A. Bioinformatics in Africa (sorted by countries):

Burkina Faso

1. Institutions:

1.1 WHO Onchocerciasis Control Programme in West Africa:

The Multi Disease Surveillance Centre of Ouagadougou (MDSC) was created in 2002 at the end of the Onchocerciasis Control Program (OCP), in order to preserve the gains achieved in onchocerciasis control, while expanding acquired experience to the control of other priority diseases such as meningitis, malaria, and HIV/AIDS using the existing infrastructure. The MDSC is created to serve as a centre of excellence for disease surveillance in the WHO Africa Region. Established in the premises of the Onchocerciasis Control Programme (OCP), the MDSC inherited the facilities of OCP, including the buildings, the laboratories, the library, and telecommunication equipment.

The MDSC supports the WHO African Region countries in the surveillance of Onchocerciasis and other communicable diseases such as Meningitis, HIV/AIDS/TB and Malaria. As the Centre is strengthened, the range of targeted diseases will be expanded and further include Non Communicable Diseases. The core activities of the MDSC includes compilation and in depth analysis of surveillance data, using the information generated to establish disease trends and to predict epidemics. Particular emphasis is laid on integration of laboratory into disease surveillance. The Centre offers to WHO Member States of the region, facilities and mechanisms for training of epidemiologists, entomologists, data managers and laboratorians, as well as in laboratory quality assurance and quality control, and exchange of key epidemiological information.

The MDSC establishes linkages with the countries of the region, and with institutions and partners involved in disease surveillance. In order to allow partners and countries participation in the management/governance of the MDSC, a Steering Committee and a Scientific Advisory Committee are created.

The mission of the Multi Disease Surveillance Centre is to provide high level technical support to the WHO/African Region Member States in the surveillance of Onchocerciasis, Meningitis, HIV/AIDS, Malaria, Tuberculosis, as well as any other priority diseases assigned to it by Member States.

The goal of this proposal is provide a comprehensive description of the objectives, activities and indicators that would result in the evolution of MDSC into a Reference Centre for disease surveillance in the African Region through Epidemiological, Entomological, Laboratory, Data management, Mapping and Documentation support to Member States in the surveillance of priority diseases.

2. Training Activities:

2.1 Long-term training activities:

No long-term activities in bioinformatics.

2.2 Short-term training activities:

No long-term activities in bioinformatics.

3. Research Interest and Activities:

The Centre undertakes research activities on major communicable diseases in order to improve their prevention and control. The Centre establishes scientific collaboration with other research institutions for some research activities. The Centre issues publications as appropriate on the outcome of the researches.

To be able to keep abreast with the fast changing pace of science, the Centre undertakes research programmes to discover new materials, new methods and develop new strategies for surveillance and interventions. The new techniques and equipment being developed need to be introduced into the country surveillance system and therefore regular in-service training have to be organised to

upgrade the skills of surveillance and laboratory staff in the Region.
To this date, the centre has initiated/participated with WHO collaborating centres in many research themes including:

- Meningitis epidemics and Environmental factors in collaboration
- Introduction of a new Conjugate A vaccine
- Feasibility of Onchocerciasis elimination by Ivermectin distribution
- *Simulium sirbanum* migration from Sierra Leone to Mali river basins, and its influence on onchocerciasis epidemiology
- Ivermectine resistance in O.volvulus parasites (Using bioinformatics and applied genomics for identifying targets for drugs, diagnostics; understanding ivermectine resistance mechanisms)
- New diagnostic tools of onchoceriasis detection (Ivermectin Resistance Product Development Team (IRPDT /WHO/TDR)

4. Expected feedback from the CNRS-NEPAD partnership:

1. Establish strong teaching courses for applied public health training (skill specific and Field Epidemiology and Laboratory Training) to improve human capacity for epidemiology, entomologic surveillance, laboratory, and data management services in countries in the region.
2. Sending and receiving Post-doctoral fellows
3. Onchocerciasis, meningitis genome data analysis and other priority diseases pathogens genome using bioinformatics tools.
4. Funding research collaborations with African laboratories involved in bioinformatics
5. Develop and organize short bioinformatics courses.

5. Web site and links:
No website.

6. Publications:
No publications in bioinformatics.

7. Contacts:
Sounkalo Moussa SANFO
WHO Onchocerciasis Control Programme in West Africa
Ouagadougou
Burkina Faso
sysals@yahoo.fr

Côte d'Ivoire

1. Institutions:

1.1 Institut Pasteur de Côte d'Ivoire (IPCI):
The Institut Pasteur de Côte d'Ivoire was created by the N°72-511 law of July 27*th*, 1972 under the Presidency of his Excellency Felix Houphouët Boigny and Professor Jacques Monod, then Leading of the Pasteur institute of Paris.
Placed under the administrative tutelage of the Ministry of the Scientific Research and the financial tutelage of the Ministry of the economy Ministry of Finance, the IPCI is erected following the N°91-645 decree of October 09*th* 1991 public establishment of industrial and commercial nature.
Geographically the Pasteur institute includes 2 sites:

- Adiopodoumé (25km of the economic capital).

- Cocody near the Hospital University Centre, Abidjan

Traditionally, the IPCI has for mission (Decree N°91-645):

- The survey of the man's transferable and immune illnesses
- The prevention and the prophylaxis of these illnesses in collaboration with the services of health and all users of medical research products.
- The survey and the research of the usable natural substances to end medical and pharmaceutical
- Integration and researchers, technicians training.
- The supply of research benefits, medical analyses, exams and advice to the profit of public and private outside partners.

In order to achieve the mission that is assigned concerning research, the Institut Pasteur is organized in Departments and Units of research regrouped around a big technical and technological platform of which the one of molecular Biology consisting lately of a pole of Cytometric of flux and a pole of Bioinformatics recently opened. The pole of bioinformatics is a part of a Molecular Microbiology unit.

The objectives of the Bioinformatics pole are:

- Encourage the access to innovations in research and the best exploitation of research data management.
- Develop the critical spirit of the researchers around their axis of research.
- Bring an active help to the improvement of the public health while having for constant worries to feed it by research and the innovation.
- Start training researchers to use bioinformatics as an indispensable tool to research.
- Encourage interdisciplinary creating a network of scientific information available to the researchers of the institute and partners.

2. Training Activities:

2.1 Long-term training activities:

- To integrate the bioinformatics in the training programs of academic scientists.
- Development of West Africa Centre training in bioinformatics and genome data analysis.

2.2 Short-term training activities:

The IPCI will organize in the month of May 2007 a yearly regional course of initiation and using genome data analysis with the participation of the Centre Biotechnologique de Sfax, Tunisia (Pr Ahmed Rebaï, Department of bioinformatics).

3. Research Interest and Activities:

The following are the present axes of research of the IPCI:

- HIV and sexual transferable infections
- Tuberculosis and respiratory illnesses
- Malaria
- Buruli Ulcer
- Hemorrhagic fevers and arbovirus
- Support to the public health (epidemiological surveillance)

The following are the present research activities at the IPCI:

- Characterization of virulence genes microorganisms
- Typing and molecular epidemiology

- Microsatellites
- Sequencing
- Cytométric of flux
- Therapeutic and vaccines assays
- Genomic data analysis
- Development of specific programs of genome analysis

4. Expected feedback from the CNRS-NEPAD partnership:

- Initiation and continuing education of the local and regional researchers to the bioinformatics
- Establish an African bioinformatics network to permit scientific exchanges
- Facilitate the access to the scientific information to the graduate students starting a career of research.
- Contribute to the development of the NTIC in western Africa in the domain of research.
- To bestow some scholarships in bioinformatics for the young researchers
- Development of trans-disciplinary research project.
- Backing of the capacities in informatics

5. Web site and links:
www.pasteurci.com

6. Publications:

1. N'Guessan K, Bakayoko S, **Ekaza E**, Aka N, Bouzid SA, Kouabenan H, Kouame-Elogne C, Gbonon V, Guessend N, Kacou-N'Douba A, M'Bengue-Kouassi A, Kouassi K, Kouakou K, **Dosso M** (2003) Protein immunological characteristics exploration from people, with Ulcer of Buruli disease cas of thirty patience who are living in lake region in Côte d'Ivoire Africa Chemist**.** *Pharmacien d'Afrique* **166**:3-6.
2. Köhler S, **Ekaza E**, Paquet JY, Walravens K, Teyssier J, Godfroid J, Liautard JP (2002) Induction of *dnaK* through Its Native Heat Shock Promoter Is Necessary for Intramacrophagic Replication of *Brucella suis*. *Infect. Immun.* **70**:1631-1634.
3. **Ekaza E**, Teyssier J, Ouahrani-Bettache S, Liautard JP, Kohler S (2001) Characterisation of *Brucella suis clpB* and *clpAB* Mutants and Participation of the Genes in Stress Responses. *J. Bacteriol.* **183**:2677-2681.
4. Bibollet-Ruche F, Peeters M, Mboup S, **Ekaza E**, Gandji R, Torimiro J, Mpoudi EN, Amblard J, Dibanga G, Saidou M, Esu-Williams E, Vanden Haesevelde M, Saman E, Delaporte E (1998) Molecular Characterization of the Envelope Transmembrane Glycoprotein of 13 New Human Immunodeficiency Virus Type 1 Group O Strains from Six Different African Countries. *AIDS Res. Hum. Retroviruses* **14**:1281-1285.

7. Contacts:

Dr EKAZA Euloge
Unité de Microbiologie Moléculaire
Institut Pasteur de Côte d'Ivoire
01 BP 490 Abidjan 01, Côte d'Ivoire
Phone: (225) 05546761
Fax: (225) 23457623
ekazae@yahoo.fr

Kenya

1. Institutions:

1.1 International Livestock Research Institute (ILRI):

Under a NEPAD initiative, the Biosciences Eastern and Central Africa (BECA) (www.biosciencesafrica.org) was established at ILRI. BECA consists of a hub, regional nodes, and other affiliated laboratories and partner institutes. A state of the art joint Bioinformatics Platform (www.becabioinfo.org), whose overall goal is to provide a coherent and powerful bioinformatics infrastructure for use by all scientists in East and central Africa. The Platform goal requires both physical and intellectual developments that together provide researchers with access to diverse infrastructure in a wide-area network, thereby addressing four important aspects of bioinformatics:

1) Science: bioinformatics tools for data integration and visualization, standardization of data formats and data analysis strategies, and distribution of analysis tasks over local- and wide-area networks are in development;
2) Bioinformatics Support Facility: provides assistance and custom programming to projects and those unable to establish a bioinformatics support function intrinsic to their project due to shortage of qualified personnel or lack of funding;
3) Hardware Platform: provide a powerful high performance computing platform capable of handling the largest analysis needs for projects;
4) Bioinformatics Training for East and central African scientists: While many Web-based tools are available to the wet-lab researcher, the Web is not well suited for tasks beyond single-sequence annotation. Researchers need to become productive in a server-based Unix environment with its wealth of scripting and automation tools. Even at an entry-level, this can be an intimidating task if proper guidance is not available.

1.2 International Centre of Insect Physiology and Ecology (ICIPE):

ICIPE's research focus is on insect biology, in order to improve the wellbeing of the peoples of the tropics through insect science. There is a commitment to utilise contemporary science in order to limit the impact of disease vectors, and agricultural pests. The understanding of the mechanisms associated with behaviour (e.g. attraction and repellency) is crucial. ICIPE seeks to enhance its bioinformatics capacity in order to support data from various EST projects designed to gain insights into the insect ecology and plant pathogen interactions though studies of metabolic pathways associated with production of all elochemicals.

2. Training Activities:

2.1 Long-term training activities:

Kenyatta University: An introductory course in Bioinformatics is offers to MSc Biotechnology students. This comprises of 35 hours of lectures and practicals.

University of Nairobi: A centre for Biotechnology and Bioinformatics (CEBIB), which will offer postgraduate training (diplomas, MSc and PhD) in areas of biotechnology and bioinformatics has recently been launched.

Other universities in Kenya, including Egerton, Maseno and the Jomo Kenyatta University of Agriculture and Technology offer introductory courses to undergraduates in biomedical sciences.

In addition, under the BECA platform MSc and PhD fellowships are being made available for Bioinformatics students. ILRI is forging links with Universities in South Africa and the United Kingdom to provide access to courses and training material.

2.2 Short-term training activities:

The Biochemical Society of Kenya (BSK) runs a 3-day training workshop that introduces participants to basic bioinformatics applications. This workshop is open to residents of countries in the eastern African region, and has continued to draw support from facilitators from various countries.

ILRI is in the process of becoming a specialist node of EMBnet, which will give it access to European funding and expertise in training of bioinformatics students.

3. Research Interest and Activities:

The following are the present areas of research interest:

1. EST clustering
2. Genome sequencing and annotation
3. Functional genomics and proteomics (including key tropical pathogens)
4. Structural bioinformatics
5. Development of Bioinformatics Data Management Systems
6. Gene Mining
7. High Throughput Genotyping
8. Microarray data management and analysis
9. Metagenomics
10. Immunoinformatics
11. Host-pathogen interaction
12. High performance computing and grid development
13. Parasite transfection technologies
14. Cell cycle regulation
15. Population genetics
16. Vector genomics
17. Drug, vaccine and diagnostic target discovery

4. Expected feedback from the CNRS-NEPAD partnership:

1. Establish strong teaching courses for M.Sc and PhD students in local Universities or Institutes. Diverting postgraduates abroad could weaken rather than strengthen programmes in the region.
2. Strengthen collaborations within Africa, and take advantage of regional initiatives, such as the BECA platform, to create an African network of excellence in the field of bioinformatics.
3. Facilitate establishment of a fast Internet backbone for research and education in Africa.
4. Access to training opportunities for M.Sc and PhD students, and for short-term attachments.
5. Sending and receiving Post-doctoral fellows
6. Establishing regular African short courses in bioinformatics and genome data analysis.
7. Funding research collaborations with African laboratories involved in bioinformatics and genome data analysis.
8. Establishing joint research projects.

5. Web site and links:

http://www.ilri.cgiar.org/

http://www.icipe.org/

http://www.uonbi.ac.ke/cebib

6. Publications:

1. Simon P. Graham, Roger Pellé, Yoshikazu Honda, Duncan M. Mwangi, Nyerhovwo J. Tonukari, Mat Yamage, E. Jane Glew, **Etienne P. de Villiers**, **Trushar Shah**, Richard Bishop, Evelyne Abuya,

Elias Awino, James Gachanja, Anthony E. Luyai, Ferdinand Mbwika, Anthony M. Muthiani, David M. Ndegwa Moses Njahira, John K. Nyanjui, Fredrick O. Onono, Julius Osaso, Rosemary M. Saya, Claude Wildmann, Claire M. Fraser, Ian Maudlin, Malcolm J. Gardner, Subhash P. Morzaria, Sheena Loosmore Sarah C. Gilbert, Jean-Christophe Audonnet, Pierre van der Bruggen, Vishvanath Nene & Evans L. N. Taracha.. (2006) Theileria parva candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proceedings of the National Academy of Sciences* **103:**3286-3291.

