Educating the educators: Incorporating bioinformatics into biological science education in Malaysia.

Mohd Shahir Shamsir*, Huszalina Hussein*, Siti Zaiton Mohd Hashim++ and Naomie Salim++.
*Bioinformatics Laboratory, Biology Dept, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor Malaysia
++ Faculty of Computer Science and Information Systems, Universiti Teknologi Malaysia, 81310 Skudai, Johor Malaysia

Abstract

Bioinformatics can be defined as a fusion of computational and biological sciences. The urgency to process and analyse the deluge of data created by proteomics and genomics studies has caused bioinformatics to gain prominence and importance. However, its multidisciplinary nature has created a unique demand for specialist trained in both biology and computing. In this review, we described the components that constitute the bioinformatics field and distinctive education criteria that are required to produce individuals with bioinformatics training. This paper will also provide an introduction and overview of bioinformatics in Malaysia. The existing bioinformatics scenario in Malaysia was surveyed to gauge its advancement and to plan for future bioinformatics education strategies. For comparison, we surveyed methods and strategies used in education by other countries so that lessons can be learnt to further improve the implementation of bioinformatics in Malaysia. It is believed that accurate and sufficient steerage from the academia and industry will enable Malaysia to produce quality bioinformaticians in the future.

Key words: Bioinformatics; computational biology; education

1. Introduction

The purpose of this paper is to survey methods used to incorporate Bioinformatics in biological science education and discuss possible strategies that could be utilised to achieve this in Malaysia. Bioinformatics represents a new field at the interface between computer science and molecular biology. The definition of bioinformatics is in itself an issue of contention. It has been myriadly defined, from a constricted view as a management of biological data derived from genomic research to a liberal view as everything computationally applied in molecular biology. This is because of the multidisciplinary nature of bioinformatics, encompassing biology, mathematics and computer science. According to a National Institute of Health (NIH) definition, bioinformatics is "research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioural or health data, including those to acquire, store, organise, analyse or visualise such data". For the purpose of this paper, this definition has been adopted. The objective of this paper is to present available strategies to educators who plan to teach bioinformatics, either by conducting a specialist undergraduate or postgraduate Bioinformatics course or incorporating it the existing curricula. In addition, the current scenario, challenges, requirements and future trends needed to ensure the successful teaching of bioinformatics will be discussed.

1.1 Disciplines within bioinformatics

Which area is required when teaching bioinformatics? Areas within biology include molecular biology, genomics and proteomics. The focus revolves around the Central Dogma, the encoding within the genome, its expression into protein and subsequently the protein function. Students are expected to know theories and principles within genetics, gene sequences, gene expression, protein structure, function and interactions. Bioinformatics also involves building tools to process, manage and display biological information in a meaningful manner. This requires the ability to write efficient computer programs. Therefore, a computer science curriculum is probably required in its entirety with emphasis on solving biological problems. Other areas that must be included are statistics and calculus

2. Delivering Bioinformatics education

Demand generated by expansion and development of bioinformatics in industry has spurred the creation of bioinformatics courses in many countries . Many surveys of bioinformatics education and research initiatives have been published representing countries such as the UK , US , Australia , Israel , France and Germany . Comparative studies on the curriculum in university and college levels has shown that emphasis is given to the computer science element and the importance is given to interdisciplinary teaching, either due to the complicated subjects offered or lack of trained educators . In addition, many critical points must be considered when incorporating bioinformatics or creating a bioinformatics course. The Workshop on Education in Bioinformatics (WEB), which started as an annual satellite meeting of International Conference on Intelligent Systems for Molecular Biology (ISMB) has raised several interesting points such as the shape and design of bioinformatics courses, the components within such course and the integration of bioinformatics elements into conventional biological science subjects . Overall, the emphasis of bioinformatics training can be divided into 3 levels; teaching the use of pre-existing tools, teaching basic programming with algorithm design and teaching in-depth theoretical foundations and principles behind bioinformatics . Teaching usage of tools include sequence analysis, protein structure visualisation and modelling while teaching basic programming covers writing of simple scripts using Perl and Java to sift through data. In-depth teaching includes teaching about genetic algorithm, neural networks and Hidden Markov models.

2.1 Teaching Bioinformatics as an undergraduate course

Concerns for the need for an undergraduate curriculum for bioinformatics was raised by Altman in 1998 . The editorial also included an overview and guideline of the topics that should be taken into consideration when creating a bioinformatics curriculum, although the materials were more suited to the duration and style of a US Masters course. Subsequently, a significant number of papers have discussed and described a variety of bioinformatics education, curriculum contents and its method of delivery .

