

FINAL REPORT
Workshop on Bioinformatics and Computational Biology (WEBCB)
7th International Conference on Bioinformatics (InCoB'08)
National Yang Ming University, Taipei, Taiwan.
23 October 2008
Organized by APBioNet
Sponsored by FAOBMB and IUBMB

1. INTRODUCTION

Over the past two to three decades, the world has witnessed how Information Technology (IT) has expanded the practice of biological research from the observational and experimental, to the informational and the computational. Traditional boundaries of computing in biochemistry from computational analysis in enzyme kinetics, pharmacokinetics, nerve conduction and structural biology have greatly expanded to include sequence analysis, genome informatics, microarray informatics, machine learning, natural language processing, database and text mining, image processing, cell and systems simulation, graph theory and networks in signalling and metabolic pathway analyses and so on.

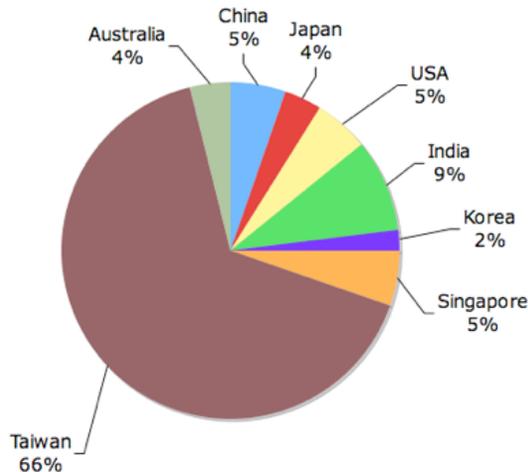
Not surprisingly, the inability of the traditional life science curriculum to cope with this shift in centre of gravity towards the informational and the computational has led to a burgeoning of post-graduate courses and training workshops in bioinformatics and computational biology (Ranganathan, 2005). While the already trained need re-training in such educational activities, the yet-to-be trained surely deserve some planning, at least, to modify and re-mould the traditional biochemistry and molecular biology curriculum in the life sciences, to include skills in bioinformatics and computational biology, which are increasingly expected of biochemists and biologists in the research laboratory and in industry. Beyond the skills level, and at a more fundamental level, the way in which biochemists and molecular biologists think about living systems, has to change.

From a reductionist approach, the so-called “antedisciplinary” revolution (Eddy, 2005) commands the future of biology to adopt approaches that are able to routinely focus on multiple molecular events from the integrative context of a systems level perspective. From the ability to monitor and measure a multiplicity of parameters to the facility of capturing, storing, curating, analyzing and maintaining these data, we need to review our curriculum and the way we educate a new generation of biologists and biochemists.

2. BIRTH OF WEBCB

To discuss issues involved in the education in bioinformatics and computational biology, the inaugural workshop on education in bioinformatics and computational biology (WEBCB) [<http://trg.apbionet.org/webcb>] was conducted on 23rd October 2008 at the National Yang Ming University, Taipei, Taiwan, following the 7th International Conference on Bioinformatics (InCoB) 2008. Jointly organized by (APBioNet), International Union of Biochemistry and Molecular Biology (IUBMB) and The Federation of Asian and Oceanian Biochemists and Molecular Biologists (FAOBMB), the workshop is tailored to educationists in bioinformatics, biochemistry and biology; scientists with keen interest in education; students in life sciences; policy makers interested in the impact of Bioinformation, Bioinformatics and Computational Biology and in promoting the development in their countries. The demography of the participants of the workshop is shown in *Figure 1*.

Keynote speakers from InCoB conference and FAOBMB symposium were invited to the workshop as panelists including Professor Vladimir Brusic, director of bioinformatics for the Cancer Vaccine Center in Dana-Farber Cancer Institute of Harvard Medical School, USA; Dr. Jong Bhak, director of the Korean Bioinformation Center (KOBIC), Korea; Professor Shoba Ranganathan, Chair Professor of Bioinformatics for the Department of Chemistry and Biomolecular Sciences, Macquarie University, Australia; Dr. Tan Tin Wee, Deputy Head of the Department of Biochemistry, National University of Singapore. On-site interviews conducted with Professor Terry Gassterland, director of Scripps Genome Center & Scripps Institution of Oceanography, University of California, San Diego, USA and also Professor Satoru Miyano from the Human Genome Center, University of Tokyo, Japan were also recorded in video format and presented during the workshop.