2. **Trushar Shah**, **Etienne de Villiers**, Vishvanath Nene, Brian Hass, Evans Taracha, Malcolm J Gardner, Clare Sansom, Roger Pelle, Richard Bishop. (2005) Using the transcriptome to annotate the genome revisited: application of massively parallel signature sequencing (MPSS). *Gene* **17:**366:104-108.
3. Richard Bishop, **Trushar Shah**, Roger Pelle, David Hoyle, Terry Pearson, Lee Haines, Andrew Brass, Helen Hulme, Simon, P. Graham, Evans L. N. Taracha, Charles Lu, Brain Hass, Jennifer Wortman, Owen White, Malcolm J. Gardner, Vishvanath Nene and **Etienne P. de Villiers**. (2005) Analysis of the transcriptome of the protozoan Theileria parva using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. *Nucleic Acids Research* **25:**5503-5511.
4. Malcolm J. Gardner, Richard Bishop, **Trushar Shah**, **Etienne de Villiers**, Jane M. Carlton, Neil Hall, Qinghu Ren, Ian T. Paulsen, Arnab Pain, Matthew Berriman, Robert J. M. Wilson, Shigeharu Sato, Stuart A. Ralph, David J. Mann, Zikai Xiong, Shamira J. Shallom, Janice Weidman, Lingxia Jiang, Jeffery Lynn, Bruce Weaver, Azadeh Shoaibi, Delia Wasawo, Jonathan Crabtree, Jennifer R. Wortman, Brian Haas, Samuel Angiuoli, Todd Creasy, Charles Lu, Bernard Suh, Joanna Silva, Teresa Utterback, Tamara Feldblyum, Miheala Pertea, Jonathan Allen, Evans Taracha, Steven L. Salzberg, Owen White, Henry A. Fitzhugh, Subhash Morzaria, J. Craig Venter, Claire M. Fraser, Vishvanath Nene. (2005) Theileria parva: genome sequence of a bovine pathogen causing a lymphoproliferative disease. *Science* **309:**134-137.
5. Arnab Pain, Hubert Renauld, Matthew Berriman, Lee Murphy, Corin A. Yeats, William Weir, Arnaud Kerhornou, Martin Aslett, Richard Bishop, Christiane Bouchier, Madeleine Cochet, Richard M. R. Coulson, Ann Cronin, **Etienne de Villiers**, Audrey Fraser, Nigel Fosker, Malcolm Gardner, Arlette Goble, Sam Griffiths-Jones, David E. Harris, Frank Katzer, Natasha Larke, Angela Lord, Pascal Maser, Sue McKellar, Paul Mooney, Fraser Morton, Vishvanath Nene, Susan O'Neil, Claire Price, Michael A. Quail, Ester Rabbinowitsch, Neil D. Rawlings, Simon Rutter, David Saunders, Kathy Seeger, **Trushar Shah**, Robert Squares, Steven Squares, Adrian Tivey, Alan R. Walker, John Woodward, Dirk A. E. Dobbelaere, Gordon Langsley, Marie-Adele Rajandream, Declan McKeever, Brian Shiels, Andrew Tait, Bart Barrell, Neil Hall.. (2005) The genome of the host-cell trans-forming parasite *Theileria annulata* and a comparison with *T. parva*. *Science* **309:**131-133.
6. B. Lambson, V. Nene, M. Obura, **T. Shah**, P. Pandit, O. Ole-MoiYoi, K. Delroux, S. Welburn, R. Skilton, **E. de Villiers** and R. Bishop. (2005) Identification of candidate sialome components expressed in tick salivary glands using secretion signal complementation in mammalian cells. *Insect Molecular Biology* **14:**403-414.
7. Nicola E. Collins, Junita Liebenberg, **Etienne P. de Villiers**, Kelly A. Brayton, Elmarié Louw, Alri Pretorius, F. Erika Faber, Henriette van Heerden, Antoinette Josemans, Mirinda van Kleef, Helena C. Steyn, M. Fransie van Strijp, Erich Zweygarth, Frans Jongejan, Jean Charles Maillard, David Berthier, Marli Botha, **Fourie Joubert**, Craig H. Corton, Nicholas R. Thomson, **Maria T. Allsopp**, and **Basil A. Allsopp**. (2005) The genome of the heartwater agent *Ehrlichia ruminantium* contains multiple tandem repeats of actively variable copy number. *Proceedings of the National Academy of Sciences* **102:** 838-843.
8. Abubakar LU, Bulimo WD, **Mulaa FJ**, Osir EO (2006) Molecular characterization of a tsetse fly midgut proteolytic lectin that mediates differentiation of African trypanosomes. *Insect Biochem Mol Biol.* **36**(4):344-52.
9. Kinyua JK, Nguu EK, **Mulaa F**, Ndung'u JM (2005) Immunization of rabbits with Glossina pallidipes tsetse fly midgut proteins: effects on the fly and trypanosome transmission. *Vaccine*. **23**(29):3824-8.
10. Hashim SO, Kaul RH, Andersson M, **Mulaa FJ**, Mattiasson B (2005) Differential scanning

calorimetric studies of a Bacillus halodurans alpha-amylase. *Biochim Biophys Acta.* **1723**(1-3):184-91.
11. Hashim SO, Delgado O, Hatti-Kaul R, **Mulaa FJ**, Mattiasson B (2004) Starch hydrolysing Bacillus halodurans isolates from a Kenyan soda lake. *Biotechnol Lett* **26**(10):823-8.
12. Baliraine FN, Bonizzoni M, Guglielmino CR, Osir EO, Lux SA, **Mulaa FJ**, Gomulski LM, Zheng L, Quilici S, Gasperi G, Malacrida AR (2004) Population genetics of the potentially invasive African fruit fly species, Ceratitisrosa and Ceratitis fasciventris (Diptera: Tephritidae).*Mol Ecol.* **13**(3):683-95.

7. Contacts:

Etienne DE VILLIERS
Bioinformatics Group Leader
International Livestock Research Institute (ILRI)
PO Box 30709, Nairobi 00100, Kenya
Phone: (254) 20 4223000
Fax: (254) 204223001
e.villiers@cgiar.org

Daniel MASIGA
Department of Biochemistry and Biotechnology
Kenyatta University
P.O. Box 43844
Nairobi 00100
Kenya
Tel: (254) 810901-12
dmasiga@icipe.org

Francis J. MULAA
Department of Biochemistry
University of Nairobi
P.O. Box 30197
Nairobi 00100
Kenya
Tel: (254) 204442534 or (254) 204441186
mulaafj@uonbi.ac.ke

Steven Ger NYANJOM
International Center of Insect Physiology and Ecology
P.O. Box 30772
Nairobi 00100
Kenya
Tel: (254) 8632000
snyanjom@icipe.org

Trushar SHAH
Bioinformatician
Biosciences eastern and central Africa
P.O. Box 30709,
Nairobi 00100
Kenya
tm.shah@africabiosciences.org

Mali

1. Institutions:

1.1 International Center for Excellence in Research (ICER):

The ICER is a research center composed of the following three programs:

1. The Malaria Research and Training Center - Parasitology Group,
2. The Malaria Research and Training Center - Entomology Group
3. The SEREFO Group.

The first two programs develop biomedical researches in malaria, Filariasis and Leishmaniasis. The third program develops biomedical researches in tuberculosis and HIV.

Bioinformatics was introduced recently to the ICER and is constantly growing. The ICER has one team, headed by Sidy SOUMARE, which supports the three programmes in all their needs in informatics and bioinformatics. This team can beneficiate from some computational facilities (4 blast servers, 15 other servers and around 200 terminals), but the ICER staff needs some training in order to be able to administrate these facilities.

2. Training Activities:

2.1 Long-term training activities:

No short-term training in bioinformatics.

2.2 Short-term training activities:

No long-term training in bioinformatics.

3. Research Interest and Activities:

The following are the present areas of research interest:

1. Functional genomics
2. Analysis of microarray data
3. Interaction between the vector and the parasite.

4. Expected feedback from the CNRS-NEPAD partnership:

1. Development of scientific collaborations into the key domains listed in the Areas of interest paragraph
2. Needs in training students and scientists in informatics and bioinformatics. Training students can be done through collaborations.
3. Write an African collaborative grant in order to support activities and various needs listed in the report. Each team can make a contribution (lab work, algorithm design, data analysis, etc...) and according to the members areas of expertise.

5. Web sites and links:

No website.

6. Publications:

No publications in bioinformatics.

7. Contacts:

Ibrahima BABER
Malaria Research and Training Center (MRTC)
Faculté de Médecine, de Pharmacie et d'Odonto-Stomatologie
Université de Bamako B.P. 1805
Bamako
Mali

Tel: (223) 2225277
baber@mrtcbko.org

Oumou NIARE
Malaria Research Training Center (MRTC)
Faculté de Médecine, de Pharmacie et d'Odonto-Stomatologie
Université de Bamako B.P. 1805
Bamako
Mali
Tel: (223) 6732523
niare@mrtcbko.org

Djibril SANGARE
Malaria Research Training Center (MRTC)
Université de Bamako B.P. 1805
Bamako
Mali
djibril_mrtc@yahoo.fr

Sidy SOUMARE
Mali-ICER (International Center for Excellence in Research)
Faculté de Médecine du point-G
B.P. 1805
Bamako
Mali
Tel: (223) 2231318
sidy@mrtcbko.org

Morocco

1. Institutions:

1.1 UFR: Artificial intelligence and Bioinformatics:

Bioinformatics, in the UFR in Artificial Intelligence and Bioinformatics, deals with the management, the analysis, the modelling and the visualization of biological databases. Since the size of the databases is often exponential, the traditional algorithms are not very effective when seeking for a good computational solution. To take care of this issue, many ways are opened to the researchers to improve the quality of the algorithms:

1. Usage of new information processing methods like artificial neuronal networks, genetic algorithms, etc.
2. Usage of Data mining to explore biochemical databases,
3. Usage of Machine learning on the biological examples to solve, for example, the problem of classification in Bioinformatics.

Our UFR offers in addition a doctoral training in Computer Science and Bioinformatics.

2. Training Activities:

2.1 Long-term training activities:

The National School of Applied Sciences of Tangier (ENSAT) offers a training of three cycles:

1. Preparatory module (two years);
2. Engineers’ module (three years) with an *Ingénieur d'Etat* certification,

3. Doctoral module which includes: a *Diplôme des Etudes Supérieures Approfondies* (DESA) of two years; and a doctorate studies program with a national Ph.D. certification.

Three specializations constitute the teaching trunk of the ENSAT: Computer engineering, Telecom engineering, and electronic systems engineering.

2.2 Short-term training activities:

A 10 days training was organised in Casablanca by the *Institut Pasteur du Maroc* in March 2006 (10-20).

3. Research Interest and Activities:

The following are the present areas of research interest:

1. Machine Learning and Profile Gene Expression of Cancer
2. Predicting Protein structure
3. Hidden Markov Models (HMMs) and multiple alignments
4. Transformational Grammar for sequence modelling
5. Physical Mapping: STSs
6. Evolutionary Computation applied to Genomic and Proteomic
7. Predicate Logic and Protein Structure

4. Expected feedback from the CNRS-NEPAD partnership:

We would like to:

1. Develop a partnership with the African countries that are interested in bioinformatics
2. Exchange professors and Ph.D. students between the labs of these countries
3. Organize yearly workshops and meetings and seminars
4. Organize Masters in bioinformatics in partnership with these institutions
5. Organization of workshops, conferences and seminars
6. Create a database to fight against the diverse diseases in Africa based on the expertise and experience of the different bioinformatics teams.

5. Web site and links:

http://www.ensat.ac.ma/udiab
http://www.pasteur.fr/pasteur/international/annonce_coursBioinfoannonce06_casa.pdf

6. Publications:

No publications in bioinformatics.

7. Contacts:

Mohammed Mounir MAOUENE
Université Abdelmalek Essaadi
Ecole Nationale des Sciences Appliqué (ENSA)
BP 1818
Tanger principale
Morocco
Tel: (212) 39393744
mmaouene@indiana.edu
mmaouen@ensat.ac.ma

Nigeria:

1. Institutions:

1.1 University of Ibadan (UIB)-Ibadan:

There has been significant increase in the number of bioinformatics activities in Nigeria (and West Africa) since 2003 when the program was initiated by the West African Biotechnology Workshops Series (WABWS, http://www.wabw.org) at the University of Ibadan, Nigeria (in collaboration with the South African National Bioinformatics Institute (SANBI, http:/www.sanbi.ac.za). Workshops that were open to scientists from all African countries have seen a very high number of applications from scientists based in West Africa. The encouraging desire to acquire cutting-edge skills to computational process data and extract useful knowledge from genome projects led to the interest of the West African Biotechnology Workshops (WABW) to develop an agenda to address the bioinformatics skills gap among scientists in West Africa. An increased commitment from agencies like NEPAD would be required in the provision of infrastructure to establish and sustain regional and national networks.

1.2 National Biotechnology Development Agency (NBDA) - Abuja:

The summary of bioinformatics activities at NBDA is not available yet.

1.3 Covenant University (CU)-Ota:

Covenant University (with her enriching and growing state-of-the-art laboratories in the area of science and technology, arts, business and social sciences) is presently the Best University in Nigeria (Private University category), based on the recent over-all rating just concluded by the Nigeria University Commission (NUC). Recently, Covenant University has initiated the establishment of a Centre for Applied Biotech, Bio-Informatics and Microbiology (CBBM) to be situated at the University. The institute has been designed to be a Public-Private Partnership for a productive synergy b/w Academia, Industry and Government. The whole concept is still evolving and more details will be release soon.

As regards CBBM, a dedicated computing lab is in plan, but even our computing capacity is presently enormous. In the department of Computer and Information Sciences, we have more than 250 Pentium 4 PCs set aside for teaching and research purposes. Furthermore, we have several moderate speed PCs at the Postgraduate research lab and our engineering departments and units. Our wet lab facilities is presently minimal (basic for teaching), the Centre requirement as it touches the wet-laboratories is also set to upgrade this to basic tools expected at an international centre of learning.

1.4 University of Ilorin (UIL)-Ilorin:

The University of Ilorin was established in 1976 by the Federal Government of Nigeria. Bioinformatics activities started at the University in February 2003 with the establishment of the West African Bioinformatics Research Initiative (WABRI). However, progress has been rather slow due to inadequate funding. We are mainly engaged in Bioinformatics training at the introductory level and proteomics studies on various species of malaria parasites. Recently, we became interested in comparative genome analysis of various species of *Plasmodium* and the comparison of chloroquine sensitive and chloroquine resistant strains of *Plasmodium falciparum*. Other activities and areas of interest can be seen on our website, http://www.wabri.org, although not all our proposed interests have been fully implemented due to our level of funding.

2. Training activities:

2.1 Short-term training activities:

The West African Biotechnology Workshops Series (WABWS) had been organizing region-wide training courses on molecular biology and bioinformatics since 2002; the National Biotechnology

Development Agency had also been organizing in-house bioinformatics training since 2003. The Covenant University's Computer and Information Science Department launched its bioinformatics activities with a workshop in 2005. The University of Ilorin had a bioinformatics training workshop program scheduled in March, 2006.

