White Paper:

Bioinformatics Curriculum Recommendations for Undergraduate, Graduate and Professional programs.

July 25, 2002

bioinformatics.ca

PRESENTS THE
Canadian Bioinformatics Workshops



CANADIAN GENETIC DISEASES NETWORK
RÉSEAU CANADIEN DE MALADIES GÉNÉTIQUES

© Canadian Genetic Diseases Network

This paper can be found at bioinformatics.ca/docs/whitepaper.pdf

Executive Summary

Working drafts of the human genome and a number of model bacterial, animal, and plant genomes have recently been published. These genomes give us the "parts lists" that, when assembled, make up living organisms. Understanding how the parts are put together is a huge challenge, one that has been traditionally undertaken through laborious, time-consuming wet-lab experiments and trials.

When we study gene expression, protein function, and other biological processes, we observe the exquisite complexity of our genetic information. Our genes, of which we have tens of thousands, are built as strings of thousands to millions of chemical units in length. These genes, and the still more complex proteins encoded by them, hide patterns within their sequences that are the building blocks of biological organisms. Computers excel at recognizing and discovering patterns. The power of computational biology or bioinformatics, as it is more commonly called, can help unlock the "secrets" of biological function by analyzing genomic and protein data, comparing it with what we already know, and providing the hypotheses that form the basis for experimentation in the laboratory setting. It is this role of bioinformatics that holds promise to interpret the enormity of gene and protein information that is accumulating in computer databases worldwide. Bioinformatics is dramatically shortening the time required to generate working models of biological regulation, development and physiology, hypotheses for drug action, models for combating infectious disease and destructive insects and developments in crop hardiness and enhancement. It is these secrets of biological processes that bioinformatics can rapidly deliver to the laboratory bench to promote the diagnosis and treatment of disease, stimulate our pharmaceutical and agricultural economies and improve the well-being of Canadians.

Analyzing and synthesizing the vast accumulation of genomic information requires individuals who have the combined expertise in biology and computer science, disciplines that have not historically shared a common attraction. If we are to take advantage of genomic information that is growing at an exponential rate, Canada needs to offer students and experienced researchers effective bioinformatics education and training programs at the university and professional levels. Academia, government, and industry must combine forces to ensure that we create effective bioinformatics programs that are available throughout Canada. Such programs must have solid molecular biology, genetics, evolutionary biology, programming, algorithm design, database design and statistical analysis foundations. They also require hands-on laboratory components, and, in the ideal, work placements in university and industrial biotechnology laboratories.

This paper provides a context for bioinformatics, lays out a detailed model of undergraduate bioinformatics education, and makes the following recommendations:

1. Integrate bioinformatics themes into foundational life science and computer science courses at all universities where appropriate.
2. Emphasize practical lab and real-life bioinformatics problem-solving activities in all curricula
3. Create university and professional educational programs that acknowledge the multi-disciplinary nature of bioinformatics, the varied skill sets that are necessary and, during the transitional phase, the differing backgrounds of the individual students.
4. Ensure that professional training programs such as the Canadian Bioinformatics Workshop series continue to fill their mandate of providing "just-in-time" bioinformatics training and upgrading to allow researchers to remain current in this fast moving field
5. Educate university leaders in the necessity of having cooperation at all levels in the university community in order to implement bioinformatics curricula. Bridge across traditional academic boundaries.
6. Create an environment in Canada that fosters and retains the top bioinformatics researchers and educators.
7. Build foundational bioinformatics degrees that incorporate a broad range of science and computer science courses as well as core ethics, logic, and writing courses.
8. Embed ethical discussion and considerations of the social impact of genetic and biological technologies into all bioinformatics curricula so that students can understand and relate to the concerns of the general public
9. Encourage industry to play a role in bioinformatics education and training by participating in presentations to university classes, presenting research seminars, sponsoring co-op positions, internships, and supporting for both wet and dry labs (funding, software, hardware)
10. Promote the creation of specific scholarships and bursaries for people seeking careers in bioinformatics. These awards can be sponsored by industry, foundations, government, and professional societies
11. Include bioinformatics in any articulation activities among universities, colleges, and secondary schools in order to promote bioinformatics awareness to younger students
12. Work with the news media to inform the general public of bioinformatics "success stories."