Total number of participants : 56

Figure 1a: The country of residence of the participants of this workshop.

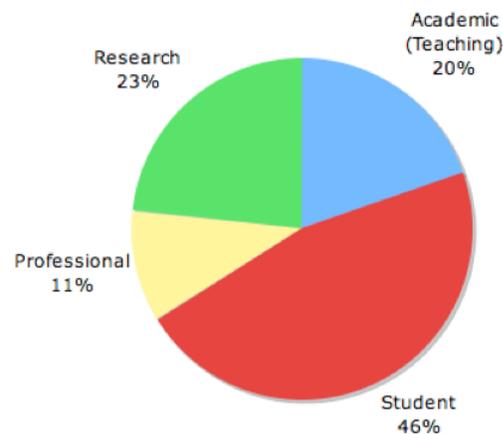


Figure 1b: Professions of the workshop participants.

3. TOPICS DISCUSSED IN WORKSHOP

Prior to this one-day workshop, the panelists and the workshop participants were provided with a set of questions. The questionnaire is available at ANNEX I. In general, the participants were asked to:

- name the top 10 challenges in Biology that are open to impact from progress in bioinformatics and computational and quantitative biology
- identify ways in which these techniques can be applied to solve these challenges
- discuss with the audience during the workshop on how these techniques and skills could be taught to a new generation of biologists who will eventually populate the laboratories of biologists solving these problems
- suggest a list of minimal skill sets for the new biologist to be proficient in bioinformatics and computational biology which will be agreed upon with the attendees during the workshop
- propose approaches to pedagogical tools and techniques for teaching or imparting these new skills to undergraduates or graduate students

On the day of the workshop, every attendee was given the same set of questionnaire which was used to guide the discussion. The workshop was conducted according to the programme schedule (see Table 1).

Table 1: Workshop programme

Thursday, 23 October 2008	
1:15pm – 1:30pm	- Registration - Welcome speech by APBioNet President Shoba Ranganathan - Introduction to Panelists
1:30pm – 3:30pm	Panel Discussion Session I
3:30am – 4:00pm	Break
4:00pm – 5:30pm	- Open Audience Discussion Session II - Wrap up and Task Force action items
5:30pm	End of workshop

4. FINDINGS FROM THE WORKSHOP

One of the aims of the workshop was to solicit the viewpoints from a panel of audience with sufficient diversity in their background in terms of education, culture, work specialization and other parameters, to provide a much balanced perspective. The workshop thus provided a suitable platform to achieve this objective. Several issues were discussed and the findings are presented in the following sections.

4.1. Impact of Bioinformatics and Computational Biology on Modern Biology Research

Biology has been a phenomenological science based on observation. Prof. Vladimir Brusic of Dana-Farber Cancer Institute remarked that biological affairs were carried out in comparatively simpler setting decades ago. However, with sequence and structure data generated and made available in large quantity and furious speed, this itself has become a challenge and has profound impact on how modern biology is carried out. Dr. Jong Bhak, director of KOBIC agreed and quipped that Biology is fast becoming an informational science and opined that bioinformatics and computational biology can help to improve Biology. Dr Chanchal K. Mitra from India, further added that the challenge now is to exploit the different granularity of the data which are now fragmented like a jigsaw puzzle waiting to be integrated to create new knowledge.