- The WABWS Advanced Viral Bioinformatics Course, first held in 2005, is set to become an annual program from 2007. Furthermore, in collaboration with USA-based institutions and scientists and the International Institute for Tropical Agriculture, IITA, the WABWS plans to organize Plant Systems Biology workshop (at Ibadan, from 2006). The Redeemer's University Biotechnology Program had been initiated with the WABWS 06 Advanced Molecular Biology Techniques Workshop (May 29–June 10, 2006).

The emerging WABWS-Applied Genome Science program in Nigeria is initially directed at UI-Bioinformatics Program: The overall goal of the UI-Bioinformatics program is to provide the infrastructure to build and enhance the ability of students and faculty at the University of Ibadan to develop and apply computational approaches to biology research and education. The objectives of the educational activities are to (i) develop undergraduate and graduate level courses in bioinformatics at (ii) establish a bioinformatics internship preparatory program, (iii)establish a bioinformatics seminar and workshop series, and (iv) co-organize training workshops with institutions in Nigeria on Applied Genomics and Bioinformatics; RUN-Biotechnology Program: The objective here are (i) to assist in identifying priority needs for establishing an internationally acclaimed biotechnology laboratory for training, education, research and development, and (ii) build the University's preparedness for participation in national and international collaborative projects; NABDA-Bioinformatics Program: The goal is to collaborate with the National Biotechnology Development Agency for capacity-building programs on biotechnology and bioinformatics; Institute of Genetic Chemistry and Laboratory Medicine: Development of a Bioinformatics Mentoring Center.

- The Convent University has organized in 2005, an International Workshop on Pattern Discovery in Biology (IWPDB). More details on this can be found at http://www.covenantuniversity.com/bioinformatics/TrainingActivities/iwpdb.htm. The Convent University plans to organize workshops in a regular basis in order to enhance its capacity building and expansion and the development of bioinformatics in Nigeria and Africa. To this ends, with the emergence of the Centre for Applied Biotech, Bioinformatics and Microbiology at Covenant University, Ota, in collaboration with the America Society of Microbiology (ASM), an encompassing Biotechnology, Bioinformatics and Microbiology is plan for next year, 2007.

- The University of Ilorin, through its Department of Physiology & Biochemistry, in collaboration with WABRI, recently organized an Introductory Bioinformatics Training Workshop from March 1–4, 2006. There were nineteen participants from various Nigerian Universities and Research Institutes at the Workshop. This is expected to be an annual event which will be organized in February/March and is aimed at capacity building in Bioinformatics. An advanced course which will be organized in November every year is also being proposed, subject to the availability of resource persons and funds.

2.2 Long-term training activities:

The University of Ilorin has introduced M.Sc. and Ph.D. programmes in Computer Science (with options in Bioinformatics). The programme is based in the Department of Computer Science and emphasis is on the development of algorithms to solve problems in bioinformatics.

The Covenant University offers M.Sc. and Ph.D in Computer Science with option in Bioinformatics (Computational Biology). Furthermore, through affiliated departments, the CBBM is been design to award Diploma and Degree certificates in Biotechnology.

3. Research interest and activities (including key publications and software):

The present areas of research interest ahost
nd activities include:

1. Viral pathogenesis, Genomics and Bioinformatics (UIB)
2. Plants System Biology (UIB)
3. Natural Product Research (UIB)
4. Genome Data Analysis (UIL)
5. Functional genomics and proteomics (CU, UIB, UIL)
6. Population studies (UIB)
7. High throughput genotyping (CU, UIB)
8. Cheminformatics (UIB)
9. Annotation of pathogen genomes (CU, UIB)
10. Advanced Molecular Biology Techniques Course (CU, UIB)
11. Genomic and proteomics studies on *Plasmodium falciparum, Plasmodium vivax* and *Mycobacterium tuberculosis* (UIL)
12. System Biology: Derivation of an intensive genes network for *Plasmodium falciparum* (CU)
13. Protein Engineering/Structural Bioinformatics (CU, UIB)
14. Comparison of genome sequences of chloroquine-sensitive and chloroquine resistant strains of *Plasmodium falciparum* (UIL)
15. Development of the African Avian Influenza Database (UIB)
16. Epidemiological and phylogenetic studies on different strains of human immunodeficiency virus (HIV) (UIL)
17. Development of Bioinformatics tools and algorithms (CU, UIL)
18. Select short and long oligos for microarray and related applications from a large EST (CU)
19. Basic and Advanced Bioinformatics Training workshops/Seminars (UIB)
20. High performance computing and grid development (CU)
21. Combinatorial Libraries and drug design (CU)
22. Detection of recombination (CU)
23. EST clustering (CU)
24. Development of an Africa *Plasmodium falciparum* database (afriPFdb). A similar database is presently at a developmental stage for the Hepatitis C Virus. (CU)
25. Forensic Research and DNA fingerprinting (CU)
26. Environmental Remediation and Sustainability (CU)
27. Host-pathogen protein interaction (our present study is on Mosquito-*Plasmodium falciparum* protein interaction) (CU)

4. Expected feedback from the CNRS-NEPAD partnership

Our expectations from the partnership include the following:

1. Strengthening of local infrastructural capacity for bioinformatics training, education, research and development in Nigerian institutions (especially at Ibadan, Ilorin, Ota, Abuja and Redemption City)
2. Curriculum development for undergraduate, graduate and professional bioinformatics programs
3. Resource persons coming to Nigeria to participate in training workshops, courses and

research activities

4. Short- to medium-term training opportunity for graduate students and Post-Doctoral Fellows
5. Funding for research collaborations with African laboratories involved in bioinformatics and genome data analysis
6. Undergraduate students visitation
7. Sending and receiving Post-doc fellows
8. Ph.D students research visitation
9. Setting up of funds to promote exchange visits (short-term and long-term) by scientists and students engaged in Bioinformatics from well developed and less developed laboratories
10. Establishment of Bioinformatics nodes that will serve various parts of Africa
11. Development of full-scale (M.Sc. and Ph.D.) courses in Bioinformatics.

5. Web sites and links:

http://www.covenantuniversity.com
http://www.run.edu.ng
http://www.uniben.edu
http://www.wabri.org
http://www.wabw.org
http://www.unilorin.edu.ng
http://www.wabri.org
http://www.asopah.org

6. Publications:

1. **Ezekiel F. Adebiyi**. Using suffix tree for efficient selection of unique oligos for large EST databases. Submit to the 6th Intl. Workshop on Algorithm of Bioinformatics, Switzerland, 2006 and International Journal on Bioinformatics and Computational Molecular Biology (IJBCB).
2. **Adebiyi, EF.** Fatumo, S and Osamor, V. A. A significant difference in output among microarray experiment front-end tools. (1 page) Poster proceeding of the 13th Annual International conference on Intelligent Systems for Molecular Biology (ISMB), 2005 and also presented at the Covenant University International Workshop on Pattern Discovery in Biology (CUWPDB 2005), April 18-27 2005.
3. **Adebiyi, EF**. Systems biology and the development of vaccines and drugs for malaria treatment. (7 pages), Biokemistri, Journal of the Nigeria Society for Experimental Biologist (NISEB) (to appear) and Presented at the Covenant University International Workshop on Pattern Discovery in Biology (CUWPDB 2005), April 18-27 2005.
4. **Adebiyi, EF**. The probability of occurrence in a single sequence. Current In Computational Biology, Posters Proc. Eighth Annual International Conferences on Research in Computational Molecular Biology (RECOMB), 451-2, 2004.
5. **Adebiyi, EF**. and Dipe T. Finding higher order motifs under the levenshtein measure. Posters Proc. of the second IEEE Computer Society Bioinformatics Conference (CSB), 583-4, 2003.
6. **Adebiyi, EF**. Pattern Discovery in Biology and Strings Sorting: Theory and Experimentation, Shaker Publisher, Aachen, Germany, 2002.
7. **Adebiyi, EF**. and Kaufmann, M. Extracting Common Motif under the Levenshtein Metric: Theory and Experimentation. Second International Workshop on Algorithms in Bioinformatics (WABI), 140-156, 2002.
8. **Adebiyi, EF**. Extracting Simple Motif under the Levenshtein Metric, Current in Computational Biology, Posters Proc. Sixth Annual International Conferences on Research in Computational Molecular Biology (RECOMB), 1-2, 2002.
9. **Adebiyi, EF**., Jiang, T., and Kaufmann, M. An efficient algorithm for finding short approximate non-tandem repeats (Extended abstract), Proc. of the Ninth International Conference on Intelligent Systems for Molecular Biology (ISMB) and appeared as Bioinformatics 17(1): S5-S13, 2001.

10. **Bewaji, C. O.** and Dawson, A. P. (1995) Thapsigargin protects human erythrocyte Ca^{2+}-ATPase from proteolysis. *Cell Calcium* **17**, 14 – 20.

11. **Bewaji, C. O.** (1995) Calcium pump isoforms of the cell surface and intracellular membranes. *Biokemistri* **5**(2), 75 – 86.

12. **Bewaji, C. O.** (1995) Ionic requirements for inositol trisphosphate- and thapsigargin-induced Ca^{2+} release from rat liver endoplasmic reticulum. *Nig. J. Biochem. Mol. Biol.* **10**, 1 – 6.

13. **Bewaji, C. O.** (1996) Inhibitor sensitivities of the Ca^{2+}-ATPase from skeletal muscle sarcoplasmic reticulum after proteolysis with trypsin. *Biokemistri* **6**(2), 52 – 58.

14. **Bewaji, C. O.** (1996) Comparison of the effects of thapsigargin, 2,5-di-(tert-butyl)-1,4-hydroquinone and cyclopiazonic acid on calcium release from rat liver endoplasmic reticulum. *Nig. J. Biochem. Mol. Biol.* 11, 27 – 30.

15. **Bewaji, C. O.**, Rich, G. and Dawson, A. P. (1998) Inhibitor sensitivities of the Ca^{2+}-ATPase and Ca^{2+}-dependent p-nitrophenyl phosphatase activities of skeletal muscle sarcoplasmic reticulum. *Biokemistri* **8**, 1 – 13.

16. **Bewaji, C. O.** (1999) Identification of drug target sites and vaccine candidates in *Plasmodium falciparum* using functional genomics. *Trop. J. Hlth. Sci.* **6,** 1 – 6.

17. **Bewaji, C. O.** (2000) Darwin's Genesis: An update on the human genome project. *African Scientist* **1(1)**, 11 – 16.

18. Bolaji, O. M.; **Bewaji, C. O.**; Oduola, A. M. J. and Bababunmi, E. A. (2000) Expression of calcium-binding proteins (Ca^{2+}-ATPase and translationally controlled tumor protein) in chloroquine resistant *Plasmodium falciparum*: A possible role of the proteins in apoptosis of the parasite. *Biokemistri* **10**, 111 – 121.

19. **Bewaji, C. O.** (2000) Rolling Back Malaria: The role of Bioinformatics. *African Scientist* **1(2):**127–141.

20. **Bewaji, C. O.** (2000) Digital representation of the amino acid sequence of proteins and comparison of sequences using the Microsoft Excel software. *African Scientist* **1(3):**196–206.

21. **Bewaji, C. O.** (2000) A bioinformatics approach to the study of the Acquired Immune Deficiency Syndrome (AIDS). *African Scientist* **1(3):**207–214.

22. Bababunmi, E. A. and **Bewaji, C. O.** (2002) Oxidative stress and antioxidant therapy in wasting conditions associated with malnutrition, HIV/AIDS, tuberculosis and malaria. *Biokemistri* **12(3):**135–139.

23. **Bewaji, C.** O. (2003) Data mining in novel genomes. *African Scientist* **4(1):**43 – 49.

24. **Bewaji, C. O.** and Bababunmi, E. A. (2004) Gene arrangements in human chromosomes: Recent findings from the Human Genome Project. *African Scientist* **5(4)**, 203 – 210. **Bewaji, C. O.** and Bababunmi, E. A. (2006) Comparison of the proteomes of three *Plasmodium* species: *Plasmodium falciparum, Plasmodium vivax* and *Plasmodium yoelii*. *Int. J. Biomed. Hlth. Sci.* **2(2)**, *in press*.

25. **Oyekanmi Nashiru**, Chuong Huynh, Raphael Isokpehi, Jessica Kissinger, Olufunso Olorunsogo (2005) Building Bioinformatics Capacity in West Africa. *African J. Med. And Medical Sci., in press.*

26. David L. Zechel, Stephen Reid, Dominik Stoll, **Oyekanmi Nashiru**, R. Antony Warren and Stephen G. Withers (2003). Mechanism, Mutagenesis, and Chemical Rescue of a Beta-Mannosidase from *Cellulomonas fimi*. *Biochemistry* **42:**7195-7204.

27. **Oyekanmi Nashiru**, D.L. Zechel, D. Stoll, T. Mohammedzeh, R.A.J. Warren, S.G. Withers (2001).Beta-Mannosynthase: Synthesis of beta-mannoside with a beta-Mannosidase mutant. *Angew.*

Chem. Int. Ed. **40:**417-20.

28. David.L. Zechel Stephen P. Reid, **Oyekanmi Nashiru**, Christoph Mayer, Dominic Stoll, David Jakeman, R. Antony Warren, Stephen Withers (2001). Enzymatic Synthesis of Carbon-Fluorine Bonds. *J. Am. Chem. Soc*. **123:**4350.
29. **Oyekanmi Nashiru**, Suk-Hoon Ko, Se Yong Lee, Dae Sil Lee. (2001) Novel alpha-Glucosidase from an extreme thermophile, *Thermus caldophilus*, GK24: Purification, Gene cloning and Expression in *E. coli*. *J. Biochem and Mol. Biol*. **34:**347-354.
30. C.H. Kim, **O. Nashiru**, J.H. Ko. (1996) Purification and Biochemical Characterization of Pullulanase Type I from *Thermus caldophilus* GK24. *FEMS Microbiol. Lett*. **138:**147.
31. Hide WA, **Isokpehi RD**. (2004) Positive selection scanning of parasite DNA sequences. *Methods Mol Biol*. **270:**127-50.
32. **Isokpehi RD**, Hide WA (2003) Integrative analysis of intraerythrocytic differentially expressed transcripts yields novel insights into the biology of Plasmodium falciparum. *Malar J*. **2(1):**38.
33. **Kissinger JC**. (2006) A tale of three genomes: the kinetoplastids have arrived. *Trends Parasitol*. [Epub ahead of print].
34. Aguero F, Zheng W, Weatherly DB, Mendes P, **Kissinger JC**. (2006) TcruziDB: an integrated, post-genomics community resource for *Trypanosoma cruzi*. *Nucleic Acids Res*. **34**:D428-31.
35. Huang J, Mullapudi N, Lancto CA, Scott M, Abrahamsen MS, **Kissinger JC**. (2004) Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol*. **5(11):**R88.
36. Striepen B, **Kissinger JC.** (2004) Genomics meets transgenics in search of the elusive Cryptosporidium drug target. *Trends Parasitol*. **20(8):**355-8.
37. Huang J, Mullapudi N, Sicheritz-Ponten T, **Kissinger JC**. (2004) A first glimpse into the pattern and scale of gene transfer in Apicomplexa. *Int J Parasitol.* 34(3):265-74.
38. **Kissinger JC**, Gajria B, Li L, Paulsen IT, Roos DS. (2003) ToxoDB: accessing the Toxoplasma gondii genome. *Nucleic Acids Res*. **31(1):**234-6.
39. **Kissinger JC**, Brunk BP, Crabtree J, Fraunholz MJ, Gajria B, Milgram AJ, Pearson DS, Schug J, Bahl A, Diskin SJ, Ginsburg H, Grant GR, Gupta D, Labo P, Li L, Mailman MD, McWeeney SK, Whetzel P, Stoeckert CJ, Roos DS. (2002) The Plasmodium genome database. *Nature*. **419(6906):**490-2.
40. **Kissinger JC**, Souza PC, Soarest CO, Paul R, Wahl AM, Rathore D, McCutchan TF, Krettli AU. (2002) Molecular phylogenetic analysis of the avian malarial parasite *Plasmodium* (Novyella) *juxtanucleare*. *J Parasitol*. **88(4):**769-73.

