

MARK D. LEBLANC

May 27, 2011

Department of Mathematics and Computer Science
Wheaton College, Norton, MA 02766
Work: (508) 286-3970 Home: (508) 285-6307
mleblanc@wheatoncollege.edu
<http://cs.wheatoncollege.edu/mleblanc>

Educational Background

- Ph.D.** Systems Design Engineering, University of New Hampshire, Durham. May, 1993.
Thesis title: *A Cognitive Model of "Mathematical Reading"*
- M.S.** Computer Science, University of New Hampshire, Durham. December, 1987.
- B.A.** Computer Science, University of Maine, Orono. May, 1984.

Employment History

- Professor of Computer Science** July, 2004 to present.
- Meneely Chair** July 2010 to present.
Department of Mathematics and Computer Science, Wheaton College, Norton, MA.
- Visiting Fellow of Computer Science** University of Wollongong, Australia, July 2004 to July 2005.
- Associate Professor of Computer Science** Wheaton College, June, 1998 to June, 2004.
- Department Chair** Mathematics & Computer Science, 1996-97, 1998-1999, 2000-2001, and 2002-2003.
- Software Engineer** Foxboro Company, June, 1999 to Sept. 2000 during full-year sabbatical.
- Assistant Professor of Computer Science** Wheaton College, July, 1993 to June, 1998.
- Software Engineer** ITEK Optical Systems, Lexington, MA June, 1984 to August, 1985.

Projects at a Glance

Authorship Attribution and Genomic Signatures – July 2004 to present.

Co-leader of an interdisciplinary research team of biologists and statisticians; designed and developed a pipeline of experimental software (C++ STL, Perl, bash, R) and local database of 700+ microbial genomes to apply techniques of genomic signature to open questions in comparative genomics.

Lexomics Research Group – January 2007 to present.

Group leader of an interdisciplinary research team comprised of an Anglo-Saxon scholar and statistician; designed and implemented a suite of experimental software to mine the tangled web of Old English authorship in poetry and prose.

Genomics Research Group -- 1998 to present.

This interdisciplinary, faculty-student research uses a linguistic metaphor to guide the work of merging computational and visualization approaches to the problem of discovering and deciphering the grammar and syntax of gene regulation. Software includes a web interface to a "DNA dictionary" or motif lexicon (based on the online Oxford English Dictionary) that allows users to search for words throughout an organism's genome and a "favGene" tool to allow researchers to query the DNA sequences in regions near their favorite collections of genes. <http://genomics.wheatoncollege.edu>

Multi-threaded Access to Historical Data

A multi-threaded DLL (dynamic-link library) written in C++ to find and collect historical real-time data from networked servers.

Scholarship in the Queue

Publications (submitted, under consideration):

Downey, Sarah J., Michael D.C. Drout, Michael J. Kahn and Mark D. LeBlanc. 'Books Tell Us:'
Lexomic and Traditional Evidence for the Sources of *Guthlac A*. *Modern Philology*.
[submitted Feb. 2010; under consideration].

Experiments (in progress):

LeBlanc, M.D., Kahn, M., Dyer, B.D. Genomic signature and horizontal transfer in microbial genomes.

Funding

External

April 01, 2011 – March 30, 2013: National Endowment for the Humanities – NEH PR-50112011 (Co-PI with Mike Drout (PI), English and Mike Kahn (Co-PI), Statistics, Wheaton College). “Lexomic Tools and Methods for Textual Analysis: Providing Deep Access to Digitized Texts.” [\$178,000 award]

July 01, 2008 – June 30, 2010: National Endowment for the Humanities – NEH HD-50300-08 (PI with Co-PIs Mike Drout, English and Mike Kahn, Statistics, Wheaton College)
“Pattern Recognition through Computational Stylistics: Old English and Beyond.”

May 2004 – May 2007: NSF DUE 0340761 (Co-PI with Betsey Dyer, Biology, Wheaton College).
“Teaching Genomics to Undergraduate Computer Science and Biology Majors:
A model involving infusion and strategic linking.”

Sept 2003 – May 2004: SIGCSE Special Projects Grant (Co-PI with Matt DeJongh, Hope College, MI).
“Course Materials and Teaching Models for Bioinformatics.”

