Calling Node Accreditation Exercise.

Both of the trainees were mentored closely by members of the HPCBio group, and took quizzes to assess how well they were prepared, and how well they assimilated the new material they were exposed to. We strongly believe that their time at UIUC will help them substantially in achieving their future career goals, and in empowering their home institutions, and African scientists, to take on computational challenges in the future.

Victor Jongeneel

Issue 17: August 2016

H3ABioNet

Internships at University of Illinois

In their own words…

1. What's the most useful thing you learnt during your internship at University of Illinois?

Deborah Fasesan: *I learnt the necessary skills in microbiome data analytics required to further my career, my node's activities and my workplace in general.*

Azza Ahmed: *The power of collaboration and importance of seeking help when needed.*

2. What was the biggest challenge of the internship?

Deborah Fasesan: *To prove my worth to myself within the frame of time scheduled for me - this included understanding my work and learning all necessary skills in 2 months.*

Azza Ahmed: *It was very smooth overall. Everyone was welcoming and very keen on helping however they could, which made it easier to adapt.*

3. What did you miss most while you were away from home?

Deborah Fasesan: *My relatives, friends and colleagues.*

Azza Ahmed: *Calling home and being in touch with my family regularly, as there was an 8 hour time difference. We talked almost only on weekends, as otherwise I would be either in the University, or our timings would be totally off.*

Deborah Fasesan presenting her findings to HPCBio group

Azza Ahmed

Issue 17: August 2016

H3ABioNet

eGenomics catalogue

eGenomics Catalogue

The rapid adoption of online learning has led to an increase in free online educational material and to the development of online aggregators for such material. However, most aggregators are not subject specific but are dedicated to the type of material e.g. Massive Open Online Course (MOOC), such as Coursera. Therefore, searching for learning/teaching online material in a specific subject area can be time consuming. Yet another major pitiful in using such freely available educational material is the lack or peer-reviewing and quality control.

To our knowledge only one aggregator exists which is domain/subject specific and is a directory to material that have been critically reviewed by a team of faculty, students and other professionals; synDRME (A synthesis of Digital Resources for Medical Education). Such domain/subject specific aggregators could be invaluable in resource limited areas to promote training and education in scarce skills areas, such as genomics.

We present the *eGenomics catalogue*, an aggregator of online free educational material relevant to genomics. Unlike synDRME, which depends on a small pool of volunteers to source and review material, eGenomics aims to be community-driven and relies on volunteers. Posting of reviews/material is entirely voluntary and the information or resources loaded onto the catalogue are personal views not endorsed by anyone.

The eGenomics catalogue was initiated by H3ABioNet by a small team of volunteers (Nicky Mulder, Vicky Nembaware, Jean-Baka Domelevo Entfellner, Lerato Magosi, Shaun Aron and Sumir Panji) with Jean-Michel Serufuri as the developer. The H3Africa Fellows' Club and some members of the Education Coordinated Working Group (especially Maryam Bibi and Barrington Burnett) voluntarily helped in designing the structure of this catalogue. The catalogue is now accessible online where the public can view and search for material and registered users can perform the following functions:

1. Upload links to free online educational material and write brief descriptions which include:
   - Types of material (Books, MOOCs, Lecture/Webinars, Databases, Software).
   - Subject area (Material is categorized into one or more of the H3Africa subjects and/or EDAM ontology Bioinformatics topics).
   - Source (Author, Institute)
2. Search for material and access links
3. Write personal reviews of material
4. Compare existing descriptions and reviews between pairs of online material

Screenshot of *eGenomics catalogue* home page

The eGenomics catalogue is fully functional and volunteers are free to use it. This aggregator has been advertised to learners in the Introduction to Bioinformatics Course (IBT) and will also be used by trainers and learners in the African Genomics Medicine Curriculum and Training Initiative. The eGenomics catalogue will be officially launched via social media platforms before the end of September 2016.

Vicky Nembaware

Back to Contents

Issue 17: August 2016

![H3ABioNet logo]

H3ABioNet

Introduction to Bioinformatics course (IBT)

Distance-based online Bioinformatics training in Africa: the H3ABioNet experience

Africa is not unique in its need for basic Bioinformatics training for individuals from a Molecular Biology background. However, unique logistical challenges must be addressed in order to meet this need on the continent. Most notable challenges include: access to comprehensive, local Bioinformatics expertise; access to Bioinformatics training; and Internet stability. A taskforce from the Education and Training working group, has therefore developed an innovative, free introduction to Bioinformatics course (IBT), taking these challenges into account.