The Promise of Bioinformatics

In 2001 the scientific community marked a significant milestone in the study of biology – the completion of the "working draft" of the human genome. The publications of the initial sequence and analysis of the human genome signalled a new beginning for modern biology, one in which the majority of biological and biomedical research will be conducted in a "sequence-based" fashion. This new approach, which uses the accumulated sequence data from all organisms – not just the human – as the basic foundation for new research, promises to lead quickly to major advances in the understanding of basic biological processes. The completion of the Human Genome Project, slated for April 2003, will truly herald the beginning of the "genome era," where the effective use of sequence-based data will have the potential to guide work leading to the prevention, diagnosis, and treatment of many disorders having a genetic basis.

On the natural resources front as well, bioinformatics is developing a major presence, opening up new possibilities. The completion of the genome of the plant, Arabidopsis thaliana (December 2000) and the first crop genome, rice (Oryza sativa) in April 2002, offers exciting new approaches to basic plant biology and to crop and forest improvement. New programs have been launched to exploit the considerable similarity between Arabidopsis and the related edible-oil rapeseed or 'Canola' – an important Canadian 'success' story on the world agricultural scene. In the same vein, rice, the staple food for two-thirds of the world's population is biologically related to the other cereal crops (e.g. wheat, barley, oats) that provide about 90% of our calories.

Canada has traditionally been reliant on the harvesting of its rich natural resources for its economic well being. However, our original forest and fishery resources are becoming depleted, and competition in the marketplace for our agricultural products is severe. If Canada is to move to a production model that is both economically successful and sustainable, it is essential that we learn to better manage and enhance our biological resources. Bioinformatics will be a key element in that strategy. The information presently being generated by new genomics projects focused on Canada's three major natural resource sectors is about to transform the way we steward, and derive value from, our fish stocks, forests and farmland. In a world that is steadily moving toward a global bio-based economy, fostering strong links between genomic information and resource managers will be a key step if Canada is to create a sustained economic comparative advantage.

Bioinformatics is a key element in these strategic approaches. The availability of DNA sequences in the last 25 years, and of protein sequences and structures for the last 40 years, has led to numerous developments in the areas of genomics and proteomics. Genomics is about genes, the information they encode and how they are regulated. Proteomics is about the tens of thousands of proteins in cells and understanding what they do and how they do it. However, it is bioinformatics that links genomes, proteomes and biological processes and allows us to study and extract knowledge from this information.

Bioinformatics or computational biology, as it is also known, allows researchers to sift through the massive data sets that are accumulating and identify information of interest. Without this analysis it would take an extremely long time before any useful information could be gleaned from these genomics and proteomic datasets.

The critical role that bioinformatics will play in the future has been highlighted in recent reports by the U.S. Biomedical Information Science and Technology Initiative (BISTI)[1]. Specifically, this task force underscored the fact that bioinformatics support will be essential as genomic science moves forward, since the biomedical community, among others, will need to have accessible computational methods with which to extract, view, and analyze genomic information, both at the population level and at the level of individual patients.

To place the future economic prospects of bioinformatics in perspective, a number of recent market research reports estimate the size of the bioinformatics market alone at anywhere between US$2 and 5 billion by 2005[2], signalling a continued expansion as sequence analysis, target validation, expanded drug development capabilities, and pharmacogenomics move into the forefront of modern science. In addition, the Federation of American Societies for Experimental Biology has recommended that the U.S. National Institutes of Health earmark a minimum of US$400 million for cutting-edge, multi-technology centres that will both supply and drive the genomics revolution.

It is bioinformatics that links genomes, proteomes and biological processes and allows us to study and extract knowledge from this information

Another example of this trend is the recent move within a major computer industry player to reposition itself. In August 2000, IBM announced an initial $US 100 million investment to spur business development in the life sciences, assuring its prominence as one of the emerging markets at the heart of IBM's growth strategy. The life sciences are important to IBM because the technology requirements of the life sciences industry directly correspond to the core of IBM's portfolio, including:

---

[1] The Biomedical Information Science and Technology Initiative, Working Group on Biomedical Computing, National Institutes of Health, June 1999. http://www.nih.gov/about/director/060399.htm

[2] Toner, B. *Will the Real Bioinformatics Market Please Stand Up?* Bioinform 6: 3-4, 2002. Market research organizations in the summary include Front Line Strategic Consulting, UBS Warburg, Frost & Sullivan, D&MD Reports, Strategic Directions International, Silico Research, and Oscar Gruss.