7. Contacts:

Ezekiel F. ADEBIYI
Department of Computer and Information Sciences
College of Science and Technology
Covent University P.M.B 1023
Ota
Nigeria
Tel: (234) 181796704 Ext. 3021
efadebiyi@yahoo.com

Clement BEWAJI, Ph.D.

Department of Biochemistry and Physiology
Faculty of Health Sciences
University of Ilorin
Nigeria
cobewaji@yahoo.com

Patrick ERAH, Ph.D.
Department of Pharmacy
University of Benin
Benin-City
Nigeria
Tel: (234) 8023360318
erah@uniben.edu

Oyekanmi NASH, Ph.D.
Program Director, West African Biotechnology Workshops Series
Applied Genomics and Bioinformatics Unit
Institute for Advanced Medical Research and Training
College of Medicine
University of Ibadan
Nigeria
Tel: (234) 8051283334
onash@wabw.org
oyekan.nash@gmail.com

Sunday WUYEP, Ph.D.
Bioinformatics and Planning Section
National Biotechnology Development Agency
Abuja
Nigeria
Tel: (234) 8037863182
wuyepsn@yahoo.com

South Africa

1. Institutions:

1.1 South African National Bioinformatics Institute (SANBI) – University of Western Cape, Cape Town:

SANBI, and through it the University of Western Cape (UWC), is the EMBnet node for South Africa, and is internationally-recognized for shaping health biotechnology outputs in SA (Nat. Biotech. 2004, 22:DC39) with strong research programmes tackling diseases such as HIV/AIDS, TB, cancer and diabetes. UWC has acquired substantial expertise in bioinformatics. It has been the site for the establishment of Africa's first graduate programme in Bioinformatics. It offers Masters and PhD training by research programmes, and is the host of the South Africa – Stanford Biomedical Informatics Training Programme (stanford.sanbi.ac.za). It devotes much effort to supporting the development of an HIV vaccine (www.saavi.ac.za), and to the understanding of the process of acute infection in HIV through the NIH funded CAPRISA programme (www.caprisa.org). It has made critical contributions already to methods for finding disease-susceptibility genes (Nature Genetics 1999, 22(3):255-9; Genome Biology 2003, 4:119), transcript expression characterization (Science 2005, 309:1559; Nucleic Acids Res. 2001, 29:234; Genome Research 2003, 13:1222), HIV research (Lancet 2002, 359:992), pathogen research (Trends in Parasitology 2005, 21:107), cancer (PNAS 2003, 100:13418), all the while maintaining a solid capacity development drive through additional regular global and national training programmes (www.sanbi.ac.za/mrc; s-star.sanbi.ac.za; and www.sanbi.ac.za/mrc/tdr2006/). SANBI hosts the SA

Medical Research Council (MRC) Unit for Bioinformatics Capacity Development, which has supported Bioinformatics development Africa-wide for the last 7 years. SANBI is the African regional bioinformatics training centre for the World Health Organisation, and is the future site for the establishment of the annotation of the Tsetse Fly genome (2008). SANBI was the lead organisation in establishment of the SA National Bioinformatics Network (NBN), and is the most high profile node of the NBN. SANBI established the African Society of Bioinformatics and Computational Biology, and has supported bioinformatics workshops and establishment of bioinformatics research sites in South Africa, Nigeria, Tunisia, Kenya, Mali, Uganda, Ghana, and Angola. SANBI, funded through several local and international granting agencies, consists of three full time Professors of Bioinformatics, support and training staff, two developers, a rapidly growing number of PhD and MSc students and two post docs.

Professor Win Hide established the SA National Bioinformatics Institute, which he directs, in 1996. He is a computational biologist and is a Kerr International Fellow of the Ludwig Institute for Cancer Research. He researches gene expression relevant to disease in the areas of multifactorial diseases and pathogens. These include diabetes, cancers, and host-pathogen diseases such as HIV.

Prof Heikki Lehvaslaiho joined SANBI October 2005. His main areas of work are human sequence variation, genotype and phenotype correlations, databases and standards development. He maintains the Human Genome Variation Database (HGVBase) together with Anthony Brookes (Leicester University, UK), and is a core developer of BioPerl bioinformatics toolkit.

Prof Vladimir B Bajic joined SANBI in January 2006. His interest is in application of bioinformatics to human health problems and biological discoveries with focus on the development of tools for transcription regulation modelling/analysis and regulatory networks. He has specific interest in gene networks controlled by steroid hormones, links of transcriptional regulators with genes and diseases, particularly cancers; diagnostic markers; and biomedical text-mining. He has developed over 40 bioinformatics tools and he publishes extensively in high-impact journals including Nature Genetics, Science, Nature Biotechnology, and PLoS Genetics (www.sanbi.ac.za/~vlad). Many bioinformatics predictions he made have been experimentally confirmed.

1.2 National Bioinformatics Network (NBN) – Cape Town:

The NBN coordinates much of the bioinformatics work. The NBN acts as a network of nodes situated on 6 campuses (University of Pretoria/Limpopo/CSIR, University of the Witwatersrand, Rhodes University, Stellenbosch University, University of the Western Cape/SANBI, University of Cape Town) and it is hoped that two other campuses will affiliate. The nodes are part of their universities and operate independently but cooperatively with each other and NBN central. All nodes will soon be connected via a fast backbone network for research and education.

The NBN has four key roles: to fund research (reviewed by an international Scientific Advisory Board), supplement and complement teaching, provide services and infrastructure development.

1.3 Biomedical Informatics Research Division of the MRC (MRC-BIRD) – Cape Town:

BIRD has a small but growing bioinformatics interest which is largely in collaboration with the HIV-1 Immune Pathogenesis and Therapeutic Research Programme (see 1.7) at the University of Pretoria headed by Professor Sharon Cassol and Dr Tulio de Oliveira. BIRD is predominantly involved in monitoring antiretroviral treatment cohorts in South Africa, virus evolution and the development of drug resistance as well as some tools and the development of ontologies. MRC-BIRD has one full time post-doctoral researcher (Dr Chris Seebregts), two senior developers (Michael Ndlokovane and Mpumelelo Msimanga) and one PhD student in biomedical informatics (Venessa Timmerman).

1.4 Computational biology group at the University of Cape Town (CB-UC) – Cape Town:

The computational biology group (http://cbio.uct.ac.za) at the University of Cape Town (UCT) is the UCT affiliate of the NBN. The group has research interests in molecular evolution, pathogen bioinformatics, comparative genomics and alternative splicing. Located at the UCT medical campus, the group is a shared initiative of the UCT Science and Health Sciences faculties and has close ties to researchers within the Institute of Infectious Diseases and Molecular Medicine and the Molecular and Cell Biology Department. In addition to research in computational biology the group coordinates bioinformatics education at undergraduate and postgraduate levels at UCT and provides bioinformatics services to the university research community. The service role of the group is currently targeted towards researchers engaged in high throughput biology, with a particular focus on the support of microarray data management and analysis.

It is a rapidly expanding research group, currently consisting of two faculty members, Associate Professor Cathal Seoighe and Dr Nicola Mulder as well as three post-doctoral researchers and ten post-graduate students.

1.5 Bioinformatics and Computational Biology Unit of the University of Pretoria (BCB-UP) – Pretoria:

The University of Pretoria Bioinformatics and Computational Biology Unit (http://www.bi.up.ac.za) was established in 2003, out of the Department of Biochemistry. It is part of the Faculty of Natural and Agricultural Sciences, of the African Centre for Gene Technologies (ACGT; http://www.acgt.co.za) and is a Node of the NBN.

Research projects include the structural modeling various malaria drug target enzymes, the development of a Functional Genomics Information Management System, the structural annotation of the malaria proteome, comparative genomics and proteomics in malaria, a tool for the development of analogous ligand libraries and a software suite for the downstream analysis of microarray results.

The Unit is involved in post-graduate training, research, short courses, support and services.

Short courses are presented on a regular basis, often together with the NBN or international scientists, and have included general courses, a course on molecular dynamics, genome annotation and microarray analysis.

Services and support include the hosting of standard bioinformatics databases and software suites, as well as a variety of specialized databases and tools. High-throughput services are available on the node’s 64x CPU cluster.

Staff members are Prof Fourie Joubert (Senior Bioinformaticist), Dr Oleg Reva (Node Manager), Mr Andrew Bwalya (System Administrator) and Mrs Loveness Dzikiti (Part-time Statistican).

The Unit offers PhD, MSc and BSc Hons programmes in Bioinformatics, and currently houses 21 post-graduate students. An undergraduate BSc IT Bioinformatics programme will be implemented in 2007 through the Department of Computer Science.

1.6 Department of Veterinary Tropical Diseases of the University of Pretoria - Onderstepoort; and Molecular Biology Department of the ARC-OVI – Onderstepoort: (VTD-MB)

Heartwater is one of the most economically important tick-borne diseases of domestic and some wild ruminants in Africa, and it also exists in parts of the Caribbean. It is caused by the obligate intracellular rickettsia *Ehrlichia ruminantium*, and research in South Africa to find a vaccine goes back over many decades. Vaccine development research for heartwater is carried out by a collaborating team of scientists working in the Faculty of Veterinary Science of the University of Pretoria and the Onderstepoort Veterinary Institute (OVI), which is an institute of the South African Agricultural Research Council (ARC). There is also collaboration with the Unit for Bioinformatics and Computational Biology of the University of Pretoria, see section 1.5 above.

The research group has recently sequenced and annotated the whole genome of this rickettsia and is currently using bioinformatics, molecular, and immunological techniques to mine the genome for genes that will be effective in a recombinant vaccine. This work is supported by grants from the South African National Research Foundation, the South African Department of Science and Technology, the South African Red Meat Research and Development Trust, and the European Union.

The team includes nine researchers and technical staff, and four postgraduate students.

More information on the E. ruminantium genome project is available at the following URL:
http://www.bi.up.ac.za/Ehrlichia_ruminantium

1.7 HIV-1 Immune Pathogenesis and Therapeutic Research Programme (HIV-PTRP)– Pretoria:

The HIV-1 Immune Pathogenesis and Therapeutic Research Programme was funded, by Prof Sharon Cassol, with a Wellcome Trust grant. The group has published 35 manuscripts in peer-reviewed journals such as *J Virology*, *AIDS*, *J Infect Dis*, *Bioinformatics*, *Genetics* (an average impact factor of 7.4) and given more than 50 oral and poster presentations at international meetings in the USA, Europe, South Africa and Thailand. The majority of these publications had a bioinformatics component.

The group has organized, and/or contributed to, ten major bioinformatics workshops held at the University of KwaZulu-Natal, Stanford, and various institutions in the U.K., Belgium, Italy and Brazil. It has established a unique website/online resouce focused on HIV-1/AIDS epidemic in Africa, www.bioafrica.net (Science 2005;310;1877). It has developed software programmes to facilitate subtyping and phylogenetic analysis of HIV-1, and related pathogens (Bioinformatics 2002;19:1-2; Bioinformatics 2005;21:3797-3800).

No formal degree in Bioinformatics per se. BSc, MSc and PhD studies in HIV-1 pathogenesis and molecular evolution as part of the programmes ongoing research in on subtype C viruses and their role in the explosive AIDS epidemic in southern Africa. This is an ongoing comprehensive programme that integrates wet lab research with bioinformatics, with wet lab research driving bioinformatics and vice verse. The bioinformatics component, initiated by Dr. Tulio de Oliveira, during is PhD programme, is conducted in strong collaboration with researchers in the Evolutionary Group at the University of Oxford.

2 MSc, 2 PhD, and 2 postdoctoral fellows were trained and graduated by the team of the HIV-1 Immune Pathogenesis and Therapeutic Research Programme. 2 additional students (one PhD, and one MSc) will be submitting their theses in June of 2006. All students received in-house and, in most cases, international training in molecular virology and bioinformatics. A third MSc student from the MRC received a full year of training in Bioinformatics. Many of the students received scholarships to present their work at international meetings, and most have several first author publications.

1.8 Council for Scientific and Industrial Research (CSIR) – Pretoria:

The CSIR activities in structural bioinformatics are based on homology modelling and the searching of structural databases for active site conformations that may be used in rational drug design target identification. These activities are part of a rational drug design programme using ligands and targets based on rational drug design protocols, and include a significant chemiinformatics component. CSIR structural bioinformatics plays a key role in enzyme reaction mechanism identification.

CSIR is also putting together all the tools required for running a genomics and proteomics platform; this includes data analysis and databasing.

1.10 Wits Bioinformatics (WITS-B) – Johannesburg:

The University of the Witwatersrand (Wits) is a major research and teaching University based in Johannesburg. Bioinformatics is part of the research and teaching programmes of the Faculties of Science and Health Sciences.

At the centre is Wits Bioinformatics, a joint project of the two faculties. It has two floors of a building on the University's main campus. The three main roles of Wits Bioinformatics are teaching, service and research. It plays a coordinating role and helps researchers and postgraduate students with projects. There are three main schools that have significant bioinformatics expertise: molecular & cell biology, computer science, and human genetics. There is also some activity in other schools in the faculties of science, health sciences and engineering (which also has a biomedical engineering programme).

Wits Bioinformatics hosts various databases for internal researchers, have compute servers available and offer a consultancy service for researchers and postgraduate students.

Wits Bioinformatics has 3 staff and a number of MSc and PhD students pursuing bioinformatics projects and an MSc by dissertation in bioinformatics has been provisionally approved.

2. Training Activities:

2.1 Long-term training activities:

University of Western Cape, University of Cape Town, University of the Witwatersrand and University of Pretoria offer undergraduate and honours courses in bioinformatics and most offer post graduate research degrees in bioinformatics. A significant teaching, research and service in bioinformatics exist in all universities, though the resources are stretched and uneven.

Lack of senior bioinformaticists is a big problem as is the high cost of Internet access.

2.2 Short-term training activities:

Short-term trainings are organized on regular bases by a number of South african institutions.

SANBI: SANBI organises several general and specialised training courses in Bioinformatics each year. The list of past and future courses is available at the following URL: http://www.sanbi.ac.za/mrc/Training.html

NBN: The NBN runs an intensive sequence of courses (approximately 16 weeks) which is aimed primarily at SA postgraduate students though researchers and external visitors also attend. These courses are given cooperatively by staff members from the different nodes of the NBN as well as by international experts who are brought out by the NBN. Services such as access to major tools and databases have been rolled out, and high performance computing is a focus and need.

MRC: The African Workshop on Virus Evolution and Molecular Epidemiology, and the South African Treatment Management Workshop, were organized at Durban by the MRC from 2nd to the 6th of July 2003.

