

# Nathan Edwards, Ph.D.

---

Department of Biochemistry and Molecular & Cellular Biology  
Georgetown University Medical Center  
(202)-687-7042    nje5@georgetown.edu

---

## EDUCATION

**Cornell University**, Ithaca, NY

**Ph.D. in Operations Research** January 2001

**M.S. in Operations Research** January 1998

**Concentration:** Mathematical Programming

**Minors:** Applied Probability and Statistics, Computer Science (Scientific Computing)

**Ph.D. Thesis:** Improved approximation algorithms for the  $k$ -level facility location problem.

**Adviser:** Dr. David Shmoys

**Minor Advisers:** Dr. Sidney Resnick, Dr. Thomas Coleman

**University of Western Australia**, Perth, Australia

**Bachelor of Science with Honors** March 1992

**Majors:** Mathematics and Computer Science

**Thesis:** Synthesizing optimal feedback control with neural networks.

**Awards:** University of Western Australia Hackett Postgraduate Studentship.

## GRANTS

**Tandem Mass Spectral Biomarkers for Plant Pathogens**

PI: **Edwards**

January 2007 — January 2008

USDA-ARS

\$46,524

**Proteomic characterization of alternate splicing and cSNP protein isoforms**

PI: **Edwards**

September 2006 — August 2009

NIH-NCI (HHS): 1R01CA126189-01

\$835,442

**Methods for rapid detection of genetic engineering in microorganisms**

PI: **Fenselau/Edwards**

February 2006 — November 2006

DOD-DARPA-SPO: HR001106C0038

\$194,679

## PUBLICATIONS (Peer Reviewed)

1. S. Swatkoski, P. Gutierrez, C. Wynne, A. Petrov, J.D. Dinman, **N. Edwards**, and C. Fenselau. Microwave accelerated residue-specific acid cleavage for proteomic applications. *Journal of Proteome Research*. 7(2):579-586, 2008
2. S. Swatkoski, P. Gutierrez, J. Ginter, A. Petrov, J.D. Dinman, **N. Edwards**, and C. Fenselau. Integration of residue-specific acid cleavage into proteomic workflows. *Journal of Proteome Research*. 6(11):4525-4527, 2007.
3. J. Wang, P. Gutierrez, **N. Edwards**, and C. Fenselau. Integration of  $^{18}\text{O}$  Labeling and Solution Isoelectric Focusing in a Shotgun Analysis of Mitochondrial Proteins. *Journal of Proteome Research*. 6(12):4601-4607, 2007