A distance-based learning model has been selected for this 3 month course (mid July - beginning October 2016) in order to increase access to expert African and European Bioinformatics trainers. While developed for a resource limited setting, IBT could easily be adapted to other settings. The course covers several Bioinformatics topics, namely: Databases and Resources; Genomics; Linux; Sequence Alignment; Multiple Sequence Alignment; Protein Structure Bioinformatics; and Molecular Evolution and Phylogenetics. Classrooms with a total of >350 enrolled participants are hosted at 19 institutions (mostly H3ABioNet nodes, but others have been welcomed), across 11 African countries. The classrooms provide local administrative and academic support - read more about the classrooms for 2016 *here*.

**Faces of the IBT course - a few 'class of 2016' photos**

H3ABioNet

Introduction to Bioinformatics course (IBT)

Classroom selection was based on certain infrastructure criteria, including: computer resources; Internet access; and availability of local teaching assistants. To ensure that classroom success does not rely on stable Internet, classrooms watch pre-recorded and pre-downloaded lecture videos, as well as work through practical assignments on the lecture content, during biweekly contact sessions. Lecture recordings, lecture slides, and practical assignments - for each session - are publicly available on the *IBT course website* under creative commons license. We encourage anyone with an interest in Bioinformatics training materials to access, use, and share the resources. In fact, it appears that the materials are being accessed by people from all over the world already (see below)!

Total number of website sessions* per country between 13 July and 10 August 2016 - people from all over the world are accessing these open access Bioinformatics materials!*login not required. Data captured, and images, generated by Google Analytics

In addition to locally run contact sessions, our trainers are available via the Mconf video conferencing system to take questions from the classrooms during the contact session time - a 'virtual classroom'. Further online 'question and discussion' forums are also available. These are hosted on the course management platform, Vula. Participants, trainers and teaching assistants from the various classrooms are able to engage in these forums and discuss topics covered in the sessions, or any other Bioinformatics-related topics that they are interested in. Assignments and assessments are also administered, submitted by participants, and graded, via Vula.
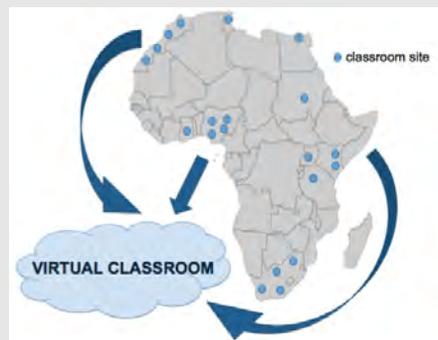
Classrooms meet each other virtually during biweekly contact sessions

IBT really has been a collaborative effort by the *over 70 volunteer staff* - we thank the IBT core team, expert trainers, expert consultants (we did not try and reinvent the wheel with this course!) and classroom staff (PIs, teaching assistants, and system administrators) who have helped to make, and continue to make, the IBT course a success! We wish all classrooms well with the final month of the course!

Kim Gurwitz, IBT project coordinator

Announcements

- Congratulation to Amel Ghouila from IPT, Tunisia (Co-Chair of the H3ABioNet Research Working Group) who received the best Regional Ambassador award of the *Technovation international competition* that she is leading in Tunisia. Technovation aims to close the gender gap for women in technology and business. Over the course of the program young women all over the world will be creating solutions to issues in their community - no longer being just consumers of technology, but creators.

- Congratulations to PhD student Ms. Chaimae Samtal from the University of Mohammed First working on the "Genomics Stratification of Prostate Cancer in African population" who won the Moroccan National Competition "My Thesis in 3 minutes" for the 2016 edition. She will be representing Morocco at the International competition on the 29th of September 2016 to be organized in Rabat Morocco. The project was started at the last H3ABioNet AGM meeting as sub-network between Prof. Ghazal, Prof. Fadlelmola and Prof. Oyekanmi.