HIV-1 Immune Pathogenesis and Therapeutic Research Programme: Workshops given in South Africa included:

- The HIV Data Management and Data Mining Workshop, Durban, South Africa, December 13-16th, 2004.
- Introduction to HIV Sequencing and Phylogenetic Analysis, Durban, South Africa, Feb 9-13, 2004.
- http://bioafrica.mrc.ac.za/AfricanWorkshop.html. South African Workshop on Virus Evolution/Molecular Epidemiology, Durban, South Africa, 2003
- http://bioafrica.mrc.ac.za/AfricanWorkshop.html. South African HIV/AIDS Treatment Management Workshop Durban, South Africa, 2003
- http://bioafrica.mrc.ac.za/AfricanWorkshop.html. Resistance and Subtyping Workshop, Cape Town, South Africa, 2003
- http://wwww.sanbi.ac.za/tdrcourse/faculty.html;http://www.sanbi.ac.za/tdrcourse/schedule.html Workshop sponsored by UNDP/World Bank/WHO. Special Program for Research and Training in Tropical Diseases at the South African National Bio-informatics Institute, Jan 21-25, 2002 (two lectures)

3. Research Interest and Activities:

The following are the present areas of research interest:

1. Recombination detection and analysis in a range of virus species and cellular organisms (Darren Martin)
2. Development of recombination and detection analysis software (WITS-B, Darren Martin)
3. Recombinant vaccine discovery (VTD-MB)
4. EST clustering (SANBI, WITS-B)
5. Annotation of bacterial genomes (VTD-MB)
6. Development of Functional Genomics Information Management Systems (BCB-UP)
7. Cheminformatics (CSIR, BCB-UP)
8. Functional genomics and proteomics (SANBI, CB-UC, BCB-UP)
9. Molecular evolution (CB-UC)
10. Virus evolution (MRC-BIRD)
11. Pathogen bioinformatics (SANBI, CB-UC)
12. Ontologies (SANBI, MRC-BIRD)
13. Systems Biology (SANBI)
14. Monitoring antiretroviral treatment cohorts (MRC-BIRD)
15. Development of drug resistance (MRC-BIRD)
16. Rational drug design (CSIR)
17. Structural bioinformatics (BCB-UP , CSIR)
18. Microarray data management and analysis (SANBI)
19. Structure annotation and modelling of malaria proteins (BCB-UP)
20. Factors controlling the rapid, global expansion of the HIV-1 subtype C epidemic (HIV-PTRP)
21. Origins, adaptive evolution and increased virulence of C viruses (HIV-PTRP)
22. Implications of HIV-1 genetic variation on vaccine/microbicide development, and on the efficacy of antiretroviral therapy, including the emergence of drug resistance (HIV-PTRP)
23. Impact of HIV-1 genetic diversity on viral tropism, compartmentalization and disease progression (HIV-PTRP)
24. Genetics/immunology of host-virus interactions (HIV-PTRP)
25. Molecular interactions between HIV-1 and associated opportunistic infections, including tuberculosis and HHV-8 (Kaposi's Sarcoma) (HIV-PTRP)

4. Expected feedback from the CNRS-NEPAD partnership:

The South African delegates do not believe that programmes for South African students to pursue postgraduate programmes abroad are a priority, or even necessarily desirable. There is capacity to supervise research in SA and it is key to strengthen it. Diverting postgraduates abroad could weaken rather than strengthen our programmes. There are numerous opportunities already for South Africans to study abroad, and we do not see additional programmes as being useful. On the other hand, exposure to top-level international research and researchers would be extremely desirable, and the possibility for our postgraduate students to spend some months in foreign labs would be welcomed enthusiastically.

The South African delegates welcomed the NEPAD initiative and felt that there could be a number of positive outcomes:

1. Visitors coming to South Africa to give courses and help with research collaboration

2. Encouragement of Post-Doc programmes both for South Africans abroad and visitors coming to South Africa
3. Accommodation of South African postgraduate students and researchers for short-term visitors abroad.

5. Web sites and links:

http://www.mrc.ac.za/
http://www.licr.org/
http://www.sanbi.ac.za/
http://www.saavi.org.za/
http://www.nbn.ac.za/
http://s-star.sanbi.ac.za/
http://cbio.uct.ac.za/
http://www.bioafrica.net/
http://stanford.sanbi.ac.za/
http://www.bi.up.ac.za/
http://www.acgt.co.za/
http://www.caprisa.org/
http://www.up.ac.za/academic/veterinary/dept_vtd_research/progr11.htm
http://www.sanbi.ac.za/~vlad/
http://www.bioinf.wits.ac.za/

6. Publications:

1. E. Van der Ryst, C. Gray, C. Williamson, L. Morris, Q. Abdool Karim, **W. Hide**, J. Esparza. Promoting HIV vaccine research and development in Southern Africa. *South African Medical Journal*. 1997. Aug;87(8):1015-6
2. **Hide**, **W.**, Burke, J., Christoffels, A., Miller, R.. A novel approach towards a comprehensive consensus representation of the expressed human genome. In *Genome Informatics* 1997, pp187-196. Satoru Miyano and Toshihisa Takagi Eds. Universal Academy Press Inc. Tokyo, Japan. ISSN 0919-9454
3. Burke, J., Wang, H., **Hide, W**. and Davison D. Alternative gene form discovery and candidate gene selection from gene indexing projects. *Genome Research.*, 1998 Mar;8(3):276-90.
4. Karlak, B.R. and **Hide, W.A.**, Assessment of utility of ESTs for nucleotide diversity using available assembled alignments from dbEST, STACK 2.0 and STACK-INDEX . In *Genome Informatics* 1998, pp266-268. Satoru Miyano and Toshihisa Takagi Eds. Universal Academy Press Inc. Tokyo, Japan. ISBN 0919-9454
5. Lori S. Sullivan, John R. Heckenlively, Sara J. Bowne, Jian Zuo, **Winston A. Hide**, Andreas Gal, Michael Denton, Chris F. Inglehearn, Susan H. Blanton, Stephen P. Daiger. Mutations in a novel retina-specific gene cause autosomal dominant retinitis pigmentosa *Nature Genetics*, 1999 Jul;22(3):255-9
6. R. T. Miller, AG. Christoffels, C. Gopalakrishnan, J. A. Burke, A. A. Ptitsyn, T. R. Broveak and **W. A. Hide**. A comprehensive approach to clustering of expressed human gene sequence: The Sequence Tag Alignment and Consensus Knowledgebase. *Genome Research*. 1999 Nov;9(11):1143-1155.
7. J. Burke, D. Davison, and **W. Hide**. D2_cluster: A Validated Method for Clustering EST and Full-length cDNA Sequences. *Genome Research*. 1999 Nov; 9 (11):1135-1142.
8. *Science*: SCOPE, **Hide, W**. ‘Malaria and the Developing Worlds of Genome Bioinformatics’. *Science SCOPE*: 2000. scope.educ.washington.edu/research/malaria (invited by the editors of Science)
9. Alan Christoffels, Antoine van Gelder, Gary Greyling, Robert Miller, Tania Hide and **Winston Hide**.

STACKdb: Sequence Tag Alignment and Consensus Knowledgebase. *Nucleic Acids Research* 2001 Jan 1;29(1):234-8

10. **Winston Hide**. Exscript: An 'Ex'-centric approach to the description of transcript diversity. Editorial. *Bioinformatics*. 2001 Jun;17(6):485-6..
11. R. M. Warren, M. Richardson, S. L. Sampson, G. D. van der Spuy, W. Bourn, J. H. Hauman, H. Heersma, **W. Hide**, N. Beyers, P. D. van Helden. Molecular evolution of Mycobacterium tuberculosis: phylogenetic reconstruction of clonal expansion; *Tuberculosis* (Edinb). 2001;81(4):291-302.
12. **Winston Hide**, Vladimir N. Babenko, Peter A. van Heusden, and Janet F. Kelso. The contribution of exon skipping events on Chromosome 22 to protein coding diversity. *Genome Research. 2001 Nov;11(11):1848-53.*
13. J. Gamieldien, A. Ptitsyn, **W. Hide**. Eukaryotic genes in Mycobacterium tuberculosis could have a role in pathogenesis and immunomodulation. *Trends in Genetics*. 2002 Jan;18(1):5-8.
14. Carlton JM, Muller R, Yowell CA, Fluegge MR, Sturrock KA, Pritt J Vargas-Serrato E, Galinski MR, Barnwell JW, Mulder N, Kanapin A, Cawley SE, **Hide WA**, Dame JB. Profiling the malaria genome: a gene survey of three species of malaria parasite with comparison to other apicomplexan species. *Molecular Biochemistry and Parasitology*. 2001 Dec;118(2):201-10.
15. Nico C. Gey van Pittius, Junaid Gamieldien, **Winston Hide**, Gordon D. Brown, Roland J. Siezen, and Albert D. Beyers. The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria. *Genome Biology*. 2001;2(10):RESEARCH0044. Epub 2001 Sep 19
16. Mashishi T, Loubser S, **Hide W**, Hunt G, Morris L, Ramjee G, Abdool-Karim S, Williamson C, Gray CM. Conserved domains of subtype C nef from South African HIV type 1-infected individuals include cytotoxic T lymphocyte epitope-rich regions. *AIDS Research and Human Retroviruses.* 2001 Nov 20;17(17):1681-7.
17. Carpenter JE, Christoffels A, Weinbach Y, **Hide WA**. Assessment of the parallelization approach of d2_cluster for high-performance sequence clustering. 2002 *Journal of Computational Chemistry*. 2002 May;23(7):755-7
18. **Hide W**, Mizrahi V, Venkatesh B, Brenner S, Simpson A, Blatch G, Soodyall H, Denby K, Wingfield M, Wingfield B, van Helden P, Ramesar R, Dorrington R, Kelso J, Oppon, E, Goyvaerts E, Ramsay M, de Villiers E, van Heerden C, Allsopp B, Seoighe C. A platform for genomics in South Africa. *South African Medical Journal*. 2001 Dec;91(12):1006-7
19. Appel S, Filter M, Reis A, Hennies HC, Bergheim A, Ogilvie E, Arndt S, Simmons A, Lovett M, **Hide W**, Ramsay M, Reichwald K, Zimmermann W, Rosenthal A. Physical and transcriptional map of the critical region for keratolytic winter erythema (KWE) on chromosome 8p22-p23 between D8S550 and D8S1759. *European Journal of Human Genetics*. 2002 Jan;10 (1):17-25
20. Abdool Karim; Miriam Adhikari; Hoosen Coovadia; Anna Coutsoudis; Stanley Thula, **Winston Hide**, Carolyn Williamson. Vertical HIV transmission in South Africa: translating research into policy and practice. *Lancet*. 2002 Mar 23;359(9311):992-3.
21. Wagied Davids, Junaid Gamieldien, David Liberles and **Winston Hide**. Positive selection scanning reveals decoupling of enzymatic activities of carbamoyl phosphate synthetase in *H. pylori*. *Journal of Molecular Evolution*. 2002 Apr;54(4):458-64
22. Vincent VanBuren, Yulan Piao, Dawood B. Dudekula, Yong Qian, Mark G. Carter, Patrick R. Martin, Carole A. Stagg, Uwem C. Bassey, Kazuhiro Aiba, Toshio Hamatani, George J. Kargul, Amber G. Luo, Janet Kelso, **Winston Hide**, and Minoru S.H. Ko.. Assembly, Verification, and Initial Annotation of the NIA Mouse 7.4K cDNA Clone Set. *Genome Research*. 2002 Dec;12(12):1999-2003
23. Kelso J, Visagie J, Theiler G, Christoffels A, Bardien S, Smedley D, Otgaar D, Greyling G, Jongeneel CV, McCarthy MI, Hide T, **Hide W**. eVOC: A Controlled Vocabulary for Unifying Gene Expression Data. *Genome Research*. 2003 Jun;13(6):1222-30
24. Mark I McCarthy, Damian Smedley and **Winston Hide**. New methods for finding disease-susceptibility genes: impact and potential. *Genome Biology*. 2003;4(10):119. Epub 2003 Sep 19
25. Helena Brentani, Otávia L. Caballero, Anamaria A. Camargo, Aline M. da Silva, Wilson Araújo da

Silva Junior, Emmanuel Dias Neto, Marco Grivet, Arthur Gruber, Pedro Edson Moreira Guimaraes, **Winston Hide**, Christian Iseli, C. Victor Jongeneel, Janet Kelso, Maria Aparecida Nagai, Elida Paula Benquique Ojopi, Elisson C. Osorio, Eduardo M.R. Reis, Gregory J. Riggins, Andrew John George Simpson, Sandro de Souza, Brian J. Stevenson, Robert L. Strausberg, Eloiza H. Tajara, Sergio Verjovski-Almeida. Marcio Luis Acencio, Mário Henrique Bengtson, Fabiana F. Bettoni, Walter F. Bodmer, Marcelo R.S. Briones, Luiz Paulo Camargo, Webster Cavenee, Janete M. Cerutti, Luís Eduardo Coelho Andrade, Paulo César Costa dos Santos, Israel Tojal da Silva, Marcos Roberto H. Estécio, Frank B. Furnari, Milton Faria Junior, Pedro A.F. Galante, Gustavo S. Guimaraes, Adriano Jesus Holanda, Edna Teruko Kimura, Maarten R Leerkes, Xin Lu, Rui M. B. Maciel, Elizabeth A. L. Martins, Katlin Brauer Massirer, Analy S.A. Melo, Carlos Alberto Mestriner, Elisabete Cristina Miracca, Leandro Lorenco Miranda, Francisco G. Nobrega, Paulo S Oliveira, Apuã C.M. Paquola, José Rodrigo C. Pandolfi, Maria Inês de Moura Campos Pardini, Fabio Passetti, John Quackenbush, Beatriz Schnabel, Mari Cleide Sogayar, Jorge E Souza, Sandro R. Valentini, Andre C Zaiats. Elisabete Jorge Amaral, Liliane A. T. Arnaldi, Amélia Goes de Araújo, Simone Aparecida de Bessa, David C. Bicknell, Maria Eugenia Ribeiro de Camaro, Dirce Maria Carraro, Helaine Carrer, Alex F. Carvalho, Christian Colin, Fernando Costa, Maria Cristina Ramos Costa, Cyntia Curcio, Ismael Dale Cotrim Guerreiro da Silva, Neusa Pereira da Silva, Márcia Dellamano, Hamza El-Dorry, Enilza Maria Espreafico, Ari José Scattone Ferreira, Cristiane Ayres Ferreira, Maria Angela H.Z. Fortes, Angelita Habr Gama, Daniel Giannella-Neto, Maria Lúcia C.C. Giannella, Ricardo R. Giorgi, Gustavo Henrique Goldman, Maria Helena S. Goldman, Christine Hackel, Paulo Lee Ho, Elza Myiuki Kimura, Luiz Paulo Kowalski, Jose E Krieger, Luciana C. C. Leite, Ademar Lopes, Ana Mercedes S.C. Luna, Alan Mackay, Suely Kazue Nagahashi Mari, Adriana Aparecida Marques, Waleska K. Martins, André Montagnini, Mario Mourão Neto, Ana Lucia T. O. Nascimento, A. Munro Neville, Marina P. Nobrega, Mike J. O'Hare, Audrey Yumi Otsuka, Anna Izabel Ruas de Melo, Maria Luisa Paçó-Larson, Gonçalo Guimarães Pereira, Neusa Pereira da Silva, João Bosco Pesquero, Juliana Gilbert Pessoa, Paula Rahal, Claudia Aparecida Rainho, Vanderlei Rodrigues, Silvia Regina Rogatto, Camila Malta Romano, Janaína Gusmão Romeiro, Benedito Mauro Rossi, Monica Rusticci, Renata Guerra de Sá, Simone Cristina Sant'Anna, Míriam L. Sarmazo, Teresa Cristina de Lima e Silva, Fernando Augusto Soares, Maria de Fátima Sonati, Josane de Freitas Sousa, Diana Queiroz, Valéria Valente, André Luiz Vettore**,** Fabiola Elizabeth Villanova, Marco Antonio Zago, Heloisa Zalcberg. The Generation and Utilization of a Cancer Oriented Representation of the Human Transcriptome, Using Expressed Sequence Tags. *Proceedings of the National Academy of Sciences* U S A. 2003 Nov 11;100(23):13418-23.