- high-performance information infrastructure to support high-speed computing and high-capacity storage
- data and knowledge management solutions to manage data and support scientific collaboration;
- clinical development integration solutions to improve efficiency and shorten the cycle time of clinical trials

This represents a market opportunity that could exceed $US 40 billion globally by 2004 (based on a study by Frost & Sullivan for IBM). In Canada, Frost and Sullivan estimate Life Sciences companies will spend more than US$2 billion on information technology between 1999 and 2004.

One of the earlier initiatives in bioinformatics was the European Bioinformatics Institute (EBI) (see www.ebi.ac.uk), that serves both industry and research by maintaining biological databases. EBI has also launched an industry support program aimed at assisting companies in adapting rapidly to advances in bioinformatics.

To address growing demands, the United Kingdom has established bioinformatics academic programmes at all levels. As well, five of the United Kingdom's research councils have banded together to cooperate in strengthening research in bioinformatics.

Japan has set up a National Institute for Genetics (www.nig.ac.jp/labs/home.html) that involves a Center for Information Biology. This centre includes laboratories for DNA data analysis, gene-product informatics, gene function research, and molecular classification.

These activities mark our arrival at a critical juncture in genomic science, where the focus of research will turn away from large-scale sequencing and assembly, and towards identifying biological applications that will have tangible academic and economic benefits. Making these advances, however, will require that Canada quickly address the "talent gap," the current dearth of individuals who can successfully perform bioinformatics research. While a small number of training programs have appeared in the past few years, these programs cannot meet the demand for bioinformatics professionals that we will face in the next 10-15 years. This is a call for dramatic developments in bioinformatics training. New training initiatives must provide for rigorous degree programs that instil in their students the ability to properly formulate biological questions and develop tractable strategies for unequivocally answering them.

The key to creating new, forward-looking bioinformatics training programs rests in the development of relationships between bioinformatics specialists and faculty in allied fields, since more biological science will now be piloted on computers, and computer-generated results will need to be verified in the laboratory. The call to action will involve faculty in the life sciences (biomedical to agriculture), computer science, mathematics, physics, and other closely-related disciplines, since instruction in each of these fields will be critical to providing trainees at the undergraduate and graduate level with the proper

arsenal of skills needed to perform biological science. While this vision is a noble one, it challenges host institutions to blur the traditional boundaries among academic departments and even among institutions themselves. Without broad-based training such as this, the type of individuals that will be needed to support – and *lead* – Canada's growing biotechnology industry will not be generated.

Challenges

There are a number of major challenges facing us. Foremost is to find enough bioinformatics-trained personnel to fill positions now and for the next few years until university programs launched in 1999/2000 start graduating students enrolled in the few undergraduate bioinformatics programs in the country.

For example, the bioinformatician needs for Genome Canada projects alone is estimated to be 215 positions in the first two years (2001/2002), and this is predicted to grow to 262 after four years. In a 2001/2002 bioinformatics survey (http://bioinformatics.ca/survey/indexbk.php) carried out by the Canadian Genetic Diseases Network (CGDN), 77 researchers with hiring responsibilities estimated their current demand for bioinformaticians to already be greater than 487, with 1083 positions needing to be filled by 2006.

A second challenge is a lack of models of good undergraduate curricula in bioinformatics. Various programs have been proposed, including one by University of Stanford professor Russ Altman in *A Curriculum for Bioinformatics: The Time is Ripe*, published in 1998[3]. However, Altman was concerned with educating graduate students in bioinformatics, whereas it has become clear that bioinformatics education needs to begin at the undergraduate and even the pre-university level.