4. X. Wu, C.-W. Tseng, and **N. Edwards**. HMMatch: Peptide Identification by Spectral Matching of Tandem Mass Spectra using Hidden Markov Models. *Journal of Computational Biology*. 14(8):1025-1043, 2007.
5. J. Xu, O. Laine, M. Masciocchi, J. Manoranjan, J. Smith, S. Du, J. Shao, **N. Edwards**, X. Zhu, C. Fenselau, and L.-Y. Gao. A unique *Mycobacterium* ESX-1 protein co-secretes with CFP-10/ESAT-6 and is necessary for inhibiting phagosome maturation. *Molecular Microbiology*. 66(3):787-800, 2007.
6. C. Fenselau, C. Havey, N. Teerakulkittipong, S. Swatkoski, O. Laine, and **N. Edwards**. Identification of  $\beta$ -Lactamase in Antibiotic-resistant *Bacillus cereus* spores. *Applied and Environmental Microbiology*. 74(3):904-906, 2007
7. S. Russell, **N. Edwards**, and C. Fenselau. Detection of Plasmid Insertion in *Escherichia coli* by MALDI-TOF Mass Spectrometry. *Analytical Chemistry* 79(14):5399-5406, 2007.
8. **N.J. Edwards**. Novel Peptide Identification using Expressed Sequence Tags and Sequence Database Compression. *Molecular Systems Biology*. 3(102), 2007.
9. D. Fasulo, A.-K. Emde, L.-Y. Wang, K. Noy, and **N.J. Edwards**. Alignment of Mass Spectrometry Data by Clique Finding and Optimization. *Proceedings of the RECOMB Satellite Workshop on Computational Proteomics: Systems Biology and Computational Proteomics*, Lecture Notes in Computer Science 4532:119-129, 2007.
10. S. Swatkoski, S. Russell, **N. Edwards**, and C. Fenselau. Analysis of a Model Virus Using Residue-Specific Chemical Cleavage and MALDI-TOF Mass Spectrometry. *Analytical Chemistry*, 79(2):654-658, 2007.
11. C. Fenselau, S. Russell, S. Swatkoski, and **N. Edwards**. Proteomic Strategies for Rapid Analysis of Microorganisms. *European Journal of Mass Spectrometry*. 13(1):35-39, 2007.
12. S. Swatkoski, S.C. Russell, **N. Edwards** and C. Fenselau. Rapid chemical digestion of small acid-soluble spore proteins for analysis of *Bacillus* spores. *Analytical Chemistry* 78(1):181-188, 2006.
13. **N. Edwards** and R. Lippert. Sequence database compression for peptide identification from tandem mass spectra. *4th Workshop on Algorithms in Bioinformatics (WABI)*, Lecture Notes in Computer Science 3240:230-241, 2004.
14. S. Istrail, G.G. Sutton, L. Florea, A.L. Halpern, C.M. Mobarry, R. Lippert, B. Walenz, H. Shatkay, I. Dew, J.R. Miller, M.J. Flanigan, **N.J. Edwards**, R. Bolanos, D. Fasulo, B.V. Halldorsson, S. Hannenhalli, R. Turner, S. Yooseph, F. Lu, D.R. Nusskern, B.C. Shue, X.H. Zheng, F. Zhong, A.L. Delcher, D.H. Huson, S.A. Kravitz, L. Mouchard, K. Reinert, K.A. Remington, A.G. Clark, M.S. Waterman, E.E. Eichler, M.D. Adams, M.W. Hunkapiller, E.W. Myers, and J.C. Venter. Whole-genome shotgun assembly and comparison of human genome assemblies. *Proceedings of the National Academy of Sciences* 101(7):1916-1921, 2004.
15. B.V. Halldorsson, V. Bafna, **N. Edwards**, R. Lippert, S. Yooseph, and S. Istrail. A survey of computational methods for determining haplotypes. *Computational Methods for SNPs and Haplotype Inference: DIMACS/RECOMB Satellite Workshop*, Lecture Notes in Computer Science 2983:26-47, 2004.

16. V. Bafna and **N. Edwards**. On *de novo* interpretation of peptide spectra. *7th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, pages 9-18, 2003.
17. **N.J. Edwards** and R. Lippert. Generating peptide candidates from amino-acid sequence databases for protein identification via mass spectrometry. *2nd Workshop on Algorithms in Bioinformatics (WABI)*, Lecture Notes in Computer Science 2452:68-81, 2002.
18. V. Bafna and **N.J. Edwards**. SCOPE: a probabilistic model for scoring tandem mass spectra against a peptide database. *9th International Conference on Intelligent Systems for Molecular Biology (ISMB)*, Bioinformatics 17(S1):S13-S21, 2001.
19. C.J. Goh and **N.J. Edwards**. Approximate output feedback optimal control of higher-order dynamical systems. *Optimal Control Applications and Methods*, 18(2):123-137, 1997.
20. **N.J. Edwards** and C.J. Goh. A direct method for continuous-time nonlinear optimal feedback controller. *Journal of Optimization Theory and Applications*, 84(3):509-528, 1995.
21. C.J. Goh and **N.J. Edwards**. Synthesis of optimal feedback controller by neural networks. *International Journal of Systems Science*, 25(8):1235-1248, 1994.
22. C.J. Goh, **N.J. Edwards** and A.Y. Zomaya. Feedback control of minimum-time optimal control problems using neural networks. *Optimal Control Applications and Methods*, 14(1):1-16, 1993.