The biggest challenge in teaching bioinformatics is trying to fulfil the breadth of knowledge created by the fusion of multidisciplinary areas within bioinformatics. These different areas relate to biological molecules and therefore require knowledge in the fundamentals of biochemistry, molecular cell biology, genetics, thermodynamics, biophysics, and statistical mechanics. Students are then required to apply analytical capabilities obtained from their knowledge of computer science, mathematics, and statistical principles to sift the deluge of data produced by genomic and proteomic studies. The multidisciplinary nature not only transcends the established areas of sciences, it also forces integration of knowledge and cross-field utilisation of techniques where researchers use it interchangeably . Consequently, teaching bioinformatics will require a specialist educator with in-depth knowledge of all the different components: mathematics, biology and computer science. This is quite a daunting task, therefore most universities lacking specialist and experience bioinformatics staffs resort to interdisciplinary and cross faculty teaching. This is logical route as utilising experts in their respective areas will not only ensure a well-taught course but also teach the necessary breadth and depth. Inter-faculty teaching will then raises the issue of ‘ownership’ and the placement of the course in an institution. The management has to determine which faculty or department to house the course, facilities to use, in biology, mathematics or computer science departments. Consequently, this will cause concerns on the teaching and learning culture, as it is obvious each discipline has its own inherent culture.

Another challenge is the depth at which to teach bioinformatics. Pevzner and Pearson raised the issue of depth commenting that broad introduction to bioinformatics without the necessary depth will produce bioinformatics technicians rather than bioinformatics scientists. The authors also highlighted the importance of teaching principles of algorithm, statistics and creating a biologically motivated problems-based learning in order to effectively teach bioinformatics . Most biologists are comfortable in using softwares such as BLAST and are contented simply by either finding a match or not, without even understanding the underlying principles behind it. This treatment of bioinformatics merely as computational tools is prone to erroneous assumptions if derived from flawed understanding of the algorithm behind the tools. This is further compounded by the rise in publication of bioinformatics textbook that is cook-book styled and protocol-centric . Hence, failure to produce a course with the necessary depth will produce students severely lacking in skills for pursuing careers in bioinformatics. The necessary depth of the curricula is achievable and commensurate to an undergraduate course time span (3 or 4 years). This can be achieved by paring down the non-related biological subjects and focusing on cellular molecular biology, genetics and gene expression studies.

Teaching a distinct bioinformatics course is highly desirable as it gives the opportunity to inculcate a proper understanding of programming techniques and the ‘discipline’ of programming. This ‘culture’ is critical as undergraduates are expected to design and create bioinformatics softwares . Students are also taught biological facts in context, e.g. introns and its effect on automated gene annotation, rather than isolated facts taught separately. This ability along with teaching of an over-arching theoretical framework would generate a coherent understanding in the students. A well-designed course is intrinsically integrated and has been adopted by majority of universities. Curriculum structures of foreign universities have been surveyed with 78 in the US, 10 in Australia and 25 in the UK. A quantitative analysis was not performed, as there were difficulties in acquiring the curriculum details and overlapping of content between differently named modules. However, a list of subject areas repeatedly found in bioinformatics curriculum is listed in Table 1 for guidance.

Table 1. List of subject areas/modules taught in curriculum surveyed

Biology • Molecular Cell Biology • The concept of Central dogma • Organ and systems level Biology • Cellular Biochemistry • Microbiology • Genetic Engineering • Genomics and Genetics • Protein structure and function	**Others** • Microarray data analysis • Bioethics • Current Issues and Future challenges • Molecular Modeling • Proteomics
Computer science • Dynamic programming • Bounded search algorithms • Cluster analysis • Classification • Neural Networks • Genetic Algorithms • Bayesian Inference • Database Structures, concept, design and mining methods • High performance computing • Large Scale Programming • Client-server Architecture and HTTP • CGI, Perl, Java and PHP • Automating tasks with shell scripting • Optimization (Expectation Maximization, Monte Carlo, Simulated Annealing, gradient-based methods)	**Bioinformatics** • Introduction to Bioinformatics • Profiles and motifs • Pair-wise sequence alignment and Multiple sequence alignment • Hidden Markov Models (construction, use in alignment, prediction) • Predicting protein structure and function from sequences • Fragment and map assembly and combinatorial approaches to sequencing • Tools for Visualizing sequences • Tools for Protein visualization • Structural visualization • Phylogenetic Trees • RNA Secondary structure prediction • Sequence feature extraction/annotation • Protein homology modeling • Protein threading • Protein molecular dynamics • Protein *ab-initio* structure prediction
Mathematics • Calculus • Statistics	