In Canada at the present time, there are 14 named undergraduate bioinformatics programs being offered in Alberta, Ontario, and Quebec (http://bioinformatics.ca/program_listings.php). The other universities are in the process of developing programs, but often lack necessary faculty with a strong understanding of bioinformatics who could help to design effective programs. There is also little consensus on what comprises effective bioinformatics education at the university and post-university level.

A third challenge for the bioinformatics community is "retrofitting" the large pool of active life science researchers who were educated before the word "bioinformatics" was even coined. This includes most current undergraduate and graduate students who still receive no specific bioinformatics education. These cohorts are ill prepared to fully

[3] Altman, R Bioinformatics, **14**(7), 549-550 (1998)

exploit the continually increasing and changing data coming from genomics and proteomics initiatives. Upgrading programs and continuing education programs will therefore have to be an essential part of bioinformatics training.

In a 2001/2002 bioinformatics survey carried out by the CGDN, 77 researchers with hiring responsibilities estimated their current demand for bioinformaticians to already be greater than 487, with 1083 positions needing to be filled by 2006.

Is there money to pay for the cost of bioinformatics training? Researchers, life science graduates, and, computer science graduates have seen the promise of an exciting and rewarding career in bioinformatics and have been taking professional upgrading workshops from the Canadian Bioinformatics Workshop (CBW) series (http://bioinformatics.ca) to fill the gaps in their knowledge and abilities. The CBW had 130 applicants for 50 positions for its February 2002 Bioinformatics Workshop. At CDN$2500 per person, that represented a CDN$325,000 demand for bioinformatics training. The fall 2001 Cold Spring Harbor workshop "Bioinformatics: Writing Software for Genome Research" (http://nucleus.cshl.org/meetings/2002courses.htm) saw 200 applicants competing for 20 positions. The cost for this workshop was US$2300 per person so there is definitely "money on the table" for bioinformatics training. The lesson is that money will be found for training if it overcomes barriers to accomplishment and competitiveness. This premise must be carried to industry and governments as the critical motivation to move forward at the urgent pace required to competitively exploit the burgeoning collections of genome and protein information.

Solutions

If we are to take advantage of the exponential growth of genomics and proteomics information, Canada needs to offer students and experienced researchers effective bioinformatics education and training programs at the undergraduate, graduate and professional levels. Academia, government, and industry must combine forces to ensure that we create effective bioinformatics programs that are available throughout Canada.

Undergraduate programs

There is a need to more clearly delineate the minimum requirements for bioinformatics training in Canada. Bioinformatics programs should include a solid foundation of courses in such core disciplines as biology, computer science, math, and molecular biology. Ideally, bioinformatics would be introduced as a topic in many of these core courses to emphasize the necessary integration of bioinformatics into these disciplines. However, specialized bioinformatics courses will also be required. New approaches may be needed to help students cross the boundaries from more process- and algorithm-

oriented thinking (computer science) to more experiment-oriented thinking (life sciences), and vice versa.

The experience that was gained from teaching the Canadian Bioinformatics Workshop series since 1999 indicates that a base level of common training will enable bioinformaticians to carry out their tasks more effectively. Designing algorithms for use in bioinformatics and then building the tools that use these algorithms cannot be carried out by computer scientists in isolation; they must understand sufficient biology in order for their algorithm/tool to be effective. In order for the biologist to interact effectively in this team design process, they must understand computer programming and algorithm development at a sufficient level to be able to "speak the same language" as the programmer/software engineer.

Core concepts and technical skills for undergraduate bioinformatics curriculum

Because the field of bioinformatics is new and rapidly changing, the undergraduate curriculum should emphasize the core concepts and technical skills that are the least likely to become redundant in the next few years. This set of core concepts can be incorporated into a range of courses, depending on how a particular university decides to mount its bioinformatics program. These core concepts and skills, along with the suggested undergraduate curriculum in the next section, will also allow institutions developing an undergraduate program in bioinformatics to more easily capitalize on existing courses. The core concepts outlined provide the breadth appropriate for an undergraduate curriculum but also provide guidance for development of graduate and professional programs where these same concepts can be treated in a more advanced and focused manner.