Jan 2002 – June 2003: NSF DUE 0126643 (Co-PI with Betsey Dyer, Biology, Wheaton College).
“Two Workshops for Professors teaching Undergraduate Biology or Computer Science with
an interest in incorporating ‘Genomics’ (the analysis of DNA sequences) into their curricula.”

1997-1998: NSF DUE 9751660 (Co-PI with Tommy Ratliff, Math/CS, Wheaton College).
“Student-Managed Web Servers and the Delivery of Quantitative Results --
Redefining Computer Literacy for the 21st Century.”

Internal (most recent only)

Summer 2010: Mars Faculty/Student Research Grant. With Co-PIs Dyer (Biology) and Kahn (Statistics) and two student researchers to update a database and pipeline of software for genomics experiments.

Summer 2008: Mars Faculty/Student Research Grant. With Co-PIs Dyer (Biology) and Kahn (Statistics) and three student researchers to develop a new database and pipeline of software for genomics experiments.

Summer 2007: Mellon funding for summer research on “Computational Text Analysis of Old English Poetry”.

Publications

Research

- Drout, M.D.C., M. J. Kahn, M.D. LeBlanc, and Nelson, C. '11. Of Dendrogrammatology: Lexomic Methods for Analyzing the Relationships Among Old English Poems. *Journal of English and Germanic Philology*, July 2011, 301-336.
- Drout, M., Kahn, M., LeBlanc, M.D., Jones, A. '11, Kathok, N. '10, and Nelson, C. '11. Lexomics for Anglo-Saxon Literature. *Old English Newsletter*, 2010.
- Dyer, B.D., Kahn, M.J., and LeBlanc, M.D. (2007). Classification and regression tree (CART) analyses of genomic signatures reveal sets of tetramers that discriminate temperature optima of archaea and bacteria. *Archaea* 2:159–167.
- Russell, S.W. and LeBlanc, M.D. (2004). Learning By Seeing By Doing: Arithmetic Word Problems. *The Journal of the Learning Sciences*, v13 (2), 197-220.
- Dyer, B.D., LeBlanc, M.D., Benz, S. '05, Cahalan, P.'04, Donorfio, B.'04, Sagui, P. '04, Villa, A.'03, and Williams, G. '03 (2004). A DNA motif lexicon: cataloguing and annotating sequences. *In Silico Biology*, v4, 0039. 471-478.
- LeBlanc, M.D. (2004). favGene v2.0 -- A Tool for Extracting and Searching Upstream, Downstream, and Genic Regions of DNA for Your "Favorite Genes". Abstract published in the Proceedings of the International Conference on Bioinformatics, Auckland, New Zealand.
- LeBlanc, M.D., Baron, M., Christoforou, A. '01, Doolittle, N. '03, Kimball, M. '02, Villa, A. '03, Williams, G. '02, and Dyer, B. (2002). The DNA Motif Lexicon -- cataloguing and annotating genomes. Abstract published in the Proceedings of the 14th International Genome Sequencing and Analysis Conference (TIGR), p92.
- LeBlanc, M.D., Aspeslagh, G. '00, Buggia N. '01, and Dyer, B.D. (2000). An annotated catalogue of inverted repeats of *Caenorhabditis elegans* Chromosome III with observations concerning odd/even biases and conserved motifs, *Genome Research*, v10(9): 1381-1392.
- LeBlanc, M.D. and Weber-Russell, S. (1996). Text integration and mathematical connections: a computer model of arithmetic word problem solving. *Cognitive Science*, v20(3), 357-407, 1996.
- LeBlanc, M.D. (1992). From natural language to mathematical representations: a model of mathematical reading. In *Mathematical Intelligent Learning Environments*, Nwana, H.C. (Ed.), Intellect Press, 1992. Also appears in a Special Issue of *Intelligent Tutoring Media*, v2(3-4), 149-158, 1991.

Pedagogy

Book

- LeBlanc, M.D. and Dyer, B.D. (2007). [Perl for Exploring DNA](#). Oxford University Press.

Papers

- Maloney, M., Parker, J., LeBlanc, M.D., Woodard, C.T., Glackin, M., and Hanrahan, M. (2010). Bioinformatics and the Undergraduate Curriculum *CBE Life Sci Educ*: 172–174.
- LeBlanc, M.D., Gousie, M. and Armstrong, T. (March 2010). Connecting Across Campus. *Proceedings of the 41st SIGCSE Technical Symposium on Computer Science Education*, Milwaukee, WI.