#### **PUBLICATIONS (Other)**

1. X. Wu, **N. Edwards**, C.-W. Tseng. Peptide Identification via Tandem Mass Spectrometry. Chapter 6 of *Advances in Computers*, volume 68, ed. C.-W. Tseng and M. Zelkowitz, pages 253-278, 2006.
2. C. Fenselau, **N. Edwards**, and S. Russell. Strategy for Rapid Recognition of Bioengineered Bacteria. *US Patent Application US60/862,105*.
3. B.V. Halldorsson, V. Bafna, **N. Edwards**, R. Lippert, S. Yooseph, and S. Istrail. Combinatorial problems arising in SNP and haplotype analysis. *4th International Conference on Discrete Mathematics and Theoretical Computer Science*, Lecture Notes in Computer Science 2731:26-47, 2003.
4. R.J. Turner, K. Chaturvedi, **N.J. Edwards**, D. Fasulo, A.L. Halpern, D.H. Huson, O. Kohlbacher, J.R. Miller, K. Reinert, K.A. Remington, R. Schwartz, B. Walenz, S. Yooseph, and S. Istrail. Visualization challenges for a new cyberpharmaceutical computing paradigm. *Proceedings of the IEEE 2001 Symposium on Parallel and Large-Data Visualization and Graphics*, pages 7-18, 2001.

#### **ACCEPTED TALKS (Peer Reviewed)**

1. Improving the Reliability of Peptide Identifications using Machine Learning. *US Human Proteome Organization Annual Meeting*. (2008)
2. Novel Peptide Identification using Expressed Sequence Tags and Sequence Database Compression. *RECOMB Satellite Workshop on Computational Proteomics* (2006)
3. Novel Peptide Identification using ESTs and Genomic Sequence. *9th Annual Conference on Computational Genomics* (2006)

## INVITED TALKS

1. Improving the Reliability of Peptide Identification by Tandem Mass Spectrometry for Clinical Proteomics and Genome Annotation. *CCMB Lecture Series, Center for Computational Molecular Biology, Brown University* (2008)
2. Proteomic Characterization of Alternative Splicing and Coding Polymorphism. *Informatics Seminar, Department of Biomedical Informatics, Vanderbilt University School of Medicine* (2007)
3. Direct Experimental Observation of Functional Protein Isoforms by Tandem Mass Spectrometry. *Department of Biochemistry and Molecular and Cellular Biology, Georgetown University Medical Center* (2007)
4. Improving Genome Annotation using Proteomics. *J. Craig Venter Institute* (2007)
5. Proteomic Characterization of Alternative Splicing and Coding Polymorphism. *Department of Biochemistry and Molecular and Cellular Biology, Georgetown University Medical Center* (2007)
6. Optimal k-mer Superstrings for Protein Identification and DNA Assay Design. *Department of Computer Science, George Mason University* (2007)
7. Large Scale Combinatorial Optimization Problems in Bioinformatics. *Industrial and Systems Engineering, Lehigh University* (2007)
8. Proteomic Characterization of Alternative Splicing and Coding Polymorphism. *Department of Bioinformatics, George Mason University* (2007)
9. Novel Peptide Identification using ESTs and Genomic Sequence. *J. Craig Venter Institute* (2007)
10. Statistical Significance for Peptide Identification from Tandem Mass Spectra. *Statistical Analysis of Metabonomic and Proteomic Data, MolPAGE Training Workpackage* (2007)
11. Novel Peptide Identification using ESTs and Genomic Sequence. *International BCB-Workshop of Computational Mass Spectrometry* (2006)
12. Faster, more sensitive peptide identification by sequence database compression. *NIH Proteomics and Mass Spectrometry Interest Groups Seminar* (2005)
13. Optimal k-mer superstrings for peptide identification from tandem mass spectra. *Capital Area Theory Seminar (CATS), University of Maryland, College Park*. (2005)
14. Faster, more sensitive peptide identification by sequence database compression. *Workshop on Biomarker Discovery. Mathematical Biosciences Institute, Ohio State University* (2005)
15. Bioinformatics challenges for protein identification from tandem mass spectra. *Colloquium, Department of Biochemistry, Georgetown University Medical Center*. (2004)
16. Generating peptide candidates from protein sequence databases for protein identification via mass spectrometry. *Proteomics: Sequence, Structure, Function. Institute for Pure and Applied Mathematics, UCLA*. (2004)
17. Adventures in bioinformatics and computational biology. *Colloquium, School of Operations Research and Industrial Engineering. Cornell University*. (2004)