Table 1: Core concepts and technical skills for bioinformatics

Biology core concepts	Statistics and computer science core concepts
Foundations of molecular biology: how genetic information is transmitted and used in living cells	Discrete math, linear algebra and multivariable calculus
Systems biology: organism, cells, subcellular organization, genetic material, environment, and the temporal and spatial considerations linking these	Generalized statistics and specialized distributions
Basic evolutionary theory	Applied probability (mass function, density, expectation, variance)
Common laboratory biology techniques, their limitations, and their computational aspects (e.g. sequencing, structure determination, DNA arrays, 2D protein gels, etc.)	Statistical approaches for computational biology
Computational strategies for assigning putative function to protein sequences, determining gene families and orthologous relationships	Statistical inference and estimation (likelihood-based and Bayesian)
Navigation of genomic data (e.g. interrogate chromosomal sequence, explore gene predictions, assess putative protein function, determine gene and	Introduction to empirical problem solving

Biology core concepts	Statistics and computer science core concepts
protein expression profiles as well as the genetic or biochemical pathways affected)	
Established methods for molecular structure analysis and simulation, such as structure generation, structure comparison, structure modeling, and molecular dynamics	Modeling (generalized linear models, hidden Markov models, classification and discriminant analysis, neural networks, cluster analysis)
Sequence comparison (pairwise alignment, multiple alignment)	Role of models, such as probability models, for representing biological data, and awareness of common methods for analyzing data in the context of a model
Phylogenetic analysis	Algorithms for combinatorial optimization, such as bounded search algorithms, neural network algorithms, dynamic programming, stochastic local search algorithms (such as genetic algorithms, Monte Carlo algorithms, simulated annealing), and expectation maximization
Fragment and map assembly and combinatorial approaches to sequencing	Probabilistic machine learning
RNA secondary and tertiary structure prediction	Design and analysis of algorithms and data structures
Sequence feature extraction/annotation	Design, implementation, integration, querying and searching (such as SQL) of biological databases/knowledge bases
Single Nucleotide Polymorphism (SNP) detection and utilisation	Databases
Gene expression analysis	Strategies for data-mining and visualisation
Regulatory network modelling	Graph theory
Protein *ab initio* structure prediction	Programming languages such as C, C++ and Java and a scripting language, such as Perl or Python
Protein structure comparison	Encryption
Protein homology modeling	Data security
Protein threading (2D and 3D)	**Other core concepts**
Protein molecular dynamics	A brief history of bioinformatics
Peptide mass fingerprinting and mass spectrometry analysis	Critical reading and writing/science communication
2D gel annotation and analysis	Bioethics
Image (spot) analysis as it pertains to gels and microarrays	
Practical applications of bioinformatics such as LIMS, pharmacogenomics, drug development, diagnostics (arrays)	

Suggested undergraduate bioinformatics curriculum

Table 2 lists a set of core and elective courses that could be included in an undergraduate bioinformatics degree that is balanced, but covers both computational and biological material quite comprehensively. We believe that universities will offer several types of bioinformatics undergraduate degrees including

1. degrees in Bioinformatics that are balanced between computational and life sciences offerings
2. degrees in Bioinformatics that are more computer science focused
3. degrees in Bioinformatics that are more biologically focused

We maintain, however, that there should be a primary focus on "option 1". The intent is to train students at the undergraduate level to be well rounded and to use this degree as a base for further studies that are either more computational or biological in nature. We encourage curriculum developers to draw from this set of proposed courses to design a bioinformatics curriculum that capitalizes on their existing resources. At the same time, the design process may highlight specific areas where additional expertise needs to be added to the instructional staff. We also encourage institutions to work across traditional academic boundaries in designing bioinformatics programs; the best programs will result when traditional discipline "silos" are bridged.

It is worth emphasizing that a single four-year degree program is unlikely to contain *all* of these courses. Thus, while the proposed set of required and elective courses is intended as a guide in developing a four-year undergraduate degree, circumstances may dictate that other models for the undergraduate degree are more appropriate for specific institutions.