Professor Terry Gaasterland, Director of Scripps Genome Center & Scripps Institution of Oceanography, University of California, San Diego commented that computational biology, together with myriad high-throughput technologies for measuring presence and/or quantities (relative or absolute), have revolutionized the way many biologists think about their systems of interest, gene of interest, and how to pose questions. Computational tools for data post-processing and hypothesis generation have fundamentally altered what is possible in molecular, cellular, developmental, and genetic biology. Bioinformatics allows scientists to investigate by measuring many items simultaneously and in a system wide basis, thus allowing us to pose many questions. Nonetheless, the questions that can be asked by traditional techniques and by high-throughput measurement platforms are very different but complementary. It is therefore important for computational biologists to understand the merits, strengths and limitations of both approaches. Dr Tan Tin Wee explained that hypotheses formulated from the many different data sources will eventually begin and end with computational aspects where quantitative measurements have to be analyzed computationally. This will surely widen the gap between the different fields and subjects. Cross-disciplinary knowledge in biomedicine, engineering, computer science and many other areas are needed and these fields will possibly need to converge. Professor Shoba Ranganathan from Macquarie University, Australia who is also the President of APBioNet added that in today's context, it is inevitable and getting increasingly impossible for

biologists to get through without computational effort/tools. Furthermore, Biology can definitely gain advantage with the help of technology to sharpen and deepen the research.

Modern biological science is largely teamwork and multidisciplinary and interdisciplinary. This has allowed modern biology to grow at an unprecedented pace with scientists and researchers attacking biological problems from the different background and bringing with them their own fresh perspectives. New developments in technology, techniques and methodology and scientific instrumentation have resulted in even faster growth in data acquisition, and a heightened need for more data analysis, data organizational tools. This is where bioinformatics and computational biology contribute and make an impact.

4.2. Top Challenges in Biology

With more data generated over the years and the advent of new technologies, we are now able to take on new challenges which were probably not feasible before, in addition to the many existing problems. To be sure, we face new challenges as well with the introduction of these new technologies and new data. The panelists were asked to name 10 challenges that they think are open to the positive impact bring about by bioinformatics and computational biology.

Some frequently raised challenges include

- How to extend bioinformatics approaches to multi-scale modeling, all the way from the molecule to cell to tissue to organ to organ system to organism and finally to the population?
- Unraveling the current alchemy of systems biology
- Current definition of "systems biology" addresses only molecular level. New approaches are necessary for translational applications
- Organization of Scientific knowledge in a disciplined way that can support future growth. This may include sustainable infrastructure to store the data, information and knowledge in perpetuity. It may include databases to ontologies, controlled vocabularies to persistent registries, identifiers and handles.
- Genomic science driven by third and fourth generation sequencing machines, from genomics to meta-genomics
- Structural biology and structural bioinformatics
- Solving the epigenetic code and perigenetics
- Elucidating the complexity of the cell and inter-cellular communication and information transfer
- Unraveling immunological regulatory circuits
- Solving the pathways of organismal development, cell differentiation, stem cell totipotency and pluripotency
- Integrating computational concepts into the biology curriculum of colleges and universities
- Deciphering metabolic pathways
- Understanding how cells maintain a steady state

4.3. Participants' Views on Current Curriculum

In response to the new challenges, participants were sought for their views on the adequacy of the current undergraduate biology curriculum in their respective institutions in equipping the students with core bioinformatics and computational biology skills, in particular, the skills to solve biological problems informatically and computationally. There was a consensus that the current system is in need of review and changes. Often, many of the students are found to be insufficiently quantitatively, informatically and computationally literate upon their enrolment to university. Thus, syllabus and curriculum that provides ample training should be implemented at

the pre-university level. However, on-going training and practice should be encouraged at the undergraduate and postgraduate levels as well. Essentially, this would also require a whole new generation of better equipped biology teachers who are themselves literate in these new techniques.

Dr. Chanchal K. Mitra lamented that Biology is often considered as the last choice of study by many students in India. Typically, many of the students who do eventually take up Biology as their studies are often lacking in basic mathematic skills, which is often required in resolving many of the problems in science. Mathematics is commonly regarded by the students as an elite subject with no application or future. Fortunately, there is a gradual shift observable and more students are beginning to appreciate the importance of Biology and Mathematics.