Papers continued

- LeBlanc, M.D. and Leibowitz, R. (2006). Discrete Partnership – A case for a full-year of Discrete Math. Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education, Houston, TX, 313-317.
- LeBlanc, M.D. and Dyer, B.D. (2004). Bioinformatics and Computing Curricula 2001 -- Why Computer Science is well positioned in a post-genomic world. *ACM SIGCSE Bulletin*, v36 (4), Dec. 2004, 64-67.
- Benz, S.'04, Grossman, R.'07, Dyer, B.D., and LeBlanc, M.D. (2004). Genomics Research and the Liberal Arts: Building a Database for Exploring Your Favorite Set of Genes (favGene v2.0). *Transformations-Liberal Arts in the Digital Age*, v2 (1), May 2004.
- LeBlanc, M.D. (2004). Bioinformatics in the Undergraduate Curriculum: Opportunities for Computer Science Educators. Abstract appears in the *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education*. ACM Press, 229-230.
- LeBlanc, M.D. and Dyer, B.D. (2003). Teaching together: A three-year case study in genomics. *The Journal of Computing Sciences in Colleges*, v18 (5), 85-95. This paper won the Best Paper Award for 2003.
- Dyer, B.D. and LeBlanc, M.D. (2002). Meeting Report: Incorporating Genomics Research into Undergraduate Curricula. *Cell Biology Education*, Winter Issue, 2002, 101-104.
- LeBlanc, M.D. and Dyer, B. (2002). Collaborations in Genomics - Connecting Courses in Genetics and Computer Science. Published in the workshop proceedings of *BIO 2010: Undergraduate Biology Education to Prepare Research Scientists for the 21st Century*, American Society of Cell Biology, 29-44.
- LeBlanc, M.D. and Baron, G. (1999). Service learning in computing. *The Journal of Computing in Small Colleges*, v14 (4): 173-181. This paper won the Best Paper Award for 1999.
- LeBlanc, M.D. (1996). Breadth += depth; Augmenting breadth in CS1 and CS2 with in-depth projects published on the Web. *The Journal of Computing in Small Colleges*, v11 (4), 109-116, May 1996.

Conference and Seminar Presentations

- Computer Science @Home: Programming with Alice Animation. Presented at the Massachusetts Home School Convention. Worcester, MA, April 30, 2011.
- Bioinformatics in the computer science curriculum. Panel presentation at CCSCNE 2011 -- Conference for Computing in Small Colleges. Western New England College, Springfield, MA, April 16, 2011.
- Recruiting via a First-Year Seminar: Storytelling Through Computer Animation. Presented at SIGCSE 2011 -- The Technical Symposium on Computer Science Education, Dallas, TX, March 11, 2011.
- medPing: Data Structures with Embedded Medical Devices. Presented at CCSCNE 2010 – Conference for Computing in Small Colleges. Hartford, CT, April 16, 2010.
- Computer Science @Home: Programming with Alice Animation. Presented at the Massachusetts Home School Convention. Worcester, MA, April 24, 2010.
- Connecting Across Campus. Presented at SIGCSE 2010 – The Technical Symposium on Computer Science Education, Milwaukee, WI, March 11, 2010.
- Computing for Poets. Presented at SIGCSE 2010 – The Technical Symposium on Computer Science Education, Milwaukee, WI, March 12, 2010.

Presentations continued

Fishing for patterns in a sea of texts: Lexomics for Anglo-Saxon Literature. Presented at the International Society of Anglo-Saxonists (with M. Drout, M. Kahn, and C. Nelson '11) at St. John, Newfoundland, Canada, July 25, 2009.

Horizontal Transfer: Across Disciplines and Genomes. Presented at Rhode Island College Computer Science Seminars (with B. Dyer and M. Kahn), Providence, RI, March 19, 2009.

Bio Meets Big Oh: Turning Research into Programming Assignments. Poster presentation at SIGCSE 2008 – The Technical Symposium on Computer Science Education. Portland, OR, March 14, 2008.

Extremophiles and Spider Webs: Adventures in Genomics. Keynote address at the National Institute for Technology and Liberal Education (NITLE) Bioinformatics Practicum (with B. Dyer), Bates College, July 19, 2007.

If Darwin had Undergraduate Interns on the Beagle.... Keynote address at the National Institute for Technology and Liberal Education (NITLE) Bioinformatics Practicum (with B. Dyer), Bates College, June 19, 2006.