## POSTERS

1. X. Wu, **N.J. Edwards**, and C.-W. Tseng. Peptide Identification by Spectral Matching of Tandem Mass Spectra Using Hidden Markov Models. *RECOMB Satellite Workshop on Computational Proteomics*, 2006.
2. X. Wu, **N. Edwards**, and C.-W. Tseng. Experimental Comparison of Peptide Identification Algorithms, *9th Annual Conference on Computational Genomics*, 2006.
3. **N.J. Edwards** and F. Pineda. Rapid Microorganism Identification Database ([www.RMIDb.org](http://www.RMIDb.org)). *54th American Society of Mass Spectrometry Conference (ASMS)*, 2006.
4. S. Swatkoski, S. Russell, **N. Edwards**, and C. Fenselau. A new proteomics strategy for the identification of viruses. *54th American Society for Mass Spectrometry Conference (ASMS)*, 2006.
5. **N.J. Edwards**, X. Wu, and C.-W. Tseng. Novel peptide identification using ESTs and genomic sequence. *2nd Annual Symposium, US Human Proteome Organization (USHUPO)*, 2006.
6. E. Patton, **N. Edwards**, B. Oktem, and C. Fenselau. A microwave and detergent procedure to detect high molecular mass proteins from vegetative bacteria by MALDI-TOF MS. *American Chemical Society National Meeting*, 2005.
7. **N.J. Edwards**. Faster, more sensitive peptide identification from tandem mass spectra by sequence database compression. *1st Annual Symposium, US Human Proteome Organization (USHUPO)*, 2005.
8. I.P. Smirnov, K.C. Parker, D. Patterson, A.J. Tomlinson, P.S. Juhasz, T. Taylor, **N. Edwards**, D. Fasulo, A. Graber, S.A. Martin, J. Ramirez, T. Colpitts, S.D. Stroupe. Development of MALDI based human serum biomarker workflows. *52nd American Society of Mass Spectrometry Conference (ASMS)*, 2004.
9. R. Nutter, L. Johnston-Dow, J. Sorenson, H.C. Gire, S. Glanowski, C. Fosler, J. Hoglund, M.A. Rydland, P.-H. Shen, K. Glasser, C. Forbes, I. Singh, V. Sathineni, L. Wu, B. Small, P. Dunn, **N. Edwards**, C.-C. Chang, J. Ni, B. Jones, K. Woodford, P. Vatta, R. Loboda, G. Wang, B. Murphy, P. St. John, S. Ferreira, A. Parker, E. Gonzalez, C. Kosman, Q. Doan, L.Z. Pham, O. Bell. A resequencing primer set for 3,000 genes implicated in cancer genetics. *European Society of Human Genetics (ESHG)*, 2003.
10. J. Sorenson, S. Glanowski, **N. Edwards**, C. Fosler, A. Parker, L. Johnston-Dow, M. Furtado, C. Gire, L.-Z. Pham. Development and validation of a predictive model for robust primer design. *American Society of Human Genetics (ASHG)*, 2003.
11. R. Nutter, J. Sorenson, C. Gire, P. Shen, M.A. Rydland, S. Glanowski, **N. Edwards**, M. Furtado, R. Fang and L.-Z. Pham. Development and performance of the VariantSeqr resequencing system in high throughput DNA sequence variation studies. *Genome Sequencing and Analysis Conference (GSAC)*, 2003.

