

J. Gregory Caporaso
gregcaporaso@gmail.com

EDUCATION

Ph.D., May 2009

Dissertation title: *Extracting signal from noise in biological data: Evaluations and applications of text mining and sequence coevolution*

Mentor: Lawrence Hunter, Ph.D.

Program in Biophysics and Structural Biology

Department of Biochemistry and Molecular Genetics

University of Colorado Denver, Aurora, CO

(Formerly U of C Health Sciences Center.)

Bachelor of Arts in Biochemistry, Minor in Chemistry, August 2004

University of Colorado at Boulder, College of Arts and Sciences, Boulder, CO

Bachelor of Science in Computer Science, May 2001

University of Colorado at Boulder, College of Engineering and Applied Science, Boulder, CO

AWARDS AND HONORS

NLM Informatics Pre-Doctoral Training Fellowship

July 2007 – July 2009

Provided full funding, including tuition and stipend, for my pre-doctoral training. Grant number: T15LM009451.

Outstanding Dissertation Award Finalist

University of Colorado Denver, May 2009.

Outstanding Research Award, 23rd Annual Student Research Forum, University of Colorado Denver
Sequence co-occurrence and covariation suggest a model of the Type VI Secretion System, January, 2009.

NSF Travel Fellowship, December, 2008

Provided registration and travel costs to attend the 2008 Rocky Mountain Bioinformatics Conference.

Outstanding Research Award, 22nd Annual Student Research Forum, University of Colorado Denver
A Comparison of Coevolution Detection Methods on Protein Alpha Helices, January, 2008.

Best Presentation, Rocky Mountain Bioinformatics Conference

A Sentence Recognizer for Mutant Protein Structure Studies: Toward Intelligent Systems for the Management of Structural Biology Data, December, 2005.

CURRENT POSITIONS

Research assistant, February 2005 – Present.

Lawrence Hunter Laboratory, Computational Bioscience Program, Department of Pharmacology,

University of Colorado Denver, Aurora, CO.

PREVIOUS POSITIONS

Research assistant (as rotation student), December 2004 – February 2005.

Mark Duncan Laboratory, Biomolecular Structure Program, University of Colorado Health Sciences Center, Aurora, CO.

Research assistant (as rotation student), August 2004 – November 2004.

Robert Garcea Laboratory, Biomolecular Structure Program, University of Colorado Health Sciences Center, Aurora, CO.

Research assistant, January 2003 – August 2004.

Michael Yarus and Rob Knight Laboratories, Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, Boulder, CO.

Contract Web Developer, 1999 – 2002.

American Business English Internet School, Boulder, CO.

Software Engineer, 2001 – 2002.

Freshwater Software (acquired by Mercury Interactive), Boulder, CO.

Student Intern, 2000 – 2001.

Freshwater Software, Boulder, CO.

Web Developer, System Administrator, 1998 – 2000.

Web Communications, University of Colorado at Boulder, Boulder, CO.

TEACHING EXPERIENCE

High School Mathematics Tutor, December 2008 – Present.

Boulder, CO.

Mentor to Junior Graduate Student, February 2008 – December 2008.

Computational Bioscience Program, University of Colorado Denver, Aurora, CO.

Guest Lecturer, Spring semester, 2007.

Bioinformatics 7712, University of Colorado Health Sciences Center, Aurora, CO.

Mentor to Undergraduate Summer Student, Summer, 2006.

Lawrence Hunter Laboratory, University of Colorado Health Sciences Center, Aurora, CO.

Curriculum Developer, 2003 – 2004.

Department of Human Resources, University of Colorado at Boulder, Boulder, CO.

INVITED PRESENTATIONS

Intermolecular Coevolution Suggests Interactions Between Type VI Secretion System Proteins.

Bioinformatics Supergroup, April, 2009, Boulder, CO.

Rapid Pattern Development for Concept Recognition Systems.
DMAP meeting, April, 2008, Aurora, CO.

Detecting Coevolution by Disregarding Evolution?
Bioinformatics Supergroup, April, 2008, Boulder, CO.