Pals in DNA Land. Presented at Bridgewater State College (with R. Leibowitz). Sponsored by the Department of Mathematics. February 27, 2006.

A Linguistic Analysis of Microbial Genomes. Presented at the University of New Hampshire (with B. Dyer). Sponsored by Department of Computer Science, Hubbard Genome Center, and the Department of Microbiology. February 23, 2006.

Poets, Delta, and DNA: inferring prokaryotic relationships from non-coding regions. Presented at the Colby College Computer Science Seminar Series, Waterville, Maine, September 15, 2005.

In silico research in Bioinformatics -- Why computing is so well positioned in a post-genomic world. Address at the School of Information Technology and Computer Science Seminar Series, University of Wollongong, NSW, Australia, Feb. 18, 2005.

favGene v2.0 -- A Tool for Extracting and Searching Upstream, Downstream, and Genic Regions of DNA for Your "Favorite Genes". Presented at the International Conference on Bioinformatics, Auckland, New Zealand, Sept. 6, 2004.

Course Materials and Teaching Models for Bioinformatics. Presented at the National Technical Symposium on Computer Science Education, Norfolk, Virginia, March 4, 2004.

Bioinformatics in the Undergraduate Curriculum: Opportunities for Computer Science Educators. Presented at the National Technical Symposium on Computer Science Education, Norfolk, Virginia, March 5, 2004.

Towards a DNA Dictionary. The 42nd Annual Meeting of the American Society for Cell Biology with Benz, S. '05, Cool, J. '04, and Dyer, B., San Francisco, CA, December 16, 2002.

The DNA Motif Lexicon -- cataloguing and annotating genomes. Presented at the 14th International Genome Sequencing and Analysis Conference (TIGR) with Baron, M. '03, Christoforou, A. '01, Doolittle, N. '03, Kimball, M. '02, Villa, A. '03, Williams, G. '02, and Dyer, B., October 4, 2002, Boston, MA.

Genomics: Biology Meets Computer Science. Presented at the Association of American Colleges & Universities (AACU) Conference on Rethinking Scientific Literacy in an Age of Diversity and Specialization. Charlestown, SC, April 2000.

Genomes and Computing -- From Homework to Research. Presented at MathFest'99 with Betsey Dyer, Providence, RI, July 1999.

Service Learning in Computing. Presented at the 1999 Northeastern Conference on Computing in Small Colleges, Providence, RI, April, 1999.

WebPals – a mathematics education tool, with Dana Breslau and Sylvia Weber-Russell. Presented at WebNet97 World Conference of the WWW, Internet, and Intranet. Association of Advancement of Computing in Education, Toronto, Canada, November, 1997.

Doing Their Own Math: Computer Support of Discursive Approaches to "Real" Math Problems, with Sylvia Weber-Russell. Presented at and published in the Proceedings of the Second International Conference on the Learning Sciences, Association of Advancement of Computing in Education, July, 1996, 324-331.

Making mathematical connections through natural language: a computer model of text comprehension in arithmetic word problem understanding, with Sylvia Weber-Russell. Presented at and published in the Proceedings of the 15th Annual Meeting of the Cognitive Science Society, Lawrence Erlbaum Associates, June, 1993, 641-646.

When more is less -- interactive tools for relational language. Presented at and published in the Proceedings of the 16th International Group of the Psychology of Mathematics Education, (abstract appears in volume 3, p. 171), Durham, NH, August, 1992.

Monitoring the role of linguistic processes in the comprehension of arithmetic word problems. Presented at and published in the Proceedings of the Third University of New Brunswick Artificial Intelligence Workshop, Fredericton, New Brunswick, Canada, October, 1990, 117-126.

Workshops

Perl for Biologists. Howard Hughes funded workshop for retaining students in science. Hampton College, VA, June 27-July 1, 2011.

Perl for Biologists. Howard Hughes funded workshop for retaining students in science. Hampton College, VA, June 27-30, 2010.

Palindromes in DNA Land. Presented at 'DNA and Health: Education, Community and Business Perspectives' (with B. Dyer) at the Center for the Advancement of Science Exploration, Bridgewater State College, February 5, 2009.

Alice in High School Land. Presented a hands-on introduction to the Alice programming environment to computing and mathematics teachers at Norton Public High School, Norton, MA, February 1, 2009.