Finally, we recommend that the following elements be an integral part of any undergraduate bioinformatics degree:

- bioinformatics research experience in academic or industry labs (summer jobs, co-op placements)
- considerable molecular biology wet-lab experience
- team projects

Table 2: Proposed undergraduate bioinformatics degree: course requirements

Year I and II requirements	Year III and IV requirements
Calculus I and II	Statistics (i.e. Analysis of Experimental and Observational Data)
Elementary Linear Algebra	Metabolism
Discrete Probability and Statistics	Molecular Biology
Modern Physics and Mechanics	Microbiology
Intro Biology I and II	Molecular Biology and Biochemistry

Year I and II requirements	Year III and IV requirements
	Laboratory
Genetics	Molecular Physiology
Molecular Biology and Biochemistry	Bioinformatics
General Chemistry & Laboratory I	Protein Structure & Function
Organic Chemistry & Laboratory I	Genomic Analysis
Introduction to Programming	Molecular Phylogeny and Evolution
Data Structures	Algorithms and data structures
Software Engineering	Databases I
Introduction to Computer Design	Artificial Intelligence
Discrete Mathematics I and II	Computational Biology
Critical Thinking or other writing or logic course	Ethics or social implications course
	Thesis project course

Electives – based on specialization – biological, statistical or computational sciences		
Simulations	Biostatistics	Nucleic Acids
Multimedia systems	Statistics for bioinformatics	Physical Biochemistry
Algorithms	Numerical analysis	Immunology
Software engineering	Web-based information systems	Proteomics
Distributed systems	Human-computer interactions	Advanced genetics
Computer graphics	Data communications and networking	Biotechnology
Database II	Computational physics	Virology
		Biochemistry II

Canadian graduate programs

Initially, graduate students may be drawn from a pool of students who don't have a thorough grounding in bioinformatics. These students may need upgrading in courses in bioinformatics, computer science, and evolutionary biology, for example. We would encourage institutions to be creative in assembling the required set of courses so that these students are able to pursue challenging advanced coursework as part of their degree.

Once students start entering graduate studies with an undergraduate bioinformatics degree, then the graduate degree will be governed by the same rules as other institutional graduate degrees. We suggest that graduate bioinformatics degrees be research and thesis focussed to emphasize that the study of bioinformatics is about science and not solely about tools.

A primary goal should be to create a bioinformatics culture that will make it attractive for bioinformaticians to work in Canada. We are already making good strides in this direction. See http://bioinformatics.ca/people/ for a list of current Canadian bioinformatics researchers.

Professional upgrading – Canadian Bioinformatics Workshop series

The Canadian Bioinformatics Workshop series was a direct outgrowth of the report of a Bioinformatics Task Force, created in 1998 by the BiotecCanada Human Resources Council (now the Biotechnology Human Resource Council, BHRC). Initial funding from the BHRC enabled CGDN to offer a pilot series of workshops from August 1999 to June 2000. The four workshops were:

- Bioinformatics (80 hours of lectures and hands-on laboratory, i.e. computer, sessions)
- Genomics (35 hours as above)
- Proteomics (35 hours as above)
- Developing the Tools (35 hours as above)

A Certificate in Bioinformatics was established, accredited by the Universities of British Columbia, Toronto, and New Brunswick. The Certificate requires students to complete Bioinformatics and two of the other three workshops.

Since 1999, 54 students have completed their certificate, 491 workshop-seats have been filled, and 300 individual students have completed one or more workshops. Demand for these workshops remains high. Students are already registering for the 2003 series (see http://bioinformatics.ca). The series has been excellent at providing much needed bootstrap training for a wide range of students and researchers from industry, government, and academia. The need for such professional upgrading will continue as current bioinformatics techniques become replaced by techniques that are not yet envisioned. This type of training complements the undergraduate and graduate programs and allows for a rapid deployment of rapidly changing material.

Summary of Recommendations

To prepare individuals for a career in which bioinformatics knowledge is key, we recommend the following actions:

1. Integrate bioinformatics themes into foundational life science and computer science courses at all universities where appropriate.
2. Emphasize practical lab and real-life bioinformatics problem-solving activities in all curricula
3. Create university and professional educational programs that acknowledge the multi-disciplinary nature of bioinformatics, the varied skill sets that are necessary and, during the transitional phase, the differing backgrounds of the individual students.