26. **Winston Hide**, Damian Smedley, Mark McCarthy and Janet Kelso. Application of eVOC: controlled vocabularies for unifying gene expression data. BIOLOGIES: issue N°11-12. *Comptus Rendus Biology*. 2003 Oct-Nov;326(10-11):1089-96.
27. Isokpehi, RD, **Hide, W.A**. Integrative Analysis of Intraerythrocytic differentially expressed transcripts yields novel insights into the biology of Plasmodium falciparum. *Malaria Journal*, 2003. Nov 14;2(1):38
28. Yun Ping Lim, Jan-Olov Hoog, Phyllis Gardner, Shoba Ranganathan, Siv Andersson, Subramanian Subbiah, Tin Wee Tan, **Winston Hide** and Anthony S. Weiss. The S-star Trial Bioinformatics Course. An online Learning Success. *Biochemistry and Molecular Biology Education* 2003. 31(1): 20-23
29. Alexei A. Sharov, Yulan Piao, Ryo Matoba, Dawood B. Dudekula, Yong Qian, Vincent VanBuren, Geppino Falco, Patrick R. Martin, Carole A. Stagg, Uwem C. Bassey, Yuxia Wang, Mark G. Carter, Toshio Hamatani, Kazuhiro Aiba, Hidenori Akutsu, Lioudmila Sharova, Tetsuya S. Tanaka, Wendy L. Kimber, Toshiyuki Yoshikawa, Saied A. Jaradat, Serafino Pantano, Ramaiah Nagaraja, Kenneth R. Boheler, Dennis Taub, Richard J. Hodes, Dan L. Longo, David Schlessinger, Jonathan Keller, Emily Klotz, Garnett Kelsoe, Akihiro Umezawa, Angelo L. Vescovi, Janet Rossant, Tilo Kunath, Brigid L.M. Hogan Anna Curci, Michele D'Urso, Janet Kelso, **Winston Hide**, and Minoru S.H. Ko. Transcriptome analysis of the early mouse embryo. *Public Library of Science Biology. 2003 Dec; 1(3): e74.*
30. Swart EC, **Hide WA**, Seoighe, C. FRAGS: estimation of coding sequence substitution rates from fragmentary data. Estienne *BMC Bioinformatics* 2004, Jan 29;5(1):8

31. Imanishi T, Itoh T, Suzuki Y, O'Donovan C, Fukuchi S, Koyanagi KO, Barrero RA, Tamura T, Yamaguchi-Kabata Y, Tanino M, Yura K, Miyazaki S, Ikeo K, Homma K, Kasprzyk A, Nishikawa T, Hirakawa M, Thierry-Mieg J, Thierry-Mieg D, Ashurst J, Jia L, Nakao M, Thomas MA, Mulder N, Karavidopoulou Y, Jin L, Kim S, Yasuda T, Lenhard B, Eveno E, Suzuki Y, Yamasaki C, Takeda J, Gough C, Hilton P, Fujii Y, Sakai H, Tanaka S, Amid C, Bellgard M, Bonaldo Mde F, Bono H, Bromberg SK, Brookes AJ, Bruford E, Carninci P, Chelala C, Couillault C, de Souza SJ, Debily MA, Devignes MD, Dubchak I, Endo T, Estreicher A, Eyras E, Fukami-Kobayashi K, Gopinath GR, Graudens E, Hahn Y, Han M, Han ZG, Hanada K, Hanaoka H, Harada E, Hashimoto K, Hinz U, Hirai M, Hishiki T, Hopkinson I, Imbeaud S, Inoko H, Kanapin A, Kaneko Y, Kasukawa T, Kelso J, Kersey P, Kikuno R, Kimura K, Korn B, Kuryshev V, Makalowska I, Makino T, Mano S, Mariage-Samson R, Mashima J, Matsuda H, Mewes HW, Minoshima S, Nagai K, Nagasaki H, Nagata N, Nigam R, Ogasawara O, Ohara O, Ohtsubo M, Okada N, Okido T, Oota S, Ota M, Ota T, Otsuki T, Piatier-Tonneau D, Poustka A, Ren SX, Saitou N, Sakai K, Sakamoto S, Sakate R, Schupp I, Servant F, Sherry S, Shiba R, Shimizu N, Shimoyama M, Simpson AJ, Soares B, Steward C, Suwa M, Suzuki M, Takahashi A, Tamiya G, Tanaka H, Taylor T, Terwilliger JD, Unneberg P, Veeramachaneni V, Watanabe S, Wilming L, Yasuda N, Yoo HS, Stodolsky M, Makalowski W, Go M, Nakai K, Takagi T, Kanehisa M, Sakaki Y, Quackenbush J, Okazaki Y, Hayashizaki Y, **Hide W**, Chakraborty R, Nishikawa K, Sugawara H, Tateno Y, Chen Z, Oishi M, Tonellato P, Apweiler R, Okubo K, Wagner L, Wiemann S, Strausberg RL, Isogai T, Auffray C, Nomura N, Gojobori T, Sugano S. Integrative Annotation of 21 037 Human Genes Validated by Full-Length cDNA Clones. (2004). *Public Library of Science Biology*. 2004 Jun;2(6):e162. Epub 2004 Apr 20
32. Davila AM, Majiwa PA, Grisard EC, Aksoy S, **Hide W**. Response to Hertz-Fowler and Berriman: Continuing tsetse and Trypanosoma genome sequencing projects. *Trends in Parasitology*. 2004 Jul;20(7):309-10
33. Tanino M, Debily MA, Tamura T, Hishiki T, Ogasawara O, Murakawa K, Kawamoto S, Itoh K, Watanabe S, de Souza SJ, Imbeaud S, Graudens E, Eveno E, Hilton P, Sudo Y, Kelso J, Ikeo K, Imanishi T, Gojobori T, Auffray C, **Hide W**, Okubo K. The Human Anatomic Gene Expression Library (H-ANGEL), the H-Inv integrative display of human gene expression across disparate technologies and platforms. *Nucleic Acids Research*. 2005 Jan 1;33 Database Issue:D567-72
34. Serap Aksoy, Matt Berriman, Neil Hall, Masahira Hattori, **Winston Hide**, Michael J. Lehane. A case for *Glossina* genome project. *Trends in Parasitology*, Volume 21, Issue 3, March 2005, Pages 107-111 2005
35. N. Tiffin, J. Kelso, A. Powell, H. Pan, V. Bajic, **W. Hide**. Integration of Text- and Data Mining Using Ontologies Successfully Selects Disease Gene Candidates. *Nucleic Acids Research,* 2005 33(5): 1544-1552.
36. The Transcriptional Landscape of the Mammalian Genome. The FANTOM Consortium. *Science*. September 2;309(5740):1559-63
37. Alexander Mehrle, Heiko Rosenfelder, Ingo Schupp, Coral del Val, Dorit Arlt, Florian Hahne, Stephanie Bechtel, Jeremy Simpson, Oliver Hofmann, **Winston Hide**, Karl-Heinz Glatting1, Wolfgang Huber, Rainer Pepperkok, Annemarie Poustka and Stefan Wiemann. The LIFEdb database in 2006. Nucleic Acids Research 2006 34 (Database Issue):D415-D418; doi:10.1093/nar/gkj139
38. Vladimir B. Bajic, Sin Lam Tan, Alan Christoffels, Christian Schönbach, Leonard Lipovich, Liang Yang, Oliver Hofmann, Adele Kruger, **Winston Hide**, Chikatoshi Kai, Jun Kawai, David A. Hume, Piero Carninci, Yoshihide Hayashizaki. PLoS Genetics Vol. 2, No. 4, e54 DOI: 10.1371/journal.pgen.0020054
39. Nicki Tiffin, Euan Adie, Frances Turner, Han Brunner, Marc van Driel 6) Martin Oti 7) Nuria Lopez-Bigas, Christos Ouzounis, Carolina Perez-Iratxeta, Miguel Andrade-Navarro, Adebowale Adeyemo, Mary Elizabeth Patti, Colin Semple, **Winston Hide**. Computational disease gene identification: a concert of methods prioritises type 2 diabetes and obesity candidate genes. *Nucleic Acids Research in press*.
40. **Martin, DP.**, and Rybicki, EP (2000) RDP: detection of recombination amongst aligned sequences. *Bioinformatics* **16**:562-563.
41. **Martin, DP.**, J. van Harmelen, Rybicki, EP., and C. Williamson. (2000) Software for identifying

recombinant HIV genomes. Proceedings of the XIII international AIDS conference, Durban, South Africa 9-14 July 2000, Monduzzi pp 81-85.

42. Schnippenkoetter, WH., **Martin, DP.**, Willment, JA., and Rybicki, EP. (2001) Forced Recombination Between Distinct Maize Streak Virus Strains. J*ournal of General Virology* **82**:3081-3090.
43. **Martin, DP.**, and Rybicki, EP. (2002) Investigation of Maize streak virus pathogenicity determinants using chimaeric genomes. *Virology* **300**:180-188.
44. **Martin, DP.**, Williamson, C., and Posada, D. (2005) RDP2: Recombination detection and analysis from sequence alignments. *Bioinformatics* **21**:260-262.
45. **Martin, DP.**, Posada, D, Crandall, KA., and Williamson, C. (2005) A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research and Human Retroviruses* **21**:98-102.
46. Tsaousis, A., **Martin DP.**, Posada, D., Ladoukakis, E., and Zouros, E. (2005) Mitochondrial DNA recombination is widespread in animals. *Molecular Biology and Evolution* **22**:98-102.
47. **Martin, DP.**, Van der Walt, E., Posada, D., and Rybicki, EP. (2005) The evolutionary value of recombination is constrained by genome modularity. *PLoS Genetics* **1**:475-479.
48. Poke, FS., **Martin, DP.**, Steane, DA., Vaillancourt, RE., and Reid, JB. (2006) The impact of intragenic recombination on phylogenetic reconstruction at the sectional level in Eucalyptus when using a single copy nuclear gene (cinnamoyl CoA reductase). *Molecular Phylogenetics and Evolution* **39**:160-170.
49. Collins, NE., Liebenberg, J., **de Villiers, EP.**, Brayton, KA., Louw, E., Pretorius, A., Faber, FE., van Heerden, H., Josemans, AI., van Kleef, M., Steyn, HC., van Strijp, MF., Zweygarth, E., Jongejan, F., Maillard, JC., Berthier, D., Botha, M., **Joubert, F.**, Corton, CH., Thomson, NR., **Allsopp, MTEP.**, and **Allsopp, BA.** (2005) The genome of the heartwater agent, /Ehrlichia ruminantium/, contains multiple tandem repeats of actively variable copy number. *Proceedings of the National Academy of Sciences, USA*. **102**:838-843.
50. Collins, NE., Pretorius, A., Van Kleef, M., Brayton, KA., **Allsopp, MT.**, Zweygarth, E. and **Allsopp, BA.** (2003). Development of improved attenuated and nucleic acid vaccines for heartwater. *Developments in Biologicals* **114**:95-110.
51. **De Oliveira T**, Deforche K, **Cassol S**, Salminen M, Paraskevis D, **Seebregts C**, Snoeck J, Janse van Rensburg E, Wensing AMJ, van de Vijver DA, Boucher CA, Camacho R and Vandamme A. (2005) An Automated Genotyping System for Analysis of HIV-1 and other Microbial Sequences. *Bioinformatics*.
52. Doherty R, **de Oliveira T**, **Seebregts C**, Danaviah S, Gordon M and **Cassol S**. (2005) HIV Proteomics Resource: combining HIV-1 protein data with bioinformatics tools. *Retrovirology*, 2:18.
53. Fredman D, Munns G, Rios D, Sjoholm F, Siegfried M, Lenhard B, **Lehvaslaiho H**, Brookes AJ (2004) HGVbase: a curated resource describing human DNA variation and phenotype relationships. *Nucleic Acids Res.* **32**:D516-9.
54. Jason E. Stajich, David Block, Kris Boulez, Steven E. Brenner, Stephen A. Chervitz, Chris Dagdigian, Georg Fuellen, James G.R. Gilbert, Ian Korf, Hilmar Lapp, **Heikki Lehväslaiho**, Chad Matsalla, Chris J. Mungall, Brian I. Osborne, Matthew R. Pocock, Peter Schattner, Martin Senger, Lincoln D. Stein, Elia Stupka, Mark D. Wilkinson, and Ewan Birney (2002) The Bioperl Toolkit: Perl Modules for the Life Sciences. *Genome Res.* **12**(10):1611-1618.
55. **Joubert, AM.**, Louw, AI., **Joubert, F.** & Neitz, AW. (1998) Isolation and characterization of an anticoagulant from the salivary glands of the soft tick, *Ornithodoros savignyi*. *Exp Appl Acarol* **22:**603-19.
56. **Joubert, AM.**, Panzer, A., **Joubert, F.**, Lottering, ML., Bianchi, PC. & **Seegers, JC**. (1999) Comparative study of the effects of polyunsaturated fatty acids and their metabolites on cell growth and tyrosine kinase activity in oesophageal carcinoma cells. *Prostaglandins Leukot Essent Fatty Acids* **61:**171-82.
57. **Seegers, JC.**, **Joubert, AM.**, Panzer, A., Lottering, ML., Jordan, CA., **Joubert, F.**, Maree, JL., Bianchi, P., de Kock, M. & Gelderblom, WC. (2000) Fumonisin B1 influenced the effects of

arachidonic acid, prostaglandins E_2 and A_2 in cell cycle progression, apoptosis induction, tyrosine- and cdc2-kinase activity in esophageal cancer cells. *Prostaglandins Leukot Essent Fatty Acids* **62**:75-84.