Extremophiles and Spider Webs: Adventures in Genomics. Presented at 'Bringing Big Science to Small Schools: Genomics Curriculum Development Workshop' (with B. Dyer) at Vassar College, July 22, 2007.

Travels in "DNA Land – Approaching DNA sequence analysis with word play. Presented with Betsey Dyer (Biology, Wheaton College) at the 38th SIGCSE Technical Symposium on Computer Science Education, Covington, KY, March 9, 2007.

Regular Expressions and DNA. Presented with Betsey Dyer (Biology, Wheaton College) at the National Institute for Technology and Liberal Education (NITLE) Bioinformatics Practicum (with B. Dyer), Bates College, June 20, 2006.

Workshops continued

Travels in “DNA Land – Approaching DNA sequence analysis with word play. Presented with Betsey Dyer (Biology, Wheaton College) at the 37th SIGCSE Technical Symposium on Computer Science Education, Houston, TX, March 3, 2006.

Travels in "DNA Land" -- Approaching DNA Sequence Analysis Through Word Play. Presented with Betsey Dyer (Biology, Wheaton College) at the Consortium for Computing in Colleges Eastern Conference, Iona College, Iona, NY, Oct. 10, 2005.

Exploring DNA Land with Regular Expressions. Presented at Geospiza’s “Bio21: Teaching Biology with Bioinformatics” with Betsey Dyer, Research Triangle Park, NC, Oct. 17-18, 2003.

Moving Research to the Classroom: Linking courses in Biology and Computer Science. Presented at a two-day short course on “Reading the Book of Life: How Bioinformatics Makes Sense of Molecular Messages”, MathFest 2003, Boulder, CO, July 29-30, 2003.

Two Workshops for Professors teaching Undergraduate Biology or Computer Science with an interest in incorporating Genomics (the analysis of DNA sequences) into their curricula. (NSF DUE #0126643), Wheaton College, Norton, MA June 2002 and June 2003.

Genomics in the Undergraduate Curriculum. Presented at “Bioinformatics in the Undergraduate Curriculum”, with Betsey Dyer. Dickinson College, Carlisle PA, March 21-22, 2003.

Collaborations in Genomics - Connecting Courses in Genetics and Computer Science. Presented at the workshop “New Paradigms in Teaching Introductory and Cell Biology” at the 42nd Annual Meeting of the American Society for Cell Biology, with Betsey Dyer, San Francisco, CA, December 14, 2002.

Teaching

Wheaton College: Department of Mathematics and Computer Science. **July, 1993 - present.**

- Introduced two-hour, hands-on labs in introductory Computer Science courses (1993).
- Changed the introductory year in computer science from Pascal to C++ (1994).
- Upgraded the minor in Computer Studies to a minor in Computer Science (1995).
- Introduced a new major in Computer Science (1997).
- Included a service learning component into the computer science program; students work as one company for clients in the local town (e.g., Town Hall, Norton Historical Society).
- NSF funding for a new 20 workstation lab (Windows/Linux machines) (1998).
- Added a new “writing intensive” course: Software Engineering (Spring 2001).
- Helped justify and secure two tenure track positions for additional computer scientists, 1998 and 2001.
- Developed a new discrete mathematics course for computer science majors (Spring 2003).
- Altered the computer science major to include a full-year of discrete mathematics. (Fall 2003).
- Developed two new non-majors “connected” courses to satisfy the Quantitative Analysis requirement: “DNA” (Fall 2003, team-taught with a biologist) and “Computing for Poets” (Spring 2004).
- Established interdisciplinary “connections” with colleagues in the Humanities: ‘Algorithms’ and ‘DNA’ connected with Ethics (Philosophy); ‘Computing for Poets’ connected with ‘Anglo-Saxon Literature’ and ‘Tolkien’ (English).
- Led computer science faculty to develop new suite of “initial courses” to broaden interest (2008).

Wheaton course experience

- Developed new first-year seminar “Storytelling through Computer Animation” featuring the reading and analysis of graphic novels and hands-on experience with the Alice programming environment (2008).

- Taught each of the following core and advanced courses:

CS1 (Robots)	CS2 (Data Structures)
Algorithms	Object-Oriented Programming
Artificial Intelligence	Operating Systems
Computer Organization	Programming Languages
Java and Security	Software Engineering

University of Wollongong, Australia, NSW: Department of IT and Computer Science. 2004-2005.

