

Bioinformatics:



**An instructional opportunity for
academic science and
engineering libraries**

Erja Kajosalo

What will I talk about today...

- Present findings of a study on bioscience at MIT
- Discuss academic science/engineering libraries roles in teaching bioinformatics tools

Background

- "... a paradigm shift in biological research driven by the emergence of molecular biology in combination with computational techniques."
- "Research, education and infrastructure development at MIT increasingly involve efforts with a primary or significant bioscience component."

Proposal for Support of the CSBi (MIT) Technology Platform", 2003

Bioscience at MIT and the Engineering and Science Libraries

- ❑ A team of ESL librarians has studied the growth and change of bioscience and its place at MIT.
- ❑ Reviewed the current ESL programs for supporting bioscience
- ❑ Studied other academic library programs
- ❑ Spoke with members of the MIT community, invited experts to visit our campus, and hosted discussions within ESL.
- ❑ Begun to develop new partnerships with bioscience information experts at MIT, and exploring partnerships with Harvard University to support bioscience information access and services.

Faculty involvement in bioscience by departments

- Core bioscience departments and divisions at MIT
 - Biology
 - Brain and Cognitive Sciences (BCS)
 - Health Sciences and Technology Division (HST)
- Other departments involved in bioscience research
 - Biological Engineering [Division] (BE, Course 20), 1998-
 - Chemistry (55% of faculty)
 - Chemical Engineering (53%)
 - Nuclear Engineering (30%)
 - Materials Science (27%)
 - EECS (25%).

Faculty involvement in biosciences by areas of interdisciplinary study

- ❑ *Biological Engineering* – BE, Biology, Chemistry, Chem. Eng, Materials Science, Mech Eng.
- ❑ *Bioinformatics* - Biology, EECS, Chemistry, HST, Math
- ❑ *Computational biology* - Biology, EECS, BE, Math, EECS
- ❑ *Computational neurosciences* - BCS, Biology, Physics, Math, EECS
- ❑ *Environmental Science* - Biology, EAPS, Civil Eng
- ❑ *Genomics and proteomics* - Biology, Chemistry, EECS, Math
- ❑ *Gene sequencing* - Biology, EECS, Math
- ❑ *Medical informatics* - Biology, HST, Math
- ❑ *Systems biology* - Biology, EECS, Physics, Chemistry, BE, Math
- ❑ *Statistical models of biomedical applications* - Biology, BCS, EECS, Chemistry, Math (probability and statistics)

Faculty involvement in bioscience by courses taught

- Several departments outside of biological engineering and the bioscience core offer bioscience courses:
 - Chemical Engineering (26 courses)
 - Civil and Environmental Engineering (19)
 - Electrical Engineering & Computer Science (22)
 - Mechanical Engineering (20)
 - Chemistry (7)
 - Aeronautics & Astronautics; Physics, Materials Science & Engineering (each 5)
 - Earth and Planetary Sciences (4)

Faculty involvement in bioscience through centers, labs, and programs

- ❑ Identified over 20 interdisciplinary centers, labs and programs at MIT with a bioscience focus.
- ❑ Some of the newest programs illustrate the rapid growth of interdisciplinary bioscience activity on the MIT campus:
 - Broad Institute (2003)
 - Computational and Systems Biology Initiative (CSBi) – a virtual center (2002)
 - The McGovern Institute for Brain Research (2000)
 - Picower Institute for Learning and Memory (1994 and 2002).

We knew it was big...

- ... but it was still surprising to realize the full scope and breadth of this activity, including the nature and scope of bioscience work within the School of Engineering
- Final report
 - appoint a liaison for biological engineering division
 - Create a more formal collaboration on biosciences between ESL librarians (e.g. small task forces working on specific projects)

Bioinformatics - definition

- Bioinformatics merges molecular biology, computer science, and information technology into a single discipline.

Bioinformatics & E/S Libraries?

- ❑ Molecular biology in combination with computational techniques has drastically changed the amount and type of data in biology research, as well as how research is carried out.
- ❑ A single microarray experiment can generate hundreds of thousands data points, *GenBank* has over 46 million sequence records (NCBI, 2006) and the *Protein Data Bank* has records for over 35,000 proteins and other biological macromolecules (PDB, 2006).
- ❑ The appearance of “-omes” and “-omics” words also illustrates to the data explosion in biology and chemistry, and the consequent need researchers have to use new bioinformatics and cheminformatics tools.
- ❑ Question – what role could engineering/science libraries play in bioinformatics?