58. Joubert, AM., Panzer, A., **Joubert, F.**, Lottering, ML., Bianchi, PC., du Toit, PJ. & Seegers, JC. (2001) Identification of a tyrosine kinase-phosphorylated protein i Arachidonic acid- and Prostaglandin A_2 - treated cells *in vitro*. *Prostaglandins Leukot Essent Fatty Acids* **65**:173-7.
59. **Joubert, F.**, Neitz, A. W. & Louw, AI. (2001) Structure-based inhibitor screening: a family of sulfonated dye inhibitors for malaria parasite triosephosphate isomerase. *Proteins* **45**:136-43.
60. **Seegers, JC.**, Lottering, L., Joubert, AM., **Joubert, F.**, Koorts, A., Engelbrecht, CA. & van Papendorp, DH. (2002) The effects of a small, pulsed DC electric field on ATP and cAMP levels in /in vitro/ and /in vivo/ systems. *Med Hypotheses* **58**:171-176.
61. Van Rensburg, H., Haydon, D., Joubert, F., Bastos, A., Heath, L. & Nel, L. (2002) Genetic heterogeneity in the foot-and-mouth disease virus Leader and 3C proteinases. *Gene* **289**:19-29.
62. Birkholtz, LM., **Joubert, F.**, Neitz, AWH. & Louw, AI. (2003) Comparative properties of a three-dimensional model of *Plasmodium falciparum* ornithine decarboxylase. *Proteins* **50**:464-473.
63. Birkholtz LM, Wrenger C, **Joubert F**, Wells GA, Walter RD, Louw AI. (2003) Parasite-specific inserts in the bifunctional S-adenosylmethionine decarboxylase/ornithine decarboxylase of *Plasmodium falciparum* modulate catalytic activities and domain interactions. *Biochem J.* **377**:439-448.
64. Joubert, A, Maritz, C, **Joubert, F** (2005) Influence of prostaglandin A2 and 2-methoxyestradiol on Bax and Bcl-2 expression levels in cervical carcinoma cells. *Biomed. Res.* **26**:131-134.
65. Wells GA, Birkholtz LM, **Joubert F**, Walter RD, Louw AI. (2006) Novel properties of malarial S-adenosylmethionine decarboxylase as revealed by structural modelling. *J. Mol. Graph. Model.* 24:309-318.
66. **De Beer TA**, Louw AI, **Joubert F**. (2006) Elucidation of sulfadoxine resistance with structural models of the bifunctional Plasmodium falciparum dihydropterin pyrophosphokinase-dihydropteroate synthase. In press. *Bioorg. Med. Chem.*
67. **Martin DP**, Willment, JA, Billharz, R, Velders, R, Odhiambo, B, Njuguna, J, James, D, and Rybicki, EP. (2001) Sequence Diversity and Virulence in Zea mays of Maize Streak Virus Isolates. Virology 288:247-255. (90% of laboratory work and 100% of authorship).
68. Willment JA, **Martin DP**, van der Walt E and Rybicki, EP. (2002) Biological and genomic characterisation of Maize streak virus isolates from wheat. Phytopathology 92:81-86.
69. Heath L, **Martin DP**, Williamson AL, and Rybicki EP. (2004) Evidence of unique genotypes of Psittacine beak and feather disease virus in southern Africa. J. Virol. 78:9277-9284.
70. Shepherd D, **Martin DP**, McGivern D, Boulton M, Thomson J, and Rybicki E. (2005) A three-nucleotide mutation altering the Maize streak virus Rep pRBR-interaction motif reduces symptom severity in maize and partially reverts at high frequency without restoring pRBR-Rep binding. J. Gen. Virol. 803-813.
71. Delatte H, **Martin DP**, Naze F, Goldbach R, Reynaud B, Peterschmitt M, and Lett JM. (2005) South West Indian Ocean islands tomato Begomovirus populations represent a new major monopartite begomovirus group. J. Gen. Virol. 1533-1542 (10% of laboratory work, 30% of authorship).
72. Fanigliulo A, Comes S, Pacella R, Harrach B, **Martin DP** and Crescenzi, A. (2005) Characterisation of Potato virus Y nnp strain inducing veinal necrosis in pepper: a naturally occurring recombinant strain of PVY. Arch. Virol. 150:709-720.
73. Varsani, A, Heath L, van der Walt, E, Rybicki EP, Williamson AL and **Martin DP**. (2006). Evidence of ancient pappillomavirus recombination. In Press J. Gen. Virol.
74. Shepherd, DN, **Martin, DP**, Varsani, A, Thomson, JA, Rybicki, EP and Klump, HH (2006) Restoration of native folding through reverse mutations: an indication of a new mechanism of epigenetics. In Press Arch. Biochem. Biophys.
75. Scheffler K, **Seoighe C.** (2005) A bayesian model comparison approach to inferring positive selection. Molecular Biology & Evolution. 2005 Dec;22(12):2531-40.

76. **Seoighe C.**, Scheffler, K. (2005) Low power to detect asymmetric divergence of duplicated genes In *Proceedings of the RECOMB 2005 International Workshop on Comparative Genomics* Eds McLysaght, A. & Huson D.H.
77. **Seoighe C**, Gehring C, Hurst LD. (2005) Gametophytic Selection in Arabidopsis thaliana Supports the Selective Model of Intron Length Reduction. PLoS Genet. 2005 Aug 5;1(2):e13.
78. Nembaware V, Wolfe KH, Bettoni F, Kelso J**, Seoighe C.** Allele-specific transcript isoforms in human. FEBS Lett. 2004 Nov 5;577(1-2):233-8.
79. **Seoighe, C.**, Gehring, C. Genome duplication led to highly selective expansion of the *Arabidopsis thaliana* proteome *Trends in Genetics* 2004 Oct;20(10):461-4.
80. Grobler, J., Gray, C.M., Rademeyer,C., **Seoighe, C.**, Ramjee, G., Karim,S.A., Morris, L. Williamson, C. The incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of subtype C infected female sex-workers. *Journal of Infectious Disease*, 2004: 190.
81. Nembaware, V, Said, M, **Seoighe, C.**, Gehring, C.A. Molecular mimicry of a plant signaling molecule enables pathogen control of plant homeostasis *BMC Evolution Biol,* 2004 Mar 24;4(1):10.
82. Swart, E.C. , Hide, W.H., **Seoighe, C.** FRAGS: estimation of coding sequence substitution rates from fragmentary data (*BMC Bioinformatics,* 2004 Jan 29;5(1):8.
83. **Seoighe, C.** Turning the clock back on ancient genome duplication. Curr Opin Genet Dev. 2003 Dec;13(6):636-43.
84. **Seoighe, C.** Johnston, C.R. Shields, D.C. Significantly different patterns of amino acid replacement after gene duplication as compared to after speciation Mol Biol Evol. 2003 Apr;20(4):484-90.
85. Nembaware, V. Crum, K. Kelso, JF, **Seoighe, C.** Impact of the presence of paralogues on sequence divergence in a set of mouse-human orthologues Genome Res. 2002 Sep;12(9):1370-6.
86. Ludidi N.N., Heazlewood J.L., **Seoighe C.**, Irving H.R. and Gehring C.A. Expansin-like molecules: Novel functions derived from common domains. Journal of Molecular Evolution 2002 May;54(5):587-94.

7. Contacts:

Basil ALLSOPP
Department of Veterinary Tropical Diseases
Faculty of Veterinary Science
University of Pretoria
Private Bag X04
Onderstepoort 0110
South Africa
Tel: (27) 125298426
basil.allsopp@up.ac.za

Sharon CASSOL
Molecular Virology and Bioinformatics Unit
Department of Medical Virology
University of Pretoria
South Africa
sharon.cassol@up.ac.za

Scott HAZELHURST
Wits Computational Molecular Biology Programme
University of the Witwatersrand
Johannesburg
Private Bag 3
2050 Wits
South Africa
Tel: (27) 117176181
scott@cs.wits.ac.za

Winston HIDE
South African National Bioinformatics Institute
University of the Western Cape
Private Bag X17
7535 Bellville
South Africa
Tel: (27) 219593645
Fax: (27) 219592512
winhide@sanbi.ac.za

Dan JACOBSON
National Bioinformatics Network
University of the Western Cape
Private Bag X17
7535 Bellville
South Africa
Tel: (27) 219592991
Fax: (27) 219593573
dan@nbn.ac.za

Fourie JOUBERT
Bioinformatics and Computational Biology Unit
University of Pretoria
Lynwood rd
Pretoria
South Africa
Tel: (27) 124205802
fjoubert@postino.up.ac.za

Colin KENYON
Council for Science and Industrial Research (CSIR)
Private Bag x2
1645 Modderfontein
South Africa
Tel: (27) 116052702
Fax: (27) 116083020
ckenyon@csir.co.za

Heikki LEHVÄSLAIHO
South African National Bioinformatics Institute
University of the Western Cape
Private Bag X17
7535 Bellville
South Africa
Tel: (27) 219592096
Fax: (27) 219592512
heikki@sanbi.ac.za

Darren MARTIN
Institute of Infectious Diseases and Molecular Medicine
University of Cape Town
Observatory
Cape Town

South Africa
darren@science.uct.ac.za

Chris SEEBREGTS
Research Information Systems Division
South African Medical Research Council
Cape Town
South Africa
chris.seebregts@mrc.ac.za

Cathal SEOIGHE
Department of Molecular and Cell Biology
University of Cape Town
Private Bag
Rondebosch 7701
South Africa
Tel: (27) 214066837
Fax: (27) 216897573
cathal@science.uct.ac.za

Vladimir B BAJIC
South African National Bioinformatics Institute
University of the Western Cape
Private Bag X17
7535 Bellville
South Africa
Tel: (27) 219592360
Fax: (27) 219592512
vlad@sanbi.ac.za

Sudan

1. Institutions:

1.1 Commission for Biotechnology & Genetic Engineering - Khartoum:

The Commission for Biotechnology and Genetic Engineering was established in 9/2/1993 as research unit. In addition to research activities it acts as focal point for the International Center for Biotechnology and Genetic Engineering.

The commission conducts researches in order to play a part in solving economical, environmental, health and nutritional problems using modern research techniques with an emphasis on the applied researches in these areas.

The laboratories were well furnished with the essential equipments and the catalyst infrastructure to facilitate emergence of a successful for research. The Commission equipped with a computer center and information to serve as informatics and Digital library.

2. Training Activities:

2.1 Long-term training activities:

No short-term training in bioinformatics.

2.2 Short-term training activities:

No long-term training in bioinformatics.

3. Research Interest and Activities:

1. Plant Genetic Transformations

2. Molecular Population Genetics
3. Detection of human and Animals diseases
4. Breast Cancer-specific protein marker
5. Phytochemical
6. Genomic map
7. Bioremediation
8. Tissue Culture.

4. Expected feedback from CNRS-NEPAD partnership:

We believe that, due to the lack of bioinformaticists, this field should be develop in Sudan through these priorities:

1. Long term training at level of M.Sc and PhD
2. Establish bioinformatics laboratory
3. Organize short courses in bioinformatics

5. Web site and links:

www.Geocity.cbge.com

6. Publications:

1. Intisar E;Yassir O.; Mubarak M; **Khalil M.**; Suad M.; Khitma H. (2002). Studies on the prevalence of sleeping sickness in Southern Sudan. 4 th Annual EANETT work-shop –Tororo, Uganda.
2. **Khalil MK.**; AA. Ahmed; 1. E. El Rayh (2003). Prevalence of *Toxoplasma gondii* in camels and their herders in three different ecologically areas in Sudan. Endemic and infectious diseases conference – Khartoum – 2-6/ May 2003.
3. **Khalil M**; Atif E; Khitma H; Mubarak M ; Intisar E. (2003).Biochemical changes in sleeping sickness sernm samples detected by CATT in Sudan. Accepted in ICTRC Conference In South Africa (September 2003).
4. **Khalil M.**; Atif E; Mubarak M.; Yassir O; Khitma H; Intisar E. (2003). Application of card agglutination test for trypanosomasis (CATT) and card indirect agglutination antigen test (CIATT) for detection of camels trypanosomasis in West Sudan. Accepted in ICTRC Conference In South Africa (September 2003).
5. Intisar E. Khitma H; Mubarak M.; **Khalil M.**; Yassir O; Suad M.; Nabil A. (2003). Updating the situation of sleeping sickness in Bahr EL Jabel State, Southern Sudan. Accepted in ICTRC Conference In South Africa (September 2003).
6. **Khalil MK.**; AA. Ahmed; 1E. El Rayh. February (2004). Some Epidemiological Studies on Toxoplasmosis in Khartoum State. M. Sc. Thesis – Faculty of Science- University of Khartoum.
7. Intisar E. Khitma H; Mubarak M.; **Khalil** M.Kalil ; Yassir O; Suad M.; Nabil A. (2003). Updating the situation of sleeping sickness in Bahr EL Jabel State, Southern Sudan. ISCTRC ; Publication No. 122.
8. **Khalil MK**; Atif E; Khitma H; Mubarak M ; Intisar E. (2003).Biochemical changes in sleeping sickness sernm samples detected by CATT in Sudan. ISCTRC ; Publication No. 122.
9. **Khalil MK**; A.A . Ahmed; 1. E. El Rayh (2003). Prevalence of Toxoplasma gondii in camels and their herders in three different ecologically areas in Sudan . Endemic and infectious diseases conference – Khartoum – 2-6/ May 2003.

7. Contacts:

Khalil Mohamed KHALIL
National Research Center
Commission for Biotechnology & Genetic Engineering

Khartoum 11111
P.O.Box 2404
Tel: (249) 918272648
khali72@yahoo.co.uk

Moawia Elfaki YAHIA
Department of Computer Science
University of Khartoum
11115 Khartoum
Sudan
Tel: (249) 11275987
meyahia@hotmail.com

Tunisia

1. Institutions:

1.1 Institut Pasteur de Tunis (IPT):

The IPT is a research institution founded in 1883. IPT is under the supervision of the Ministry of Health and is part of the *Université El Manar* of Tunis (Ministry of high Education). The missions of the institute are: Public Health Laboratory activities (PHL), Research on infectious diseases, and R/D on vaccines. Research programs are mainly oriented towards local health problems such as leishmaniais, viral hepatitis, and scorpion venoms.

The group of Bioinformatics and Modelling of the IPT is hosted by the *Laboratoire d'Immunopathologie Vaccinologie et Génétique Moléculaire* (LIVGM), and exists since the beginning of 2005. Its present research activities include: genome annotation, EST clustering and modelling of the host/parasite response to Leishmania infection. It consists of two senior scientists, two PhD students and one MSc student.

1.2 Centre de Biotechnology de Sfax (CBS):

Bioinformatics activity started at CBS in 2001 with the setting-up of a research and service unit of bioinformatics. This unit currently includes one senior researcher, one engineer and four Phd students. Activities include sequence annotation (service) and three research programs: *ab initio* prediction of short eukaryote genes, statistical modelling by Bayesian networks approach of signal transduction pathways and statistical analysis of human sequence variation data (haplotype reconstruction and linkage disequilibrium). Activities of the Bioinformatics unit could be found at the website: http://www.cbs.rnrt.tn/ and the research activity report is available under request to Bioinformatics@cbs.rnrt.tn. Although the computing facilities are good, there is still a need for trained human resources to strengthen bioinformatics capacities at CBS, particularly in structural bioinformatics.