- Visiting Fellow in Computer Science.
- Team-taught CSCI204/MCS9204 “Unix and the C Family”.

University of New Hampshire: Department of Computer Science. 1987-1993.

- Scientific Programming with Numerical Methods. Fall 1991 - Spring 1993.
 - developed new service courses for the entire College of Engineering.
 - one of the first instructors to introduce C as the initial language for engineering students.

Student Scholarship

Undergraduate Honors Theses

diviText: Visualizing Text Segmentation for Text Mining. Amos Jones. May 2011.

The Politics of Free: Open Source Software in Government. Brian Donorfio. May 2004.

Giving DNA a Trie. Adam Villa. May 2003.

Search Algorithms for Locating Potential Regulatory Motifs in the Promoters of the Krebs's Cycle Genes of *Caenorhabditis elegans*. Glen Aspeslagh, May, 2000.

Utilizing a Genetic Algorithm to Search the Structure-space of Artificial Neural Networks for Optimal Architectures. Ken Aspeslagh, May 2000.

Presentations and Publications

Jones, Amos '11. diviText: Graphical and Automated Text Segmentation for Text Mining. Poster presentation at the Northeastern Consortium for Computing Sciences in Colleges, Western New England College, April 15, 2011.

Nelson, Christina '11, Drout, M., Kahn, M., and LeBlanc, M.D. Fishing for patterns in a sea of texts: Lexomics for Anglo-Saxon Literature. Presented at the International Society of Anglo-Saxonists at St. John, Newfoundland, Canada, July 25, 2009.

Benz, Steve '05, Grossman, Robbie '07, Dyer, B.D., and LeBlanc, M.D. (2004). Genomics Research and the Liberal Arts: Building a Database for Exploring Your Favorite Set of Genes (favGene v2.0). *Transformations- Liberal Arts in the Digital Age*, v2 (1), May 2004.

Benz, Steve '05 and Grossman, Robbie '07 (2004). FavGene 2.0 -- A Perl and MySQL System for Exploring a Set of Genes. Presented at the Northeastern Conference on Computing in Small Colleges, Union College, NY, April 2004. Abstract published in the *Journal of Computing Sciences in Colleges*, 311-312.

Student Scholarship continued

Donorfio, Brian '04 (2004). The Politics of Free: Open Source Software in Government. Presented at the Northeastern Conference on Computing in Small Colleges, Union College, NY, April, 2004. Abstract published in the *Journal of Computing Sciences in Colleges*, 279-280.

Benz, Steve '05 and Cool, Jonah '04 (2003). Using Regular Expressions to Locate Putative Zinc Finger Binding Sites. Presented at the Northeastern Conference on Computing in Small Colleges, Rhode Island College, RI, April 2003. Abstract published in the *Journal of Computing Sciences in Colleges*, v18(5), 254.

Villa, A. '03 (2003). Searching DNA Neighborhoods. Presented at the Northeastern Conference on Computing in Small Colleges, Rhode Island College, RI, April 2003. Abstract published in the *Journal of Computing Sciences in Colleges*, v18 (5), 245.

Christoforou, Andrea '01 (2001). Counting Problems in Genomics. Presented at the Northeastern Conference on Computing in Small Colleges, Middlebury, VT, April 2001. Abstract published in the *Journal of Computing Sciences in Colleges*, v16 (4), 343.

Kimball, Melissa '02 (2001). A Motif Lexicon for the Genomic Analysis of DNA. Presented at the Northeastern Conference on Computing in Small Colleges, Middlebury, VT, April 2001. Abstract published in the *Journal of Computing Sciences in Colleges*, v16 (4), 344.

Kim, Ruben '99 (1999). A Computational Experiment in Number Theory. Presented at the Northeastern Conference on Computing in Small Colleges, Providence, RI, April, 1999. Abstract published in the *Journal of Computing Sciences in Colleges*.

Aspeslagh, Ken '00 and Pasquale, Matt '99 (1998). Cracking all night - Nocturnal execution and implementation of an encryption breaking application. Presented at the Northeastern Conference on Computing in Small Colleges, Sacred Heart University, Fairfield, CT, April, 1998. Abstract published in the *Journal of Computing Sciences in Colleges*.

Student Research Work**Summer 2011**

Amos Jones '11 and Donald Bass '12. NEH grant research students. New tools for text mining experiments.