## TEACHING

- **University of Maryland, CMSC858E (Fall 2005)**  
**Algorithms in Biosequence Analysis.** Graduate level CS-Theory core course with an emphasis on sequence matching and indexing, hidden Markov models for gene finding and protein families, and haplotype phasing.

- **University of Maryland, BCHM676 & BIOE689B (Spring 2005, 2006, 2007) Biological Mass Spectrometry.** Gradual level elective, co-taught with Dr. Catherine Fenselau (Biochemistry), on mass spectrometry and associated analytical chemistry techniques applied to biological samples, and bioinformatic analysis of mass spectrometry experiments.
- **USHUPO Short-Course (2005, 2006) Bioinformatics for Proteomics.** Short-course at the USHUPO annual symposium, co-directed with Dr. Akilesh Pandey, on the bioinformatics aspects of protein identification by tandem mass spectrometry.
- **Teaching Assistant, OR&IE115, Cornell University (Fall 1995) Introduction to Operations Research.** Freshman introduction to graph algorithms and linear programming, taught by Dr. David Shmoys.
- **Teaching Assistant, OR&IE580, Cornell University (Fall 1996) Design and Analysis of Simulated Systems.** Graduate level course in discrete event simulation, taught by Dr. Lee Schruben.

## STUDENTS

- **Xue Wu**, Computer Science Ph.D. student (Dr. Chau-Wen Tseng), University of Maryland, College Park.  
Direction of thesis research in peptide identification by tandem mass spectrometry.
- **Ping Chen**, Computer Science Ph.D. student, University of Maryland, College Park.  
Supervision of thesis research in string matching using gapped seeds.
- **Elena Zotenko**, Computer Science Ph.D. student (Dr. Diane O'Leary), University of Maryland, College Park.  
Proposal committee.
- **Eli Moore**, Environmental chemistry Ph.D. student (Dr. Rodger Harvey), Chesapeake Biological Laboratory, University of Maryland.  
Thesis committee.
- **Ming-Yih Lai**, Biochemistry Ph.D. student, University of Maryland, College Park.  
Supervision of work for Fenselau laboratory rotation.

## GUEST LECTURES

- University of Maryland, College Park, Department of Computer Science. Advanced Topics in Information Processing: Exploiting Biological Resources (Dr. L. Raschid, 2006).
- George Washington University. Introduction to Bioinformatics (Dr L. Florea, 2005).
- Georgetown University. Advanced Bioinformatics (Dr C. Wu, 2005).
- George Washington University. Medical Bioinformatics (Dr S. Istrail, 2004).
- John Hopkins Bloomberg School of Public Health, Department of Molecular Microbiology and Immunology. Special topics in Protein Bioinformatics (Dr F. Pineda, 2004).
- George Washington University. Fundamentals of Proteomics (Dr F. Kashanchi, 2004).

## PROFESSIONAL EXPERIENCE

### University of Maryland, Center for Bioinformatics & Computational Biology,

College Park, MD

#### Assistant Research Scientist

October 2004—Present

- Sequence database compression for protein identification via mass spectrometry.
- Sequence databases for protein identification from genomic data.
- Bioinformatics of rapid microorganism identification using mass spectrometry.
- Algorithms (clique finding, isotope cluster deconvolution) for peptide quantitation via mass spectrometry.
- Spectral matching using HMMs for peptide identification via mass spectrometry.
- Applications of exact and inexact mer-counting (sequence database compression, species signatures, primer/probe design) using sequencing-by-hybridization (de Bruijn) graphs, Eulerian path sets, and uniqueness oracles.
- Lossless gapped seed-sets for linear expected time inexact string matching.