2. Training Activities:

2.1 Long-term training activities:

A 2 years *Diplôme d'Etudes Supérieures Spécialisées* of bioinformatics (DESS) is organised since more than two years by the University of Manouba and the IPT at the *École Nationale des Sciences Informatiques* (Tunis). This DESS is a very good start. A room with 15 computers was made available for the students. But unfortunately, the program of this master needs to be revised because problems exist at different levels due to the limited number of teachers competent in the field. This year we have installed Linux and several bioinformatics tools in all computers soon. Next year we will try to ameliorate the program content. There is a crucial need for an external support of senior bioinformaticists (visiting professors).

2.2 Short-term training activities:

The CBS (Sfax) have started a national one-week practical course of bioinformatics since 2001. The course took place at a computer room at the faculty of Medicine of Sfax with 15 computers under Linux and Windows systems. Each edition of the course brought an improvement and the course was open to African colleagues. We have had one Student from Algeria and one from Ivory Coast. Most of the participants to this course were PhD Students and senior researchers. An international theoretical and practical course on “Bioinformatics and Genome Data Analysis” was also organized from March 24 to April 02 2005, by the CBS in collaboration with the *Institut Pasteur* (Paris) and support from many international organizations (ICGEB, ICRO, IUBMB, EMBO). At the academic level, a 15h course of Bioinformatics was introduced at the master level since 2003 to the Master of “Molecular and cellular Biology” at the faculty of Science of Sfax.

A two weeks winter school of biomathematics is also organised every two years in Tunis.

3. Research Interest and Activities:

The following are the present areas of research interest:

1. Host-pathogen protein interaction (IPT)
2. Genomes annotation OR analysis of genome data (IPT, CBS)
3. Development of genes prediction programs (CBS)
4. Analysis of expression data (based on microarray or EST data) (IPT)
5. Genes regulation (IPT, CBS)
6. Annotation of protein families (IPT, CBS)
7. Modelling of protein interaction networks in signal transduction pathways (IPT, CBS)
8. Mathematical modelling of cellular mechanisms: systems biology (IPT)
9. Statistical analysis of DNA variation (STR, SNP) data (CBS)
10. Recombination for HIV retroviruses (IPT)

4. Expected feedback from the CNRS-NEPAD partnership:

We would like to highlight our interest to participate in the following personnel development projects, namely:

1. Short training of M.Sc and PhD students
2. Sending and receiving Post-doc fellows
3. Establishing regular African short courses in bioinformatics and genome data analysis by supporting existent efforts and new initiatives
4. Reinforcing the present teaching team of the DESS of bioinformatics of Tunis by visiting senior bioinformaticists
5. Funding research collaborations with African laboratories involved in bioinformatics and genome data analysis
6. Developing South-North and South-South collaborations
7. Setting up an African network of excellence in the field of bioinformatics. Such network could help other countries in their capacity building in this field
8. Financing the accommodation PhD students and Post-Docs for short-term visitors abroad.

5. Web site and links:

http://www.cbs.rnrt.tn

6. Publications:

1. Adjaye J, R Herwig, D Herrmann, W Wruck, **A Benkahla**, TC Brink, M Nowak, JW Carnwath, C Hultschig, H Niemann, and H Lehrach. (2004) Cross-species hybridisation of human and bovine

orthologous genes on high density cDNA microarrays. *BMC Genomics* **5**:83.

2. International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431:**931-945. (Consortium authors: **A Benkahla** is in the German group)
3. The International Chimpanzee Chromosome 22 Consortium. (2004) DNA sequence and comparative analysis of chimpanzee chromosome 22. *Nature* **429:**382-388. (Consortium authors: **A Benkahla** is first of the German group)
4. The International Chimpanzee Chromosome 22 Consortium. (2003) Human versus chimpanzee chromosome-wide sequence comparison and its evolutionary implication. *Cold Spring Harb. Symp. Quant. Biol.* **68:**455-60.
5. The HSA21 expression map initiative. (2002) A gene expression map of human chromosome 21 orthologues in the mouse. *Nature* **420:**586-590. (Consortium authors: **A Benkahla** is first of the German group)
6. **Benkahla A**, DL Robertson, E Beaudoing, and J-M Claverie. (2000) The detection of putative recombinant sequences using BLAST. *German Conference on Bioinformatics 2000.*
7. Ewing RM, **A Ben Kahla**, O Poirot, F Lopez, S Audic, and J-M Claverie. (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Research* **9:**950-959.
8. Kharrat N, Ayadi I. and **Rebaï A.** (2006) Sample size computation for association studies using case-parents design. *Journal of Genetics.* In press.
9. Aifa S, Miled N, Frikha F, Aniba MR, Svensson SP and **Rebai A.** (2006) Electrostatic interactions of peptides flanking the tyrosine kinase domain in the epidermal growth factor receptor (EGFR) provides a model for intracellular dimerization and autophosphorylation. *Proteins: Structure, Function and Bioinformatics.* 62:1036-1043.
10. Rebai M, Kharrat N, Ayadi I and **Rebai A.** (2006) Haplotype structure of five SNPs within the ACE gene in the Tunisian population. *Annals of Human Biology*, In press.
11. Ayadi I, Ammar-Keskes L and **Rebai A.** (2006) Haplotypes for 13 Y-chromosomal STR loci in South Tunisian population (Sfax region). *Forensic Science International.* In press.
12. Mahfoudh-Lahiani N, **Rebaï A.**, Makni H. (2006) Data for 10 autosomal STR markers in south Tunisian population. *Forensic Science International.* In press.
13. Masmoudi K, **Rebai A,** Ellouz R. (2006) AFLP and SSR fingerprinting to evaluate genetic diversity among bread wheat cultivars in Tunisia. *Cereal Res. Comm.* In press.
14. N. Grati-Kamoun, F. Lamy Mahmoud, **A. Rebaï***, A. Gargouri, O. Panaud and A. Saar. (2006) Genetic diversity assessment of Tunisian olive tree (*Olea europaea L.*) cultivars using AFLP markers. (*: corresponding author). *Genetic Resources and Crop Evolution.* 53:265-275.
15. Kharrat N., Rebaï M. and **Rebai A.** (2006) Linkage disequilibrium in the human genome and its applications: between myth and reality. *Archives de l'Institut Pasteur de Tunis.* In press.
16. Tounsi S, Aoun AE, Blight M, **Rebai A**, Jaoua S. (2006) Evidence of oral toxicity of *Photorhabdus temperata* strain K122 against *Prays oleae* and its improvement by heterologous expression of *Bacillus thuringiensis* cry1Aa and cry1Ia genes. *J Invertebr Pathol.* In press.
17. Hadj Kacem H, **Rebai A**, Kaffel N, Abid M et Ayadi H. (2006) Evidence for linkage and association between autoimmune thyroid diseases and the 18q12-q21 region in a large Tunisian family. *International Journal of Immunogenetic*s, 33:25-32.
18. Tounsi S, Dammak M, **Rebaï A**. and Jaoua S. (2005) Response of larval *Ephestia kuehniella* (Lepidoptera: Pyralidae) to individual *Bacillus thuringiensis kustaki* toxins and toxin mixtures.

Biological Control. 35:27.31.

19. Maalej A, Petit-Teiexiera E, Michou L, **Rebaï A**, Cornelis F and Ayadi H. (2005) Association study of VDR gene with Rheumatoid Arthritis in the French population. *Genes and Immunity.* 6:707-711.
20. Choura M and **Rebai A.** (2005) Identification and characterization of new members of vacuolar H^+-pyrophosphatase family from *Oryza sativa* genome. *Russian Journal of Plant Physiology.* 52:926-930.
21. Medini M., Hamza S, **Rebai A** and M. Baum. (2005) Analysis of Genetic Diversity in Tunisian Durum Wheat Cultivars and related wild species by SSR and AFLP markers. *Genetic Resources and Crop Evolution,* 52:21-31.
22. Belguith H, Hajji S, Salem N, Charfeddine I, Lahmar I, Amor MB, Ouldim K, Chouery E, Driss N, Drira M, Megarbane A, **Rebai A**, Sefiani A, Masmoudi S, Ayadi H. (2005) Analysis of GJB2 mutation: evidence for a Mediterranean ancestor for the 35delG mutation. *Clin. Genet.* 68:188-189.
23. Maalej A, Mbarki F, **Rebai A**, Karray F, Jouida J, Abid M, Ayadi H. (2004) Evidence of association between FKBP1B and thyroid autoimmune disorders in a large Tunisian family. *Autoimmunity.* 37:237-239.
24. Maalej A, **Rebai A**, Ayadi A, Jouida J, Makni H, Ayadi H. (2004) Allelic structure and distribution of 103 STR loci in a Southern Tunisian population. *Journal of Genetics*. 83:65-71.
25. Ayadi H, Kacem HH, **Rebai A**, Farid NR. (2004) The genetics of autoimmune thyroid disease. *Trends in Endocrinol and Metabolism.* 15:234-239.
26. Bougacha-Elleuch N, **Rebai A**, Mnif M, Makni H, Bellassouad M, Jouida J, Abid M, Ayadi H**.** (2004) Analysis of MHC genes in a Tunisian isolate with autoimmune thyroid diseases: implication of TNF -308 gene polymorphism. *Journal of Autoimmunity*. 23:75-80.
27. Masmoudi S, Charfedine I, Rebeh I, **Rebai A**, Tlili A, Ghorbel A, Belguith H, Petit C, Drira M, Ayadi H. (2004) Refined mapping of the autosomal recessive non-syndromic deafness locus DFNB13 using eight novel microsatellite markers. *Clinical Genetics*. 66:358-364.
28. Hamza S, Benhamida W and **Rebaï A**. (2004) SSR-based Genetic diversity Assessment among tunisian barley cultivars and relationship with morphological traits. *Euphytica.* 135:107-118.
29. H. Hadj Kaçem, **A. Rebaï**, N. Kaffel, S. Masmoudi, M. Abid and H. Ayadi. (2003) PDS is a new susceptibility gene for autoimmune thyroid diseases: association and linkage study. *Journal of Clinical Endocrinology and Metabolism,* 88:2274-80.
30. C. Bouchlaka, S. Abdlehak, A. Amouri, H. BenAbid, S. Hadiji, M. Frikha, T. Ben Othman, F. Amri, H. Ayadi, M. Hachicha, **A. Rebaï**, A. Saad, K. Dellagi & The Tunisian Fanconi Anemia Study Group. (2003) Fanconi anemia in Tunisia: high Prevalence of group A and identification of new *FANCA* mutations. *J Hum Genet.* 48:352-61.
31. Maalej A., Bougacha N, **Rebaï, A**. (2001) Lack of linkage and association of autoimmune thyroid diseases with the CTLA-4 gene in a large tunisian family. *Human Immunology*, 62:1245-1250.
32. **Rebai A**. (2000) Linkage analysis for complex diseases: a new life for an old method. *Arch Inst Pasteur Tunis*, 77:25-35.
33. **Rebaï A**. and Goffinet B. (2000) More about QTL mapping with diallel designs. *Genetical Research*, 75:243-247.
34. Hamza S., Slama S., **Rebaï A.**, Rezgui S., Kahl G. and Harrabi M. (2000) Pathotype variation for representative genotypes of *Ascochyta rabiei* in the Beja region. *Journal of Plant Pathology*, 82:23-28.
35. El Attari H., **Rebaï A.**, Hayes P.M., Barrault G., Dechamp-Guillaume G. et Sarrafi A. (1998) Potential of doubled-haploid lines and localization of quantitative trait loci (QTL) for partial

resistance to bacterial leaf streak (*Xanthomonas campestris* pv. *hordei*) in barley. *Theoretical and Applied Genetics*, 96:95-100.

36. **Rebaï A.**, Blanchard P., Perret D., Vincourt P. (1997) Mapping quantitative trait loci controlling silking date in a diallel cross of maize. *Theoretical and Applied Genetics*, 95:451-459.
37. **Rebaï A**. (1997) Comparison of methods for regression interval mapping in QTL analysis with non-normal traits. *Genetical Research*, 69:69-74.
38. **Rebaï A**. et Goffinet B. (1996) Correction: power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theoretical and Applied Genetics,* 92:128-129.
39. Doerge R. et **Rebaï A**. (1996) Significance thresholds for QTL interval mapping tests. *Heredity*, 76:459-464.
40. **Rebaï A.**, Goffinet B. et Mangin B. (1995) Comparing power of different methods for QTL detection. *Biometrics*, 51:87-99.
41. **Rebaï A.**, Goffinet B. et Mangin B. (1994) Approximate thresholds of interval mapping tests for QTL detection. *Genetics*, 138:235-240.
42. Goffinet B. et al. (Groupe MMM). (1994) Méthodes mathématiques pour l'étude des gènes contrôlant des caractères quantitatifs. *Génétique, Sélection, Evolution*, 26, Suppl, 9s-20s.
43. Mangin B., Goffinet B. et **Rebaï A**. (1994) Constructing confidence intervals for QTL location. *Genetics*, 138:1301-1308.
44. **Rebaï A**. et Goffinet B. (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theoretical and Applied Genetics*. 86:1014-1022.

7. Contacts:

Alia BENKAHLA
Group Bioinformatics and Modelling
Laboratoire d'Immunopathologie, Vaccinologie et Génétique Moléculaire
Institute Pasteur of Tunis
13, place Pasteur BP74 Belvédère
1002-Tunis
Tunisia
Tel: (216) 71 843755 Ext: 464
Alia.Benkahla@pasteur.rns.tn

Ahmed REBAÏ
Service de Bioinformatique
Centre de Biotechnology de Sfax
BP K
3038 Sfax
Tunisia
Tel: (216) 98817294
Tel: (216) 74440816
ahmed.rebai@cbs.rnrt.tn

B. African Bioinformaticists Mailing List:

TO: sysals@yahoo.fr, sanfom@oncho.oms.bf, ekazae@yahoo.fr, e.villiers@cgiar.org, dmasiga@icipe.org, mulaafj@uonbi.ac.ke, snyanjom@icipe.org, tm.shah@africabiosciences.org, baber@mrtcbko.org, niare@mrtcbko.org, djibril_mrtc@yahoo.fr, sidy@mrtcbko.org, mmaouene@indiana.edu, mmaouen@ensat.ac.ma, efadebiyi@yahoo.com, cobewaji@yahoo.com, erah@uniben.edu, onash@wabw.org, oyekan.nash@gmail.com, wuyepsn@yahoo.com, basil.allsopp@up.ac.za, sharon.cassol@up.ac.za, scott@cs.wits.ac.za, winhide@sanbi.ac.za, dan@nbn.ac.za, fjoubert@postino.up.ac.za, ckenyon@csir.co.za, heikki@sanbi.ac.za, darren@science.uct.ac.za, chris.seebregts@mrc.ac.za, cathal@science.uct.ac.za, vlad@sanbi.ac.za, khali72@yahoo.co.uk, Alia.Benkahla@pasteur.rns.tn, ahmed.rebai@cbs.rnrt.tn

CC: Philippe.Solano@mpl.ird.fr, tulio.deoliveira@zoology.oxford.ac.uk, jaco.deridder@up.ac.za, kmatthee@csir.co.za, mireilledosso@yahoo.fr, david79tr@yahoo.fr, hamid.barakat@pasteur.ma, s_mesmoudi@yahoo.fr, rgouy@voila.fr, aboutabi@yahoo.fr, a.aboutabit@ensat.ac.ma, db@interchange.ubc.ca, jkissing@uga.edu, gnbrain@yahoo.com, tm.shah@cgiar.org