Academic year 2010-2011

Emily Baldwin '14 (Wheaton Research Partner).
Learning Perl for experiments in genomics.

Summer 2010

Amos Jones '11, Brandon Waltz '11, Matt Brousseau '11 (NEH-funded), Nick Faulconer '12, Donald Bass '12 (Mars Student Fellows).

Updated software to automatically re-seed our database of 700+ microbial genomes and set up experiments with genomic signatures to detect horizontal transfer.

Updated content with WordPress and ported websites for both the Genomics and Lexomics research groups.

Fall 2009

Christina Nelson '11 (Wheaton Research Partner).

Experimental support for two new publications on untangling the web of the Anglo-Saxon corpus.

Academic year 2008 – 2009

Christina Nelson '11 and Neil Kathok '10 (Wheaton Research Partners).

Designed and ran hundreds of experimental runs as we untangle the web of the Anglo-Saxon corpus.

Student Research continued

Summer 2008

Neil Kathok '10, Christina Nelson '11, Amos Jones '11 (Mars Student Fellows).

Implemented software to automatically re-seed our database of 700+ microbial genomes.

Designed and implemented a suite of software tools for authorship attribution (genomic signature) techniques for both Anglo-Saxon and microbial genomes.

Academic year 2007 – 2008

Evan Ferri '08, Matthew Brown '10, and Neil Kathok '10 (Wheaton Research Partner).

Experimented with models to determine if genomic signatures in plasmids are similar to the host genome.

Spring 2007

Neil Kathok '10 (Wheaton Research Partner).

Tested all Perl examples found in our book 'Perl for Exploring DNA' (Oxford).

Summer 2006

Greg Williams '06 (NSF Grant Intern).

Designed and seeded a mysql database to hold 300 microbial genomes and associated annotation.

Summer 2005

Robert Goodman '07 (NSF Grant Intern).

FavGene v2.1 – upgrading software, adding genomes, and running experiments.

July 2004 – July 2005

Paul White '05 (Research Fellow, University of Wollongong, Australia, NSW).

Development of a suite of software for experiments on over 200 prokaryotic genomes.

Fall 2003 to Spring 2005

Steve Benz '05., Robert Goodman '07, Nguni Phakela '06 (Wheaton Research Partners).

FavGene v2.0 – an application to search upstream and downstream of a user's suite of "favorite genes".

Summer 2003

Brian Donorfio '04 and Pete Cahalane '04. (Mars Research Fellows).

Motif Lexicon v3.0 – a complete upgrade of the "DNA dictionary" software.

January 2003

Greg Williams '03. Perl utilities for a Favorite Gene Project.

Summer 2002

Martin Baron '03 and Patrick Sagui '04. (Mars Research Fellows).

Implementation of Relateds-module and porting Motif Lexicon v2.0 to new Linux/Win2K servers.

January 2002

Melissa Kimball '02 and Greg Williams '03 (Wheaton Research Partners)

Implementation of Statistics module and Etymology module for the Genomics Research Group website.

January 2001

Adam Villa '03, Nick Doolittle '03, Melissa Kimball '02 and Nathan Buggia '01.

Software team to implement what we believe is the first motif lexicon (DNA dictionary).

Summer 2000

Glen Aspeslagh '00 (Shouse Fellow) and Nathan Buggia '01 (Wheaton Fellowship)

Student Research continued

January 1999

Melissa Kimball '02 and Trevor Agnitti '02. (Mars Research Fellows)
Undergraduates in computer science prototype a motif lexicon (DNA dictionary).

January 1998

Ken Aspeslagh '00 (Shouse Fellow)
Designed and programmed the graphical user interface for *eXpress Math* v2.0
Prototyped by teachers and students in the Quincy Public School system.

Student Funding and Fellowships (most recent only):

Amos Jones '11 and Donald Bass '12. NEH grant text mining software development.

Donald Bass '12 and Nick Faulconer '12 – Mars Student Fellows.
Genomics (horizontal transfer) experiments.

Christina Nelson '11 – Wheaton Research Partner.
Lexomics (Old English) experiments.

Christina Nelson '11 and Neil Kathok '10 – Wheaton Research Partners.
Lexomics (Old English) and genomics (horizontal transfer) experiments.

