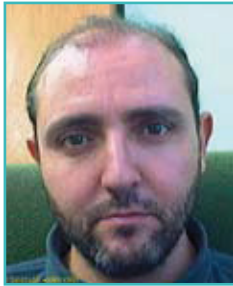# SYNAPSES:

Bridging the gap between Biologists and Bioinformaticians

2nd international Workshop on Pattern Discovery in Biology, Covenant University, Ota, Nigeria

José R. Valverde[1] and Ezekiel F. Adebiyi[2]

[1] EMBnet/CNB, Centro Nacional de Biotecnología, CSIC. Madrid, Spain
[2] Department of Computer and Information Sciences, College of Science and Technology, Covenant University, Ota, Nigeria

This has been the second Bioinformatics Workshop organized at Covenant University with the aim of bringing together Biologists and Bioinformaticians. The main goal of these workshops is to provide a way for specialists in both fields to get together and exchange knowledge and experience. The workshop is structured around a series of talks that introduce key concepts (of Biology for Bioinformaticians and of Bioinformatics for Biologists) and a series of hands-on practical sessions where relevant tools are demonstrated and tried first-hand by assistants. This basic scheme, provides a nurturing medium where specialists in both areas can talk, interact and get closer, shedding the seed of future collaborations.

The second Workshop on Pattern Discovery in Biology took place in the College of Science and Technology, Covenant University, Ota, Nigeria, from July the 6th to July 11th. Its location, very close to Lagos makes it an ideally well communicated place easy to reach. Covenant University, funded by the World Mission Agency (WMA) Intl Inc strives to become a major leading institution in the world; it is located in the broad, quiet campus of Canaanland and enjoys many modern facilities. This welcoming and relaxing environment provides a warm environment for Workshop participants to mix and interact as well as the necessary technical support needed to be able to test and get acquainted with the tools demonstrated.

The workshop started with a warm Opening Ceremony where the Vice Chancellor, the Registrar, the Dean, College of Science and Technology and the Chairman, Workshop Steering Committee welcomed all participants and that helped set the key note for the rest of the sessions, stressing the main points of the meeting: easing mutual understanding, closing the interdisciplinary gap and promoting cooperation and lasting collaborations between participants. A keynote address helped introduce the main topics that were covered in the remaining of the sessions, and the leitmotif for the workshop, which this time centred on the need to partially automate using Bioinformatics, the design of new drugs to fight major diseases in the region.

The main body of the Workshop was structured around enlightening talks and practical sessions. The role of the talks was to introduce the key points of major topics both from the perspective of Biologists and the perspective of Bioinformaticians. Since the main goal is to bridge the gap and allow efficient interdisciplinary communications, the contents must go beyond basic introductions and get into enough advanced details to allow mutual comprehension. Reaching this level of detail was the major challenge for this workshop, and imposed relevant restrictions in the way talks needed to be structured (so as to cover both introductory and advanced topics) and the effort required by participants to follow the presentations. These difficulties were eased out by providing enough room for discussion after the talks, where participants could engage

The college of Science and Technology.

Workshop participants at the Closing Ceremony.

in conversations where not only obscure points were illustrated but where practical and personal points of interest and individual applications could be brought up and openly discussed and where free brainstorming was welcome.

The descriptive talks were complemented by a series of practical sessions, in a ratio close to 50%. The practical sessions helped address several complementary goals: first, they served to illustrate the points made during descriptive presentations, further clarifying key concepts and demonstrating how these are experienced in real practice; as a second benefit, they provided first hand personal experience in the use of the tools of the trade and the interpretation of the results they yield, ensuring every participant left with the expertise needed to identify and study biologically significant patterns generated by bioinformaticians and to use them to build predictive models to guide subsequent research; and third, they provided a joint environment for practical interaction among participants from different disciplines, in a space where they could share experiences and exchange comments in an actual, real set up similar to the one they find in their day-to-day work at the lab. To ensure achievement of these goals, a large computer room was used, where every participant could work with one computer (or even more if desired) and where they could freely interact with computers, speakers and their colleagues while working on a practical problem selected from real life.