2.2 Teaching a postgraduate course

Conducting graduate courses is relatively easier as the students would already have learned science and mathematics fundamentals in their undergraduate studies. However, the difficulty arises from this very diverse background. Students enrolling into bioinformatics will be handicapped as the training was primarily on a single discipline, either biology or computer science. Graduates courses in Bioinformatics would have to conduct intensive remedial classes to make up for this lack of knowledge. Computer science students will have to take classes in biochemistry, molecular biology, and genetics whereas biology based students will generally spend the majority of their part of graduate study in course work covering introductory computer programming, databases and artificial intelligence. The faculty must be able to develop a suitable bioinformatics curriculum by plugging the skills gaps in these students from various backgrounds and making sure it fits the needs of the industry. In addition, there is a time constraint of 2-3 years associated with MSc studies compelling teaching staff to be creative in optimising this relatively short span of time. However, time limitation is compensated by the smaller nature of graduate courses that enables coaching at a more personal level. Many teaching methods employing small groups participating in active learning are found in the literature. Graduate classes at the University of Michigan uses cooperative learning to interpret a sample data set derived from various bioinformatics tools such as microarray data, two-hybrid data, homology-search results to elucidate a hypothetical signalling pathway. Reconstruction of the full pathway by integrating these data sets provided a successful and effective introduction to the field of systems biology and reinforces the bioinformatics elements of the exercise .

2.3 Integration with existing subjects

One approach to teach Bioinformatics is to integrate elements of Bioinformatics into existing ‘traditional’ subjects such as Genetics, Molecular Biology, Zoology and Biochemistry . The advantage of this method is that it covers the spectrum of the curricula and is not intrusive to a tight and packed structure. It also creates a simultaneous teaching throughout the course, allowing a rapid exposure to bioinformatics within the domain of each field and laying a foundation for higher levels of education. It is also the most realistic strategy as it does not require individuals with specialist bioinformatics training and distribute the burden between faculty staff. For example, bioinformatics exercises can be incorporated within biological sciences subject such as Biochemistry , Cell Biology , laboratory sessions or incorporating in computer science courses . Element to be introduce can be open ended inquiry exercise using online tools such as Protein Explorer or downloadable software such as PyMOL . Many of these softwares are free and Internet accessible. This integration method have been adopted by many foreign universities to complement their existing curricula . A continuous exercise can also be integrated with the student’s final year project by imbedding a bioinformatics analysis element in their projects. However, integrating bioinformatics within a subject is sometimes limited to teaching the usage of bioinformatics tools, as time is constraint.

2.4 Creating a specialized bioinformatics subject or module.

Another approach is to create a specialized subject or module that focuses on Bioinformatics within the curricula. A bioinformatics subject can be packaged to have both training on how to use bioinformatics tools and teaching the underlying principles behind the algorithm and programmes used in such tools. Creating unique course can allow universities to offer traditional courses with a minor in Bioinformatics. Many universities have adopted this method in the US and Israel .

2.5 Plugging the gaps in education and industry

Concerns were raised about the pressing short-term need for bioinformatics now. To address this need, retraining packages for experience and skilled researchers was proposed. Short intensive training courses either part time or distance learning and workshops tailor-made for people from the industry can be created to allow a quick transition into bioinformatics, either from biology, mathematics and computer science areas. This would also allow for continued professional development for postdoctoral scientists to update their computational capabilities. At very the least, to non-computer scientists, learning to write short scripts would automate tedious monotonous tasks.

3. Requirements for Successful teaching

Many critical points that need to be considered when incorporating bioinformatics or creating a bioinformatics course. The Workshop on Education in Bioinformatics (WEB), which started as an annual satellite meeting of International Conference on Intelligent Systems for Molecular Biology (ISMB), is created as a venue for bioinformatics educators to meet and discuss issues in bioinformatics education. WEB meeting has raised several interesting aspects that must be determined when creating a bioinformatics course; the shape and design of bioinformatics courses, the components within such course and the integration of bioinformatics elements into conventional biological science subjects .