In addition, other participants conveyed that the resistance or reluctance by the faculty members to change or to understand the need for a revamp curriculum to cope with the new demands can be a major deterrent to the progress in reformation.

As an early proponent of bioinformatics and computational biology, Dr. Tan Tin Wee, explained that bioinformatics and computational biology have become an integral aspect of any modern biological research programme. It is now essential for every lab to have people who are skilled in bioinformatics and computational biology. He proposed that every biologist should equip themselves with such basic skills. Thus, a review of the current curriculum is timely and urgent as the gap in basic skills between top schools and average schools will widen.

It is also noted from the participants that there is often a lack of continual effort to review the curriculum even as technology continues to advance. It remains that in many schools, there is often too much emphasis on specialisation. In short term, it would seem difficult to equip students and faculty members with multi-disciplinary capabilities or facilitate such initiatives. Thus, it would probably be more feasible to incrementally introduce combined subjects to the students.

4.4. Minimal Skill Sets Required for the Modern Biological Research

The participants have proposed the following as the minimal skill sets required for modern biological research.

- Knowledge of available bioinformatics data resources
- Proficiency in database tools
- Knowledge of available bioinformatics tools
- Proficiency in bioinformatics software and tools
- Programming skills for bioinformatics software and tool development
- Mathematical, statistical, structure or quantitative modelling techniques
- Ability to handle to handle complexities (complex systems, hierarchical levels (multiscale modeling), self-organization)
- Good knowledge and understanding of molecular biology, cell physiology, and/or the molecular basis of the biological domain

The participants also agreed that a minimum duration should be allocated for the students to equip themselves with these minimum skill sets so that they have ample time to practice.

4.5. Proposed Topics for Inclusion Into Curriculum

In order to enable students and faculty members to acquire the minimal skill sets proposed above, the participants have suggest a list of topics to be included into the bioinformatics and computational biology curriculum in the future.

- Algorithms and programming
- Mathematics
- Statistics and numerical analysis
- Machine learning
- Structural biology and molecular modeling
- Database tools should be taught prior to entry at the research level; i.e., masters level
- Bioinformatics software and their applications (optional)

4.6. Pedagogical Tools for Learning and Teaching

As bioinformatics teaching involves complex computational concepts, it is necessary that new pedagogical tools are explored and adopted to promote effective learning and teaching. The participants proposed that “chalk-and-talk” lectures should still be retained as they provide a better connection not just with the audience but also the lecturers.

However, new pedagogical methods such as problem-based learning (Schmidt, 1993) and emerging e-technology such as Learning Activity Management System (LAMS) [<http://www.lamsinternational.com/>] that can encourage collaborative effort and teamwork should be implemented. This is particularly crucial as modern biological research increasingly require good coordination and participation from researchers hailing from different backgrounds, and possibly located at different geographical locations.

In this respect, Webinars [<http://assignments.uspto.gov/assignments/q?db=tm&sno=75478683>] should also be introduced to facilitate immersive learning and to promote active discussion and exchange of ideas in real time. Hand-on training is also important and the amount of practical hands-on projects/tasks should be increased to encourage the interdependence of the various subject areas. This ensures that the students will get to know how to use bioinformatics and computational biology tools and gain confidence.

5. CONCLUSION

Based on the feedback and survey, there is a clear consensus that bioinformatics, computational biology and quantitative biology are key drivers of modern biological research, and will continue to have major impact in almost all the diverse sub-fields of biology. All top challenges highlighted by interviewees have significant or crucial involvement of bioinformatics and computational techniques. There does not appear, however, to be any consensus regarding the curriculum that needs to be emplaced in order to train the future generations of biologists capable of dealing with such quantitative aspects of biology.