Bioinformatics portals in academic libraries – mostly medical

- ❑ Cornell University: Mann Library <http://www.mannlib.cornell.edu>
- ❑ Johns Hopkins University: Welch Medical Library <http://www.welch.jhu.edu>
- ❑ Purdue University Libraries <http://www.lib.purdue.edu>
- ❑ University of California—San Diego: Science and Engineering Library <http://scilib.ucsd.edu>
- ❑ University of Florida: Health Science Center Libraries <http://www.library.health.ufl.edu>
- ❑ University of Minnesota, Twin Cities: Bio-Medical Library <http://www.biomed.lib.umn.edu>
- ❑ University of Pittsburgh: Health Sciences Library System <http://www.hslls.pitt.edu>
- ❑ University of Southern California: Norris Medical Library <http://www.usc.edu/hsc/nml/>
- ❑ University of Utah: Eccles Health Sciences Library <http://medlib.med.utah.edu>
- ❑ University of Virginia: Claude Moore Health Sciences Library <http://www.healthsystem.virginia.edu/internet/library/>
- ❑ University of Washington: Health Sciences Library, Health Links <http://healthlinks.washington.edu>
- ❑ Vanderbilt University: Eskin Biomedical Library <http://www.mc.vanderbilt.edu/biolib/>
- ❑ Virginia Commonwealth University: VCU Libraries <http://www.library.vcu.edu>
- ❑ Washington University in St. Louis: Becker Medical Library <http://library.wustl.edu>

Messersmith et al,

A Web-based assessment of bioinformatics end-user support services at US universities

J Med Libr Assoc 94(3) July 2006

When in doubt, survey...

- ... other academic E/S libraries
 - The main conclusion - few academic E/S libraries have well-developed bioinformatics support models
 - Fewer have a systematic strategy to support their communities' bioinformatics needs
 - Getting more and more questions though, e.g. about specific gene sequences or how to identify the protein binding sites for a certain drug.

When in doubt, survey...

- Only 15% of respondents indicated that they have a “bioinformatics specialist” in their library.
- Some of these specialists are librarians with a strong subject specialty, but few places have PhD level specialists (whether with or without a library degree as well).
- Most of these positions were funded by the library. Only one response indicated that a library position was funded by a different department, at a higher salary level.
- 40% - bioinformatics specialists are located outside of the libraries, but that they collaborate informally with libraries.

When in doubt, check the literature...

- ❑ Many medical libraries support bioinformatics services, so no surprise that a lot of the literature focuses on medical library based services.
- ❑ For example, we found several articles in the literature about how to create a bioinformatics librarian position in a medical library (e.g. Tennant, 2005), how to create a medical library-based services program (e.g. Yarfitz and Ketchell, 1999), or how to train medical librarians to handle bioinformatics questions (e.g. Lyon et al, 2004).

When in doubt, check the literature...

- The primary categories of bioinformatics questions in such settings are about (Yarlitz and Ketchell, 2000)
 - locating resources, programs, or databases
 - technical questions about identifying appropriate tools for research
 - analytical questions about planning experiments or assisting in-depth with data analysis.

Service Levels

- Basic level
- Intermediate level
- Expert level
- Details adopted from MacMullen (2004)

When in doubt, check the literature...

- **Basic level** – ability to locate resources, ability to offer introductory training programs, and knowledge of some databases.
- Support activities include
 - General education and training – assist researchers with managing citations and personal collections, and acquiring and organizing background research for grant applications.
 - Expand print and online collections – a traditional collection development role.
 - Portal of curated web resources - create and maintain web page(s) of bioinformatics resources.
 - Offer use of the libraries and libraries training room as a 'neutral space' - meeting, collaboration, and instruction space to interdisciplinary faculty and students.

When in doubt, check the literature...

- ❑ **Intermediate level** - some staff members able to answer technical questions, such as how to identify appropriate tools for research.
- ❑ Go beyond beyond the basic level by offering
 - Course-integrated instruction – at both undergraduate and graduate levels.
 - Consultation services and liaison programs.
 - Bioinformatics computing resources and analysis – some health science libraries have implemented services focused on the data analysis component of bioinformatics, including hardware, software and consulting; at a minimum, provide expert advice with trials, license issues, and networking databases.

When in doubt, check the literature...