Many of the literature surveyed have highlighted several requirements to ensure an effective bioinformatics teaching: fast Internet access, using a practical heavy curricula and departure from traditional passive learning by using creative instructional delivery. Internet access is important in teaching bioinformatics. The changing of trends in information access, particularly over the Internet have been shown to transform biological science education . Students now need to access online resources, usually a free central depository for biology-derived data. There are 166 online databases and resources available freely over the Internet . Access is crucial as a fast Internet access would put researchers in a developing country biologist at par to an academic biologist in an industrialized country .

The hands on nature of bioinformatics require students to repeatedly perform data or sequence analysis thus requiring a practical heavy curriculum. This is to simulate the repetitiveness of research and to familiarise the students to tools used routinely for data analysis by researchers in the field. However, utilisation of bioinformatics tools is only useful as an exposure early in the course. To create a bioinformaticist, the curricula must advance further by creating an understanding of the principles behind the analysis tools. The complexity of bioinformatics itself lends a challenge in determining the suitable instructional methods for which to apply in education . Survey of the literature have revealed numerous proposed methods and examples; e.g. Instructional Design Theory , Problem Based Learning , Inquiry-based laboratory . Initiatives to teach Bioinformatics by distance learning have also been started. Thailand for example has embarked on that model with the aim to provide bioinformatics education across geographical barriers . Singapore has also founded a distance learning course supported by its S-Star Life Science Informatics Alliance that comprises of six institutions from five continents; Karolinska Institute in Sweden, the National University of Singapore, Stanford University in the United States, Uppsala University in Sweden, the University of Sydney Australia, and the University of the Western Cape in South Africa .

4. Current scenario in Malaysia

Currently bioinformatics education in Malaysia encompasses undergraduate and postgraduate programmes. Institutions that offers Undergraduate Bioinformatics courses are UM, KUTPM and UTM . However, detailed information is lacking on each university's corresponding website. Currently, several initiatives have been implemented to generate impetus for bioinformatics research. Biotechnology based Bio-Valley, which was proposed under the Eighth Malaysia Plan, was aimed to be the nucleus of the country's biotechnology industry. The cluster was intended to comprise biotechnology research institutions, universities and specialized companies and would foster the convergence of intellectual expertise and entrepreneurship. Establishing Bio-Valley would be an excellent platform for generating bioinformatics research. However, its failed fruition reinforces the complexity involved in creating successful technology science parks . The creation of bioinformatics networks has also been initiated with the aim to enhance research and accelerate productivity. National Biotechnology and Bioinformatics Network (NBBnet) is such an initiative . Conceptualised in 1997 and realized in 1999, it aims to give R&D support to the researchers for conducting Biotechnology and Bioinformatics researches. It aims to allow sharing of resources to using a dedicated bioinformatics architecture, consequently improving research in Malaysia . In addition, an initiative named EMASGRID, an acronym for Extensible Malaysia (NBBnet)-SGE Grid is an attempt to harness the power of a centralized compute grid system operated by Sun Microsystems's Sun Grid Engine (SGE) . The latest initiative is the inter-institution network called Malaysian Research and Education Network (MYREN) launched in 2005. Malaysia is also the founding member of The Asia-Pacific Bioinformatics Network (www.apbionet.org) along with Australia, Canada, China, Japan, Korea, Singapore, and the US. The aim of this network is to encourage cross border information exchange and collaborations in the bioinformatics field. Initiatives created by APBioNet include the Biomirrors project , creating a similar model for African Bioinformatics Network and launching International Conference on Bioinformatics (InCoB), an annual international bioinformatics conference with the inaugural conference was successfully held in Bangkok in 2002 . InCoB has also been held in Malaysia at Universiti Sains Malaysia, Penang in 2003. Similarly focused conferences have also been held by other institutions such as HELP University College's called Bioinformatics Symposium 2005 entitled 'IT comes Alive' and Universiti Teknologi Malaysia's Biology Inspired Computing (BIC) in 2005. Private initiatives in Malaysia are still in its infancy although few companies such as Synamatix have made their mark in bioinformatics research

5. Future trends

There is a concern that bioinformatics has lost its impetus and is moving towards a supporting role in biological sciences. Once the darling of venture capitalist, the drop in bioinformatics has been compared with the bursting of the dot.com bubble. The first computational biology course was introduced in 1989 by Carnegie Mellon University and the existence of biological databases, computational application in biology existed since the 1950's. Therefore, the application of computers in biology is not a recent phenomenon. Only the flood of data created by the race to sequence the human genome combined with the over expectation hyped by journal publications and redressing computational biology as a 'new' field of bioinformatics has pushed computational biology/bioinformatics to the centre stage.