Neil Kathok '10, Christina Nelson '11, Amos Jones '11 – Summer 2008 Mars Student Fellows.
NEH and *Wheaton Genomics Group* authorship attribution and genomics signature project.

Neil Kathok '10 - 2007 Wheaton Research Partner.
Wheaton Genomics Group textbook project

Steve Benz '05., Robert Goodman '07, Nguni Phakela '06 - 2003-2004 Wheaton Research Partners.
Wheaton Genomics Group FavGene v2.0

Brian Donorfio '04 and Pete Cahalan '04 - 2003-2004 Mars Research Fellows.
Wheaton Genomics Group Motif Lexicon v3.0

Stephen Benz '05 and Jonah Cool '04 - 2002-2003 Wheaton Research Partners.
Wheaton Genomics Group - Finding putative Zinc Finger regulatory sites.

Martin Baron '03 and Patrick Sagui '04 - Summer 2002 Mars Research Fellows.
Wheaton Genomics Group - Motif Lexicon v2.0.

Adam Villa '03, Nick Doolittle '03, Melissa Kimball '02 and Nathan Buggia '01 - Jan. 2001
Filene Research Fellows. *Wheaton Genomics Group* - software for motif lexicon.

Melissa Kimball '02 and Trevor Agnitti '02 - Jan. 1999 Mars Research Fellows.
Wheaton Genomics Group - prototype of a motif lexicon.

Nathan Buggia '01 - Jan. 1999 Wheaton Fellowship Program.
Wheaton Genomics Group - browse and search tools for DNA sequences

David Dudek '01 - Jan. 1999 Wheaton Fellowship Program.
Disseminating course materials via the web

Nathan Buggia '01- Jan. 1998 Wheaton Fellowship Program.
eXpress Math -- v2.1 Perl database

Ken Aspeslagh, '00 - January 1997 Wheaton Fellowship Program.
eXpress Math -- v2.0 graphical user interface revision

Student Funding and Fellowships (continued):

Tim McFadden, '97 and John Crowley, '99 -- Summer 1996 Gebbie Student Researchers.
eXpress Math v1.0

Marie Chapman, '96 – Spring 1996 Wheaton Foundation Grant.
Distributed copies of eXpress Math software and documentation

Tung Nguyen, '96 - January 1996 Wheaton Fellowship Program.
Object-oriented classes for event-driven programming

Tung Nguyen, '96 - January 1995 Wheaton Fellowship Program.
Inside Macintosh

Independent Studies

<i>Experiments in Text Mining (Spring 2011)</i>	Christina Nelson '11
<i>Mobile App Development (Fall 2011)</i> <i>(medPing v2.0)</i>	Ethan Levian '14, Nick Faulconer '12, Tyler Santos '12, Tim Gwynne '12, Raul Gil '12
<i>Topics in Bioinformatics (Fall 2009)</i>	Madeline Odigie '10
<i>Microbial Genomes (Fall 2009)</i>	Neil Kathok '10
<i>Programming Mobile Devices (Fall 2005)</i>	Nguni Phakela '06, Evan Fink '06, Sarah Milewski '07, Alex Cook '08
<i>Experiments in Putative Regulatory Sites (Spring 2004)</i>	Steve Benz '05
<i>Educational software for word problems (Fall 2002)</i>	Rauny Baez '03
<i>Systems Engineering in Genomics (Fall 2000)</i>	Nathan Buggia '01
<i>Artificial Intelligence (Spring 1999)</i>	Chris Zahka '99, Ruben Kim '99, Glen Aspeslagh '00
<i>Java and Security (Spring 1999)</i>	Aaron Willis '99, Travis Riley '99, David Sartory '00
<i>Neural Networks (Spring 1998)</i>	Ken Aspeslagh '00
<i>Genomics (Spring 1998)</i>	Glen Aspeslagh '00
<i>Advanced Object-Oriented Programming (Fall 1998)</i>	Ken Aspeslagh '00
<i>Java and Security (Fall 1997)</i>	Kim McClure '98, Todd Oldenburg '98, Ruben Kim, '99, Nathan Buggia, '01
<i>Network Administration (Fall 1998)</i>	Matt Pasquale '99
<i>eXpress Math (Spring 1996)</i>	Tung Nguyen, '96, David McClintock, '95, Andrew Ullmann '95, Marie Chapman, '96
<i>Apple Scripting (Fall 1996)</i>	Tung Nguyen, '96