- **Expert level** – an expert in the library is able to handle analytical questions like how to plan experiments or offer in-depth assistance with data analysis.
- Go beyond the intermediate level by offering
 - Specialized in-depth instruction on resources, classroom & web-based tutorials, and hosted seminars by experts.
 - Dedicated specialist(s) - not necessarily librarians – with bioinformatics and/or cheminformatics background.
 - Repeatable/reusable literature searches and integration - create online tools that enable knowledge sharing across time and disciplinary boundaries.
 - Anticipatory synthesis - libraries could take a more proactive approach to organizing and synthesizing information

What do ESL have already?

- ❑ Have a good basic knowledge of bioinformatics resources
- ❑ Five ESL librarians attended the three-day NCBI course called "Introduction to Molecular Biology Information Resources"
- ❑ Biology librarian attended the week-long NIH Biomedical Informatics program in Woods Hole, MA.
- ❑ ESL collaborations with outside experts, have produced well-attended bioinformatics training sessions demonstrating a need for bioinformatics services on all levels from undergraduates to postgraduates.
 - Collaboration with a Harvard Librarian (PhD in molecular biology)
 - MIT Cancer Research Center
 - Broad Institute

What next @MIT Libraries: team of two

- ❑ Learn, learn, learn...
- ❑ Develop bioinformatics workshops with the current staff
 - Basic introduction to NCBI bioinformatics resources (geared towards undergrads)
 - Intermediate session using NCBI tools (e.g. BLAST)
- ❑ Create a bioinformatics services & resources portal
- ❑ Continue and expand collaboration with on-campus & off-campus experts
- ❑ Sponsor NCBI advanced training sessions at the MIT - <http://www.ncbi.nlm.nih.gov/Education> - in collaboration with some departments and/or centers (e.g. NCBI mini courses).
- ❑ Determine if ESL should move to the expert level support for bioinformatics services, and if so, how to accomplish this

Selected Readings

- Focus Issue July 2006 of J Med Libr Assoc:
 - Broad issues to consider for library involvement in bioinformatics.
 - A Web-based assessment of bioinformatics end-user support services at US universities.
 - A management case study: challenges of initiating an information service in molecular biology and genetics
 - Design and implementation of a library-based information service program in molecular biology and genetics at the University of Pittsburgh.
 - Developing library bioinformatics services in context: the Purdue University Libraries bioinformationist program.
 - The University of Washington Health Sciences Library BioCommons: an evolving Northwest biomedical research information support infrastructure.
 - Carving a niche: establishing bioinformatics collaborations.
 - Vignettes: diverse library staff offering diverse bioinformatics services.
 - GeneTests: an online genetic information resource for health care providers.
 - Challenges and strategies of the Genetics Home Reference.

Selected Readings

- Alpi, K. (2003). Bioinformatics training by librarians and for librarians: Developing the skills needed to support molecular biology and clinical genetics information instruction computer file. *Issues in Science & Technology Librarianship*, 37(Spring).
- Hemminger, B. M. (2005). Introduction to the special issue on bioinformatics. *Journal of the American Society for Information Science and Technology*, 56(5), 437.
- Hemminger, B. M., Losi, T., & Bauers, A. (2005). Survey of bioinformatics programs in the United States. *Journal of the American Society for Information Science and Technology*, 56(5), 529.
- Lyon, Jennifer, Giuse, Nunzia Bettinsoli, Williams Annette, Koonce Taneya, and Walden Rachel. (2004). A model for training the new bioinformationist. *Bulletin of the Medical Library Association* 92(2): 188-195.
- MacMullen, W. J., & Denn, S. O. (2005). Information problems in molecular biology and bioinformatics. *Journal of the American Society for Information Science and Technology*, 56(5), 447.
- MacMullen, W. J., Vaughan, K. T. L., & Moore, M. E. (2004). Planning bioinformatics education and information services in an academic health sciences library. *College & Research Libraries*, 65(4), 320-333.
- Tennant, M. R. (2005). Bioinformatics librarian: Meeting the information needs of genetics and bioinformatics researchers. *Reference Services Review*, 33(1), 12.
- Yarfitz, Stuart, and Ketchell Debra S. (2000). A library-based bioinformatics services program. *Bulletin of the Medical Library Association*, 88(1):36-48.

Acknowledgements

- Tracy Gabridge
- Louisa Worthington Rogers
- Howard Silver
- Amy Stout

Questions?