The mainstay of bioinformatics application has been on the premise that the genomics sequence data held information to complex diseases and interpreting these sequences would lead to the curing of these diseases either by gene therapy or protein specific drugs. The premise is that each gene is coding for a specific protein that possesses a specific function. This simplistic view accelerated the sequencing and mining of the genome. However, the human genome is more complex than anticipated. Firstly, even though the human genome is around 30 times larger than the worm or fly genome, the number of genes is only two or three times as many as the fly or worm genomes. It seems that the differences in phenotypic complexities is not due to the number of genes but to the interactions between them. Therefore, there are not that many genes to find and annotate.

Furthermore, the larger size of the human genome makes locating the genes much more difficult as the introns are bigger, decreasing the signal (exons) to noise (introns) ratio and causing erroneous prediction by computational gene finding strategies .

Another attraction of bioinformatics is obviously the ability to perform industrial scale data mining and the prospect of discovering genes of commercial values. It would envisaged that this would accelerate production and reduce drug costs . The future was envisioned with the coming of a new “theoretical, virtual biology” where purely in-silico investigation can be pursued. However, this is a naïve view and any perceived discovery and advances must be validated in ‘traditional’ laboratory work, creating a backlog of ‘tens of thousand of genes awaiting this laborious scrutiny’ . This reality has now become apparent, making investors realise that the yield obtained by database mining activities after 2001 have been below expectation and prompting anxiety over the sustainability of investments into the field of bioinformatics . This entire decline have been signalled by the closure of bioinformatics companies as early as 2000 . There are some who even considers demoting bioinformatics into tools and application status . Currently, bioinformatics tools for biologist are becoming more user-friendly. This would cause bioinformatics tools to be routinely used in laboratories, eliminating a specialised need akin to the commercialisation of kits to standardised difficult laboratory procedures by biotechnology companies.

How does this decline affect quantity and the manner of the bioinformatics courses being offered? Universities offering bioinformatics courses have also declined in numbers due to decreasing demands from students. This phenomenon has been observed in Europe and the US . Studies conducted showed either reduction in the number of courses offered or reduction in the enrolment number of students. Corresponding analysis on the number of bioinformatics vacancies advertised have also showed a marked decline. This consolidation and maturation of the education market mirrors the direction taken by the industry a few years earlier. Surprisingly, the opposite is occurring in India where Bioinformatics courses are ‘sprouting’ and becoming a money spinning enterprise teaching mediocre curriculum .

Bioinformatics research has also shifted from genomic sequences to protein expression and determination, microarray, assimilating into areas like medical informatics, phamacogenomics, and creating new areas of interest such as transcriptomics, metabolomics and systems biology. The future of bioinformatics and even biological sciences is becoming a multidisciplinary field which integrates approaches from engineering, mathematics and computer science . The field of bioinformatics itself is very fluid; therefore, the curriculum must also be very adaptable.

The focus of bioinformatics is also changing from the study of single entities such as DNA sequences and move towards more complex areas such as metabolic networks, protein structures, functions and the interconnectivity between different proteins. This shift in research focus would require high end computing, a consequence of the increase in data size and complexity. Integration of all these autonomously generated disparate data will be another added layer of complexity signalling a greater emphasis on computer sciences element in the future. Complementing the increased computing power would also require creation of massive data storage facilities and the establishment of a high performance network, both potential areas for development in the future.

6. Conclusions – the next step for Malaysian educators

For Malaysia, the obvious strategy is building capacity to create home grown skills. Human resources have been proved the bottleneck in developing bioinformatics. Even developed countries in Asia like Japan was caught flatfooted by explosion in demand . Sufficient training will create a stable of bioinformaticians and reduce dependence on imported software. As bioinformatics is a new emerging field in Malaysia, participation from every discipline is critical to ensure a rapid introduction of bioinformatics in education . Multidisciplinary aspect of bioinformatics must not a bane but must be embraced and capitalised.