4. Ensure that professional training programs such as the Canadian Bioinformatics Workshop series continue to fill their mandate of providing "just-in-time" bioinformatics training and upgrading to allow researchers to remain current in this fast moving field
5. Educate university leaders in the necessity of having cooperation at all levels in the university community in order to implement bioinformatics curricula. Bridge across traditional academic boundaries.
6. Create an environment in Canada that fosters and retains the top bioinformatics researchers and educators.
7. Build foundational bioinformatics degrees that incorporate a broad range of science and computer science courses as well as core ethics, logic, and writing courses.
8. Embed ethical discussion and considerations of the social impact of genetic and biological technologies into all bioinformatics curricula so that students can understand and relate to the concerns of the general public
9. Encourage industry to play a role in bioinformatics education and training by participating in presentations to university classes, presenting research seminars, sponsoring co-op positions, internships, and supporting for both wet and dry labs (funding, software, hardware)
10. Promote the creation of specific scholarships and bursaries for people seeking careers in bioinformatics. These awards can be sponsored by industry, foundations, government, and professional societies
11. Include bioinformatics in any articulation activities among universities, colleges, and secondary schools in order to promote bioinformatics awareness to younger students
12. Work with the news media to inform the general public of bioinformatics "success stories."

Acknowledgements

Core faculty of Canadian Bioinformatics Workshops

In developing the White Paper, we drew on a number of resources. The core faculty (see http://bioinformatics.ca/faculty.php) of the CBW, comprising six prominent Canadian bioinformaticians, were consulted throughout the process of researching and generating this document. They are:

- Fiona Brinkman, Molecular Biology and Biochemistry, Simon Fraser University.
- Christopher Hogue, Samuel Lunenfeld Research Institute, Mt. Sinai Hospital, and Department of Biochemistry, University of Toronto.
- Steven Jones, BC Genome Sequences Centre, Vancouver, BC

- François Major, Département d'Informatique et Recherche Opérationnelle, Université de Montréal
- Francis Ouellette, CGDN Bioinformatics Core Facility, Centre for Molecular Medicine and Therapeutics (CMMT) Bioinformatics Core Facility, UBC Bioinformatics Centre.
- David Wishart, Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta.

Focus group

In March 2002, we held a focus group meeting with Canadian bioinformatics scientists in industry and academia to receive their perspectives on bioinformatics education and training. They included:

- Rob Abbott, Abbott Strategies
- Ryan Brinkman, Xenon Genetics Inc.
- Stefanie Butland, Centre for Molecular Medicine and Therapeutics, University of British Columbia
- Jenny F. Bryan, Department of Statistics and Biotechnology Laboratory, University of British Columbia
- Anne Condon, Department of Computer Science, University of British Columbia
- Bill Crosby, Plant Biotechnology Institute, National Research Council, Canada
- Margaret Ellis, Canadian Bioinformatics Workshop series
- Warren Gallin, Department of Biological Sciences, University of Alberta
- Geoffrey Hicks, Manitoba Institute of Cell Biology, University of Manitoba
- Steven Jones, BC Genome Sequences Centre
- Francis Ouellette, CGDN, CMMT, UBC
- Chris Upton, Department of Biochemistry and Microbiology, University of Victoria
- David Wishart, Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta
- Michael Murphy, Department of Microbiology and Immunology, University of British Columbia

Validation meeting

In April 2002, we held a Validation Meeting with Canadian and U.S. bioinformatics scientists in industry and academia.

- Michael Cherry, Department of Genetics, School of Medicine, Stanford University
- Francis Ouellette, CGDN, CMMT, UBC
- Andy Baxevanis, National Human Genome Research Institute, National Institutes of Health, U.S.A.
- Roy Gravel, Faculty of Medicine, University of Calgary

- Janice Glasgow, Department of Computer Science, Queens University
- Claire Thifault, Executive Director, Biotechnology Human Resource Council
- Margaret Ellis, CBW/CGDN
- Keri Gammon, CBW/CGDN
- Rob Abbott, Abbott Strategies

We would like to thank the Biotechnology Human Resource Council (BHRC) for their generous financial support of this project

---