### Applied Biosystems, Informatics Research, Rockville, MD

#### Senior Staff Scientist, Bioinformatics

July 2003—October 2004

#### Staff Scientist, Bioinformatics

July 2002—July 2003

- Managed Informatics Research proteomics projects, supervising 1-3 people.
- Co-Developed *Biomarker Toolbox* for the statistical analysis of protein profiling mass spectra.
- Developed novel protein sequence databases that integrate coding SNPs, alternate splice forms, and known protein isoforms. (Celera/AB DNA Award).
- Contributed to the design and specification of protein sequence databases supplied with AB's instrument software. (Celera/AB DNA Award).
- Developed tool to verify large PCR primer sets uniqueness properties with respect to the human genome. Used in *Diomedes* production pipeline. (Celera/AB DNA Award).
- Developed bagging, cross validation, decision tree infrastructure for resequencing PCR primer design success prediction. Used for *VariantSEQr* confidence values.

### Celera Genomics, Informatics Research, Rockville, MD

#### Computer Scientist

September 2000—June 2002

- Developed novel models and algorithms for high-throughput identification of peptides via tandem mass spectrometry for a wide variety of mass spectrometry platforms, technologies, and experimental protocols.
- Implemented the SCOPE peptide identification sequence database search engine and web-based user interface for unsupervised use by Celera's mass spectrometry scientists. (Celera/AB DNA Award).
- Oversaw the development of Celera's high-throughput proteomics analysis pipeline, associated object model and relational database data model.

- Oversaw the development of a three-tiered application for browsing and manual curation of peptide identification results from Celera's proteomics analysis pipeline.
- Supervised the deployment of spectral matching for redundant spectrum elimination and peptide identification in Celera's proteomics analysis pipeline.
- Designed and supervised the deployment of the automated association of mass spectrometry experiments for quantization and identification of peptides in Celera's proteomics analysis pipeline.

**Cornell University, Field of Operations Research**, Ithaca, NY

**Research Assistant**

August 1999—August 2000, Summers 1995—1999

- Studied approximation algorithms for the  $k$ -level facility location problem.
- Studied capacitated fixed charge network design problems.
- Studied path coloring and routing in wavelength division multiplexing optical networks.
- Studied the fair scheduling of high bandwidth multicasts on trees.
- Ported parallel Fortran code for large scale discrete optimal control problems from the Intel iPSC/860 to the Cornell Theory Center's IBM SP/2. Analyzed algorithm for bottlenecks and data contention to optimize performance.

## **PROFESSIONAL SERVICE**

### **Scientific Advisory Councils**

- Australian Proteomics Computational Facility (2005–Present)

### **Program Committees**

- US HUPO (2008)
- RECOMB Satellite Workshop on Computational Proteomics (2006)
- RECOMB Satellite Workshop on Systems Biology (2006)
- International Workshop on High Performance Computational Biology (2006)

### **Grant Review Panels**

- NIH: Small Business: Computational Biology and Software Development (SRA: Malgorzata Klocek) (2008)
- NIH: Genes, Genomes, and Genetics Study Section (SRA: Michael Marino) (2006)
- NIH: Biodata Management and Analysis Study Section (SRA: Marc Rigas) (2006)
- DOE: Genomes to Life (PA: John Houghton) (2005)

### **Reviewer**

- Analytical Chemistry
- Algorithms for Molecular Biology
- Bioinformatics
- IEEE Journal of Selected Topics in Signal Processing

- IEEE Transactions on Parallel and Distributed Systems
- IEEE Transactions on Computers
- IEEE Transactions on Computational Biology and Bioinformatics
- INFORMS Journal of Computing
- ISMB 2008
- Journal of Bioinformatics and Computation Biology
- Journal of Computation Biology
- Journal of Proteome Research
- RECOMB 2001, 2002, 2003, 2005, 2008
- Operations Research Letters
- Operations Research
- Proceedings of the National Academy of Sciences
- Proteomics