Survey of the literature and examination of implementations in other universities suggests that Bioinformatics education in Malaysia is best taught integrated with traditional subjects. This method has several advantages. First, it is easiest to implement without resorting to specialists bioinformaticians that is difficult to recruit in the first place. Existing staff most probably has exposure to bioinformatics and further training or even retraining to reorientate their careers to bioinformatics is faster than waiting for existing bioinformatics courses to graduate new bioinformaticians. Secondly, integration would accelerate exposure and awareness throughout the curriculum and keep the students at pace with up to date scientific development. Biologist would be trained with tools and basic programming while computer scientists and mathematicians can be exposed to biological problems that require their expertise. It would also prepare them for future bioinformatics graduate courses. Thirdly, integrating Bioinformatics element will enhance the value of their training rather than diminish the skills associated with their degree. Malaysia is still in the 'tools' use stage and lacks the necessary bioinformatics industry to absorb a large influx of new graduates from academia. The students would still have the laboratory training associated with traditional biological science degrees to ensure a greater chance of employment.

Should institutions decide to create a bioinformatics course, they should contend with the multidisciplinary nature of bioinformatics. Courses should include all the disciplines required with the necessary depth to ensure that the students graduate with the essential skills and not caught in the netherworld between biology and computer science. Emphasis must be given to programming and computing expertise as the core skills of careers in bioinformatics are about writing programmes while in biology, emphasize must be given to molecular biology, proteomics and systems biology. The easiest route would be to use an existing computer science course and adapt it by teaching the biological concepts within bioinformatics.

Bioinformatics is necessary for Malaysia to remain up-to-date in the biomedical, biotechnological and agricultural sectors. We might have missed the initial wave but as the field of bioinformatics is maturing, we can develop our own market by avoiding mistakes made by developed countries. The future potential may lay in the rich biodiversity in Malaysia . Together with bioinformatics, it can be capitalised by managing database of indigenous species and tapping it for a potential novel drug discovery.

7. References

1. Zauhar, R. J. University bioinformatics programs on the rise. *Nature Biotechnology* **19**, 285-286 (2001).
2. Brass, A. Bioinformatics Education--A UK perspective. *Bioinformatics* **16**, 77-78 (2000).
3. Counsell, D. A review of bioinformatics education in the UK. *Briefings in Bioinformatics* **4**, 7-21 (2003).
4. Hack, C. & Kendall, G. Bioinformatics: Current Practice and Future Challenges for Life Science Education. *Biochemistry and Molecular Biology Education* **33**, 82-85 (2005).
5. Hemminger, B. M. & Anne Bauers, T. L. Survey of bioinformatics programs in the United States. *Journal of the American Society for Information Science and Technology* **56**, 529-537 (2005).
6. Zatz, M. M. Bioinformatics training in the USA. *Briefings in Bioinformatics* **3**, 353-360 (2002).
7. Littlejohn, T. Bioinformatics in Australia. *Bioinformatics* **16**, 849-850 (2000).
8. Cattley, S. A review of bioinformatics degrees in Australia. *Briefings in Bioinformatics* **5**, 350-354 (2004).
9. Samish, I. Bioinformatics Education Programs in Israel: Academic, Private and Unique Professional-retraining Programs. *Workshop on Education in Bioinformatics (WEB)* (2003).
10. Danchin, A. A brief history of genome research and bioinformatics in France. *Bioinformatics* **16**, 65-75 (2000).
11. Schomburg, D. & Vingron, M. Bioinformatics research and education in Germany. *In Silico Biology* **2** (2002).
12. Burhans, D. T. & Skuse, G. R. The role of computer science in undergraduate bioinformatics education. *SIGCSE Bulletin* **36**, 417-421 (2004).
13. Morrow, C. & Wilkins, D. in *Proceedings of the 2nd annual conference on Mid-south college computing* (192-199, Little Rock, Arkansas, 2004).
14. Ranganathan, S. Bioinformatics Education—Perspectives and Challenges. *PloS Computational Biology* **1**, 52 (2005).
15. Dyer, B. D. & LeBlanc, M. D. Meeting Report: Incorporating Genomics Research into Undergraduate Curricula. *Cell Biology Education* **1**, 101-104 (2002).
16. Altman, R. A curriculum for bioinformatics: the time is ripe. *Bioinformatics* **14**, 549-550 (1998).
17. Cohen, J. Guidelines for Establishing Undergraduate Bioinformatics Course. *Journal of Science Education and Technology* **12**, 449-456 (2003).
18. Cooper, S. Integrating bioinformatics into undergraduate courses. *Biochemistry and Molecular Biology Education* **29**, 167-168 (2001).
19. LeBlanc, M. D. & Dyer, B. D. Bioinformatics and Computing Curricula 2001: Why Computer Science is Well Positioned in a Post-genomic World. *inroads – The SIGCSE Bulletin* **36**, 64-68 (2004).
20. Dubay, C., Brundege, J. M., Hersh, W. & Spackman, K. in *Proceedings of the AMIA 2002 Annual Symposium* 220-224 (2002).