As a part of our goal of making participants proficient in the field, and ensuring they can continue working once they go back home, we selected to demonstrate major points choosing

whenever possible from a large array of services publicly available on the World Wide Web, as these are the ones they will most likely want to use, and complemented them in parallel with equivalent demonstrations using locally installed tools and programs (e.g. Using Clustal on the web and ClustalX [1] and PhyloWin [2] on the desktop) so they could realize the limitations of web based approaches and the richness of the local applications approach.

In order to ensure participants had access to all the applications needed, we used a 4GB USB key/pen that is continuously being developed at EMBnet/CNB, in Spain[3]. The USB key/pen contains a minimalistic Linux system with a lightweight, yet complete, user interface (Puppy Linux) and can be either booted directly, run as a virtual machine under other system (windows, Mac, Linux) or as a slower emulated machine (if one lacks administrative privileges). The system has been populated with a large number of software packages, including publicly accessible systems like EMBOSS [4] or PHYLIP [5] as well as free -yet usually non distributable- systems like TRITON [6] or MODELLER [7] for which we had previously obtained permission from the authors to include and distribute them for educational purposes. This emulated machine proved to work acceptably well on the resources of the computers available (256MB RAM) even when dealing with complex problems like protein homology modelling, protein-ligand docking and even quantum mechanics modeling. As this key is still a work in progress, we shall not get into more details of its contents, configuration and availability here. The key and its contents were made available

Dr. Jose   and Dr. Adebiyi with some participants at the beach.

A few participants at the beach.

to participants, and a number of keys were also distributed at the end of the Workshop.

The workshop started with a general introduction to the field of Bioinformatics, putting a large stress in explaining major Biological and Computer Science concepts, the way they are mapped out on each field and the differences in terminology and conceptual approach to a common problem. After these basic concepts were illustrated, participants could move into the major tasks planned for the workshop: the generic problem of pattern discovery in Bioinformatics and its relationship with the close Biological concepts of similarity and homology, and then by extension its many applications in sequence comparison, evolution, genomics, genetic engineering and gene manipulation, proteomics, regulomics, functional prediction, etc..

As we have already mentioned, the leitmotif selected for this second international workshop was the problem of drug discovery and design, and understanding of drug resistance mechanisms. This is a topic of major interest in the region, where multi-drug resistant strains of serious diseases have spread due to a variety of reasons. The second part of the Workshop dealt therefore with these topics in more details, jumping from traditional sequence comparison into three-dimensional pattern discovery, homology modeling, understanding of the interactions between host and microorganism and between enzymes and substrates, and the methods available to analyse drug effects and drugs resistance.

According to the guiding line for the Workshop, the practical sessions were organized around a real-life example: resistance of *Mycobacterium*

tuberculosis to Isoniazid. We selected this example as there is a well understood body of knowledge in the field, it is one of the major diseases both in developed and developing countries, multiple-resistant strains have been characterized and identified as a major health challenge by WHO, and its molecular understanding is still a topic of great current interest. Using as a reference a very recent paper published by Wahab et al. [8] analyzing the mechanism of resistance on *InhA* (PDB 1ZID), we reproduced the work described in the paper on a similar system, the *KatG* ser315thr mutant, which is the most commonly observed [9] and were able to even proceed one step further using Triton [6], Modeller [7], Autodock [10] and Mopac [11].

The results of the Workshop were summarized in the Closing Ceremony, presided by the University Vice Chancellor, Prof. Aize Obayan, and attended by other members of the University's Management and all participants, who had an opportunity to express their feelings about the outcome of this hard working workshop week. The participants at the end of the ceremony received from the Vicechancellor their Certificates of Participation. According to the views expressed at this Closing Ceremony, the workshop was favourably evaluated by participants who asserted its value to bring together both (biologists and bioinformaticians) scientific communities and stir up collaboration, and expressed their interest in future continuation of this workshop series, an outcome that is encouraging and shows the effectiveness of this kind of initiatives.