- Congratulations to Bouchra Chaouni from the University of Mohammed First, the former beneficiary of an H3ABioNet internship, who is the winner of the prestigious Fulbright fellowship to spend one year at the Marine Biological Laboratory of the Brow University, MA, USA. Bouchra will be comparing the marine microbiome of African and American oceanic costs using bioinformatics tools.

- Congratulations to Judith Kumuthini, Mamana Mbiyavanga and Emile Chimusa for their recent publication "Minimum information required for a DMET experiment reporting"; Pharmacogenomics. 2016 Sep;17(14):1533-1545 (PubMedID: *27548815*).

Any announcements for the next edition of the H3ABioNet newsletter may be sent to:

kim.gurwitz@uct.ac.za AND sumir.panji@uct.ac.za

Back to Contents

H3ABioNet

Upcoming Events

Upcoming Events

- **July 6th to October 9th 2016**: the *Introduction to Bioinformatics* online course enters its final month. Participants at the course's 20 classroom sites look forward to modules on Protein Structure Bioinformatics and Phylogenetics.

- **September 5th to September 10th 2016**: *H3ABioNet / IBM Research Africa Hackathon* on Malaria Drug Resistance for a DREAM challenge, Johannesburg, South Africa (http://www.h3abionet.org/h3abionet-ibm-research-africa-hackathon-malaria-drug-resistance-2016)

- **September 20th to September 23rd 2016**: Joint 2016 conference of the South African Societies for Bioinformatics and South African Genetics Society, Durban South Africa.

- **3rd to 5th October 2016**; *The International Conferences on Research Infrastructures* (ICRI) that provides a global forum in the Research Infrastructures domain, Cape Town, South Africa.

- **October 3rd to October 14th 2016**: INDA Hands-on NGS-Statistics course, Dakar, Senegal.

- **19th to 22nd October 2016**: *International Conference on Rare Diseases and Orphan Drugs* (ICORD) for 2016, RareX has an extended program, to incorporate patients, their families and caregivers as well as support groups and minority groups within the rare disease community, Stellenbosch, South Africa

- **October 27th to October 31st 2016**: *Ninth H3Africa consortium meeting*, Mauritius.

- **November 1st to November 4th 2016**: H3ABioNet SAB and AGM, Cape Town, South Africa.

- **November 6th to November 10th 2016**: *Ninth Annual RECOMB/ISCB Conference* on Regulatory and Systems Genomics, with DREAM Challenges and Cytoscape Workshop, Phoenix, Arizona, USA.

- **November 21st to November 23rd 2016**: *The Fourth International Society for Computational Biology Latin America Bioinformatics Conference* (ISCB-LA), jointly organized with A2B2C, will take place in Buenos Aires, Argentina.

Any upcoming events for the next edition of the H3ABioNet newsletter may be sent to:

kim.gurwitz@uct.ac.za AND sumir.panji@uct.ac.za

Back to Contents

Issue 17: August 2016

H3ABioNet

Upcoming H3ABioNet working group meeting schedule

Upcoming H3ABioNet working group meeting schedule*

*Schedule until end September 2016

Summary of H3ABioNet upcoming working group meetings

Month	Date	Day	Working Group (WG)	Time (UTC)
September	9th	Friday	Research WG meeting	11:00
September	16th	Friday	Infrastructure WG meeting	13:00
September	22nd	Thursday	H3ABioNet Seminar Series	13:00
September	23rd	Friday	User Support WG meeting	9:00
September	27th	Tuesday	Education and Training WG meeting	11:00

Timezone conversions to UTC for all H3ABioNet working group meetings

UTC Time Offset	Time Zone Name	Region/ Country in the Time Zone offset
-5 hours	CDT	Chicago, USA
0 hours	GMT	Burkina Faso, Ghana, Mali, Morocco, Senegal
+1 hour	WAT	Cameroon, Chad, Gabon, Namibia, Nigeria, Niger, Tunisia
+2 hours	CAT	Botswana, Egypt, Malawi, South Africa, Sudan, Zambia
+3 hours	WAT	Ethiopia, Kenya,Tanzania, Uganda

This edition of the newsletter was compiled and edited by Kim Gurwitz. For any corrections, please contact Kim at kim.gurwitz@uct.ac.za

Back to Contents