21. LeBlanc, M. D. & D.Dyer, B. Teaching Together: A Three-year case study in Genomics. *Journal of Computing in Small Colleges* **18**, 85-95 (2003).
22. Bednarski, A. E., Elgin, S. C. R. & Pakrasi, H. B. An Inquiry into Protein Structure and Genetic Disease: Introducing Undergraduates to Bioinformatics in a Large Introductory Course. *Cell Biology Education* **4**, 207-220 (2005).
23. Campbell, A. M. Public Access for Teaching Genomics, Proteomics, and Bioinformatics. *Cell Biology Education* **2**, 98-111 (2003).
24. Centeno, N. B., Villa-Freixa, J. & Oliva, B. Teaching Structural Bioinformatics at the Undergraduate Level. *Biochemistry and Molecular Biology Education* **31**, 386-391 (2003).
25. Ai, Y.-C., Firth, N. & Jermiin, L. in *The China Papers: Tertiary Science and Mathematics Teaching for the 21st Century* 51-56 (2002).
26. Johnson, S. B. & Friedman, R. A. Bridging the gap between biological and clinical informatics in a graduate training program. *Journal of Biomedical Informatics* **In Press, Corrected Proof** (2006).
27. Hughey, R. & Karplus, K. in *Frontiers in Education Conference, 2001. 31st Annual* F2B-15-19 vol.2 (2001).
28. Maojo, V. & Kulikowski, C. A. Bioinformatics and Medical Informatics: Collaboration on the Road to Genomic Medicine? *J Am Med Inform Assoc*, M1305 (2003).
29. Friedman, C. P. et al. Training the Next Generation of Informaticians: The Impact of `BISTI' and Bioinformatics; A Report from the American College of Medical Informatics. *Journal of American Medical Informatics Association* **11**, 167-172 (2004).
30. Fetrow, J. S. & John, D. J. Bioinformatics and computing curriculum: a new model for interdisciplinary courses. *ACM SIGCSE Bulletin* **38**, 185-189 (2006).
31. Sahinidis, N. V. et al. Establishing a master's degree programme in bioinformatics: challenges and opportunities. *Systems Biology, IEE Proceedings* **152**, 269-275 (2005).
32. Wickware, P. Training in a hybrid discipline. **413**, 6-7 (2001).
33. Doom, T., Raymer, M., Krane, D. & Garcia, O. A Proposed Undergraduate Bioinformatics Curriculum for Computer Scientists. *SIGCSE BULLETIN* **34**, 78-81 (2002).
34. Pevzner, P. A. Educating biologists in the 21st century: bioinformatics scientists versus bioinformatics technicians. *Bioinformatics* **20**, 2159-2161 (2004).
35. Pearson, W. R. Training for bioinformatics and computational biology. *Bioinformatics* **17**, 761-762 (2001).
36. Kumar, A. Teaching Systems Biology: An Active-learning Approach. *Cell Biol Educ* **4**, 323-329 (2005).
37. Boyer, R. The new biochemistry: blending the traditional with the other. *Biochemistry and Molecular Biology Education* **28**, 292-296 (2000).
38. Feig, A. L. & Jabri, E. Incorporation of Bioinformatics Exercises into the Undergraduate Biochemistry Curriculum. *Biochemistry and Molecular Biology Education* **30**, 224-231 (2002).
39. Honts, J. E. Evolving Strategies for the Incorporation of Bioinformatics Within the Undergraduate Cell Biology Curriculum. *Cell Biology Education* **2**, 233-247 (2003).