We look forward to future re-occurrences of this workshop to build up momentum for research in Bioinformatics and the Life Sciences in the region, and hope to see the results of the co-operation ties in the form of work collaborations between the participants.

The workshop was wrapped up by a final day devoted solely to personal interactions, and during which participants were able to enjoy the leisure of a visit to Badagry and the beach at Ogungbe.

Acknowledgements:

Dr. Daniel Rotimi, Registrar, Prof. James Katende, Dean, College of Science and Technology, and in-fact the entire Board of Regents are to be thanked for their support in making this event possible. The Workshop Steering Committee (Dr. Ezekiel Adebiyi, Ms. Ijeoma Dike, Mr. Conrad Omonimiyin, Dr. Wande Daramola, Mr. Oluwagbemi, Mrs. Ogunlana and Dr. Segun Fatumo) are to be thanked for the efforts invested in making this a successful and smoothly run event, the Workshop Resource Persons (Drs. Valverde, Rebai, Bewaji, Adebiyi, Osamor and Fatumo) are thanked for their labor in steering the workshop, and all participants are to be acknowledged and congratulated for the big efforts and great undertakings achieved during it.

References:

1. Larkin MA, et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics, 23, 2947-2948.

2. Galtier et al. (1996) SEAVIEW and PHYLO _ WIN: two graphic tools for sequence alignment and molecular phylogeny, Bioinformatics, 12-6, 543-548

3. Valverde, J. R. (2007) IBS-ES-07 course: The making of. embnet.news, vol 13, no. 4. 11-17

4. Rice et al. (2000) EMBOSS: The European Molecular Biology Open Software Suite, Trends in Genetics, 16(6), 276-277.

5. Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle

6. Prokop et al. (2008) TRITON: a graphical tool for ligand-binding protein engineering. Bioinformatics. 24(17): 1955–1956.

7. Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234(3):779-815.

8. Wahab et al. (2009) On Elucidating Isoniazid Resistance Using Molecular Modeling, J. Chem. Inf. Model. 49, 97-107

9. Johnson et al. (2006) Drug Resistance in Mycobacterium tuberculosis. Curr. Issues Mol. Biol. 7: 91-112

10. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. and Olson, A. J. (2009) J. Comput. Chem. in press. "Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity."

11. MOPAC2009, James J. P. Stewart, Stewart Computational Chemistry, Colorado Springs, CO, USA, http://OpenMOPAC.net (2008)

Personal Account: Training in Grid Computing for Bioinformatics

Kanchana Senanayake

Institute of Biochemistry, Molecular Biology and Biotechnology (IBMBB), University of Colombo, Cumaratunga Munidasa Mawatha, Colombo, Sri Lanka

Since the Institute of Biochemistry, Molecular Biology and Biotechnology (IBMBB) was elected as the Sri Lankan EMBNet node at the EMBNet AGM in 2006 we have been in close contact with the Swedish EMBNet node. As a result of this association I was fortunate enough to get an invitation for a short visit to the Swedish University of Agricultural Sciences (SLU), Uppsala, Sweden. This visit was funded by SIDA/SAREC and IBMBB. In this article I will be giving a brief insight in to my visit from an academic and a personal point of view.

Academic

The main objective of my visit was to get hands on training in grid computing and its applications in Bioinformatics. A 14 node cluster was given to the IBMBB under the NorduGrid project in 2006. **NorduGrid** is a Grid Research and Development collaboration aiming at development, maintenance and support of the free Grid middleware, known as the Advance Resource Connector (ARC)[1]. The aim of this collaboration was to use grid computing for bioinformatics applications as well as porting existing bioinformatics tools on to the grid. These tools then can be used for high volume data processing application in various fields of Biology. Unfortunately the know-how was not there to make the best use of this resource at the IBMBB. The best way of describing this situation the Sri Lankan way is:

"Like giving an elephant, without the goad to control it"

So for the last three years this was not utilized at all for any type of research. In order to utilize this resource IBMBB needed someone with hands on

---