PEER-REVIEWED PUBLICATIONS

1. Detecting Coevolution by Disregarding Evolution? Tree-Ignorant Metrics of Coevolution Perform As Well As Tree-Aware Metrics; **J. Gregory Caporaso**, Sandra Smit, Brett C. Easton, Lawrence Hunter, Gavin A. Huttley, and Rob Knight. *BMC Evolutionary Biology* 2008, 8:327.
2. An Integrated Approach to Concept Recognition in Biomedical Text; William A. Baumgartner, Jr., Zhiyong Lu, Helen L. Johnson, **J. Gregory Caporaso**, Jesse Paquette, Anna Lindemann, Elizabeth K. White, Olga Medvedeva, K. Bretonnel Cohen, Lawrence Hunter. *Genome Biology*, 2008, Suppl 2:S9.
3. Intrinsic Evaluation of Text Mining Tools May Not Predict Performance on Realistic Tasks; **J. Gregory Caporaso**, Nita Deshpande, J. Lynn Fink, Philip E. Bourne, K. Bretonnel Cohen, and Lawrence Hunter; *Pacific Symposium on Biocomputing* 13:640-651(2008).
4. Rapid Pattern Development for Concept Recognition Systems: Application to Point Mutations; **J. Gregory Caporaso**, William A. Baumgartner, Jr., David A. Randolph, K. Bretonnel Cohen, and Lawrence Hunter; *Journal of Bioinformatics and Computational Biology*, 2007 Dec;5(6):1233-59.
5. PyCogent: A Toolkit for Making Sense from Sequence; Rob Knight, Peter Maxwell, Amanda Birmingham, Jason Carnes, **J. Gregory Caporaso**, Brett C. Easton, Micah Hamady, Zongzhi Liu, Catherine Lozupone, Raymond Sammut, Sandra Smit, Matthew Wakefield, Jeremy Widmann, Shandy Wikman, Stephanie Wilson, and Gavin A. Huttley; *Genome Biology* 2007, 8:R171.
6. MutationFinder: A High-Performance System for Extracting Point Mutation Mentions from Text; **J. Gregory Caporaso**, William A. Baumgartner Jr., David A. Randolph, K. Bretonnel Cohen, and Lawrence Hunter; *Bioinformatics*, 2007 23(14):1862-1865.
7. Error Minimization and Coding Triplet/Binding Site Associations are Independent Features of the Canonical Genetic Code; **J. Gregory Caporaso**, Michael Yarus, and Rob Knight; *Journal of Molecular Evolution*, 2005.
8. Origins of the Genetic Code: The Escaped Triplet Theory; Michael Yarus, **J. Gregory Caporaso**, and Rob Knight; *Annual Review of Biochemistry*, 2005.

INVITED PUBLICATIONS

9. Concept Recognition, Information Retrieval, and Machine Learning in Genomics Question Answering; **J. Gregory Caporaso**, William A. Baumgartner Jr., Hyun-min Kim, Zhiyong Lu, Helen L. Johnson, Olga Medvedeva, Anna Lindemann, Lynne M. Fox, Elizabeth K. White, K. Bretonnel Cohen, and Lawrence Hunter; *TREC 2006 Conference Proceedings*, January, 2007.
10. Concept recognition and the TREC Genomics tasks; **J. Gregory Caporaso**, William A. Baumgartner Jr., K. Bretonnel Cohen, Helen L. Johnson, Jesse Paquette, Lawrence Hunter; *TREC 2005 Conference Proceedings*, January, 2006.

CONFERENCE PRESENTATIONS

Intermolecular Coevolution Suggests Interactions Between Type VI Secretion System Proteins.
Rocky Mountain Bioinformatics Conference (oral and poster presentation), December, 2008.

Intrinsic Evaluation of Text Mining Tools May Not Predict Performance on Realistic Tasks.
Pacific Symposium on Biocomputing (oral presentation), January 2008.

A Comparison of Coevolution Detection Methods on Protein Alpha Helices.
Pacific Symposium on Biocomputing (poster presentation), January, 2008.
Rocky Mountain Bioinformatics Conference (oral and poster presentation), December, 2007.

Recognizing Point Mutations in Text: Techniques and Applications in Biomedical Text Mining.
Pacific Symposium on Biocomputing (poster presentation), January, 2007.

Concept Recognition, Information Retrieval, and Machine Learning in Genomics Question Answering.

Rocky Mountain Bioinformatics Conference (oral and poster presentation), December, 2006.
Text REtrieval Conference (TREC) 2006 (poster presentation), November, 2006.

A Sentence Recognizer for Mutant Protein Structure Studies: Toward Intelligent Systems for the Management of Structural Biology Data.

Rocky Mountain Bioinformatics Conference (oral and poster presentation), December, 2005.

Concept Recognition and the TREC Genomics Tasks.

Text REtrieval Conference (TREC) 2005 (poster presentation), November, 2005.

Stereochemical and Adaptive Factors in Genetic Code Evolution.

Evolution 2004 (poster presentation), July, 2004.

OPEN-SOURCE SOFTWARE PROJECTS

Primary developer, MutationFinder
<http://mutationfinder.sourceforge.net/>

Developer, PyCogent (Python COmparative GENomics Toolkit)
<http://pycogent.sourceforge.net/>

PROFESSIONAL SOCIETY MEMBERSHIPS

International Society for Computational Biology, since 2005.

American Association for the Advancement of Science, since 2006.

Colorado Evolutionary Response Team, since 2006.

Committee for Skeptical Inquiry, since 2008.

REFERENCES

Lawrence Hunter, Associate Professor
Computational Bioscience Program, University of Colorado Denver.
larry.hunter@ucdenver.edu; (303) 724-3574

Rob Knight, Assistant Professor
Department of Chemistry and Biochemistry, University of Colorado at Boulder.
rob.knight@colorado.edu; (303) 492-1984

Mair Churchill, Associate Professor
Department of Pharmacology, University of Colorado Denver.
mair.churchill@ucdenver.edu; (303) 724-3670