40. Gibbons, N. J., Evans, C., Payne, A., Shah, K. & Griffin, D. K. Computer Simulations Improve University Instructional Laboratories. *Cell Biology Education* **3**, 263-269 (2004).
41. Doom, T., Raymer, M., Krane, D. & Garcia, O. Crossing the interdisciplinary barrier: a baccalaureate computer science option in bioinformatics. *IEEE Transactions on Education* **46**, 387-393 (2003).
42. Martz, E. Protein Explorer: easy yet powerful macromolecular visualization. *Trends in Biochemical Sciences* **27**, 107-109 (2002).
43. DeLano, W. L. (DeLano Scientific, San Carlos, CA, USA, 2002).
44. Campbell, A. M. Open Access: A PLoS for Education. *PloS Biology* **2**, 560-563 (2004).
45. Fox, J. A., Butland, S. L., McMillan, S., Campbell, G. & Ouellette, B. F. F. The Bioinformatics Links Directory: a Compilation of Molecular Biology Web Servers. *Nucl. Acids Res.* **33**, W3-24 (2005).
46. Pongor, S. & Landsman, D. Bioinformatics and the Developing World. *Biotechnology and Development Monitor* **40**, 10-13 (1999).
47. Ben-Dor, S. et al. in *Intelligent Systems for Molecular Biology (ISMB) Conference* (Brisbane, Australia, 2003).
48. Shachak, A., Ophir, R. & Rubin, E. Applying Instructional Design Theories to Bioinformatics Education in Microarray Analysis and Primer Design Workshops. *Cell Biolology Education* **4**, 199-206 (2005).
49. Abbot, C. in *Presentation at the Workshop on Education in Bioinformatics* (Edmonton, Canada, 2002).
50. Chorpothong, N. & Charmonman, S. An eLearning Project for 100,000 Students per Year in Thailand. *International Journal of The Computer, the Internet and Management* **12**, 111-118 (2004).
51. Lim, Y. P. et al. The S-Star Trial Bioinformatics Course: An On-Line Learning Success. *Biochemistry and Molecular Biology Education* **31**, 20-23 (2003).
52. Yeo, L. C., Deris, S., Illias, R. M. & Zaki, N. M. in *Advance Technology Congress* (Universiti Putra Malaysia, Serdang, Malaysia, 2003).
53. Cyranoski, D. Malaysian Biotechnology: the valley of ghost. *Nature* **436**, 620-621 (2005).
54. Rinaldi, A. More than the sum of their parts? *EMBO reports* **7**, 133-136 (2006).
55. Richard, R. J. A. & Sriraam, N. A Feasibility Study of Challenges and Opportunities in Computational Biology : A Malaysian Perpective. *American Journal of Applied Sciences* **2**, 1296-1300 (2005).
56. Raih, M. F. et al. NBBnet – The National Biotechnology and Bioinformatics Network: A Malaysian initiative towards a national infrastructure for bioinformatics. *Electronic Journal of Biotechnology* **6** (2003).
57. Raih, M. F. et al. in *LSGRID 2004* (eds. Konagaya, A. & Satou, K.) (Springer-Verlag, 2004).
58. Gilbert, D. et al. Bio-Mirror project for public bio-data distribution. *Bioinformatics* (2004).
59. Ranganathan, S. & Miyano, S. The Asia-Pacific Regional Perpective on Bioinformatics. *IEEE Intelligent Systems* (2001).

60. Salamone, S. Synamatix Touts Analysis Speed with Intelligent Database. *Bio-IT World (Online)* **http://www.bio-itworld.com/newsitems/2006/january/01-19-06-news-synamatix** (2006).

61. Birney, E., Bateman, A., Clamp, M. E. & Hubbard, T. J. Mining the draft human genome. **409**, 827-828 (2001).

62. Jones, P. B. C. The commercialization of bioinformatics. *Electronic Journal of Biotechnology* **3** (2000).

63. Leite, M. Public sphere and the sustainability of the bioinformatics promise. *Genetics and Molecular Research* **3**, 575-581 (2004).

64. Wickware, P. MAG's demise signals trouble for bioinformatics firms. *Nature Biotechnology* **18**, 144 (2000).

65. Black, G. C. & Stephan, P. E. Bioinformatics Training Programs Are Hot but the Labor Market Is Not. *Biochemistry and Molecular Biology Education* **33**, 58-62 (2005).

66. Russell, J. Bioinformatics Moves from Starring Role to Supporting Cast. *Bio-IT World (Online)* (2006).

67. Black, G. C. & Stephan, P. E. (Alfred P. Sloan Foundation, 2004).

68. Balaram, P. Bioinformatics: Blowing up a balloon. *Current Science* **82**, 1189-1190 (2002).

69. Tadmor, B. & Tidor, B. Interdisciplinary research and education at the biology-engineering-computer science interface: a perspective. *Drug Discovery Today* **10**, 1183-1189 (2005).

70. Garner, H. R. & Pertsemlidis, A. Computational biology: biological insight from Is and Os. *BIOSILICO* **1**, 27-35 (2003).

71. Crain, I. K. & Collins, N. M. in *International Conference on Biodiversity and Bioresources: Conservation and Utilization.* (Phuket, Thailand., 1997).