Undoubtedly, the revolutionary changes that biology has undergone in these past three decades has created many lines of development, the dust of which has not settled, and as such, different leading biologists coming from different backgrounds and persuasions will have their own perspectives. Notwithstanding the tendency towards convergence, the state of consensus has not yet arrived for any coherent and universally agreed minimal skill sets which a biologists should possess in the quantitative, computational and informatic realm, not least the universally agreed curriculum that will satisfy all quarters.

Without a doubt, many aspects of statistics, computing, computational science, informatics, information technology, mathematics, have been suggested for inclusion in the curriculum, far more than an average biology curriculum can handle. One solution could be a diversification and differentiation of biology into specialized subfields with the differing blends of such topics in addition to mainstream topics in “wet-lab” biology.

There is, nevertheless, significant consensus that new and productivity-generating pedagogical tools and methodology should be utilized in the delivery of such curricula, including newer techniques such as problem-based learning, e-learning tools, learning activity sequencing, and moodle-like constructionivist learning approaches, beyond the traditional chalk-and-talk lecture and passive-mode tutorials.

Finally, it is also clear that a new generation of biology teachers who are well versed in computational biology and bioinformatics, especially those who have been formally trained, and not just self-taught, should be nurtured, who will in turn be the flag-bearers of this revolution in biology education in the making.

6. ACKNOWLEDGEMENTS

The organizers of WEBCB would like to acknowledge the help of the InCoB 2008 organizers in the local arrangements for the workshop and providing the excellent support for the publicity, registration and logistics for the workshop. Most of all, we acknowledge the funding and support of FAOBMB and IUBMB without which this workshop will never have materialized.

REFERENCES

Eddy, S.R. 2005. "Antedisciplinary" Science. *PLoS Comput Biol*, 1: e6.

Ranganathan, S. 2005. Bioinformatics education—perspectives and challenges. *PLoS Comput Biol*, 1(6): e52.

Schmidt, H. G. 1993. Foundations of problem-based learning: some explanatory notes. *Medical Education*, 27:422-432.

ANNEX 1: WEBCB Workshop Questionnaires

Online form and printed versions of the following questionnaires were sent and distributed to the participants in advance of and during the workshop to guide the discussion. Results from the questionnaires are presented and discussed in this report. Other responses from selected panelists were also recorded in video format and posted on to the workshop website (<http://trg.apbionet.org/webcb>).

Questions #1:

In your opinion, how has bioinformatics and computational biology impacted the practice of modern biological research?

Questions #2:

Is the impact positive or negative and why?

Questions #3:

What are the Top 10 challenges in biology which are open to positive impact from bioinformatics and computational biology?

Questions #4:

How do you foresee bioinformatics and computational biology being used to tackle these challenges?

Questions #5:

What do you think are the minimal skill sets required for the new biologist to be proficient in the use of bioinformatics and computational biology techniques ?

Example:

- * knowledge of available bioinformatics data resources
- * proficiency in database tools
- * knowledge of available bioinformatics tools
- * proficiency in bioinformatics software and tools
- * programming skills for bioinformatics software and tool development
- * knowledge in providing online accessible resources
- * mathematical, statistical, structural or quantitative modelling techniques

Questions #6:

Do you think the current undergraduate biology curriculum in your institution or your country or elsewhere from whence you get your graduates students are adequate in equipping them with core bioinformatics and computational biology skills and the skills to solve biological problems informatically and computationally? Why and how so?

Questions #7:

Can you suggest bioinformatics and computational topics that should be included into the curriculum and to what level should they be taught?

Please cover at least the following aspects:

- * algorithms and programming
- * database tools

- * bioinformatics software
- * mathematics
- * statistics
- * machine learning
- * structural biology and molecular modelling

Questions #8:

Can you propose some pedagogical tools and techniques that can be used for improving the teaching Bioinformatics and Computational Biology to Biologists?

- * Please comment on some of the following: chalk and talk lectures
- * demonstrations
- * small group tutorials
- * problem based learning
- * LAMS sequences
- * case studies
- * practical hands-on projects
- * tutorial questions
- * small group discussion
- * paired programming
- * case studies
- * mini-project for integrative learning
- * essay writing for synthesis