

## Michael Brudno

Email: [brudno@cs.toronto.edu](mailto:brudno@cs.toronto.edu)  
Tel: 416-978-2589  
Fax: 416-978-1455  
Web: <http://www.cs.toronto.edu/~brudno>

Offices: 286C Pratt Bldg and 604 CCBR  
Mailing: 10 King's College Rd, SF 3303  
University of Toronto  
Toronto, ON M5S 3G4 Canada

### **Assistant Professor and Canada Research Chair in Computational Biology, University of Toronto** (start date: January 2006)

- Department of Computer Science
- Banting and Best Department of Medical Research
- Donnelly Centre for Cellular and Biomolecular Research

### **Adjunct Scientist, Hospital for Sick Children**

- Genetics and Genome Biology Division
- The Centre for Applied Genomics (TCAG; Scientific Management Board Member)

### **Education**

#### *Stanford University*

Ph.D. in Computer Science, September 2004.

Thesis Title: Algorithms for Comparison of DNA Sequences

Thesis Supervisor: Serafim Batzoglou

M.S. in Computer Science, March 2003.

#### *University of California, Berkeley*

A.B. degree with High Honors, May 2000.

Double major in Computer Science (Honors Thesis) and History

### **Previous Positions**

#### *Massachusetts Institute of Technology*

Visiting Scientist, CSAIL 9-12/2005

Host: Bonnie Berger

#### *University of California, Berkeley*

Postdoctoral Fellow, Computer Science Division 9/2004 – 9/2005.

Supervisor: Gene Myers

#### *GosNIIGenetika, Moscow, Russia*

Visiting Fellow, Scientific Center for Biotechnology 9/2003 – 12/2003

Host: Mikhail Gelfand

#### *Perlegen Sciences, Inc.*

Scientific Programmer. 10/2000 – 5/2001

Supervisor: Kelly A. Frazer

### **Awards & Honors**

- Ontario Early Researcher Award (2009-2014)
- Eurosyst 2009 Best Paper Award
- Canada Research Chair in Computational Biology (2006-2011)
- ISMB 2004 Best Paper Award

**Refereed Publications** (Papers where I am a PI or Corresponding author are noted with a \*, students working with me are in **bold**)

1. Dralyuk I, Brudno M, Gelfand M, Zorn M, Dubchak I. ASDB: database of alternatively spliced genes. *Nucleic Acids Res.* **28**:296-7. 2000
2. Dubchak I, Brudno M, Pachter LS, Loots GG, Mayor C, Rubin EM, Frazer K. Active conservation of noncoding sequences revealed by 3-way species comparisons. *Genome Res.* **10**:1304-6. 2000
3. Mayor C, Brudno M, Schwartz JR, Poliakov A, Rubin EM, Frazer K, Pachter LS, Dubchak I VISTA: Visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*, **16**:1046-1047. 2000
4. Brudno M, Gelfand MS, Spengler S, Zorn M, Dubchak I, Conboy JG Computational analysis of candidate intron regulatory elements for tissue-specific alternative pre-mRNA splicing. *Nucleic Acids Res.* **29**:2338-48 2001
5. Brudno M, Morgenstern, B. Fast and sensitive alignment of large genomic sequences. *Proceedings of the IEEE Computer Society Bioinformatics Conference (CSB 2002)*.
6. Brudno M, **Do CB**, Cooper GM, **Kim MF**, Davydov E, Green ED, Sidow A, Batzoglou S. LAGAN and Multi-LAGAN: Efficient tools for large-scale multiple alignment of genomic DNA. *Genome Res.* **13**:721-731, 2003.
7. Cooper GM, Brudno M, Green ED, Batzoglou S, Sidow A. Quantitative estimates of sequence divergence for comparative analyses of mammalian genomes. *Genome Res.* **13**:813-820, 2003.
8. Brudno M, **Malde S**, Poliakov A, **Do CB**, Couronne O, Dubchak I, Batzoglou S. Global alignment: finding rearrangements during alignment. *Special Issue on the Proceedings of the ISMB 2003, Bioinformatics* **19**: 54i-62i, 2003.
9. Taher L, Rinner O, Garg S, Sczyrba A, Brudno M, Batzoglou S, Morgenstern B. AgenDA: homology-based gene prediction. *Bioinformatics* **19**:1575-1577, 2003.
10. Brudno M, Chapman M, Göttgens B, Batzoglou S, Morgenstern B. Fast and sensitive multiple alignment of large genomic sequences. *BMC Bioinformatics* **4**:66, 2003.
11. Shan N, Couronne O, Pennacchio LA, Brudno M, Batzoglou S, Joy S, Bethel W, Rubin EM, Hamann B, Dubchak I. Phylo-VISTA: an interactive visualization tool for multiple DNA sequence alignments. *Bioinformatics* **20**: 636-643, 2004.
12. Rat Genome Sequencing Project Consortium (RGSPC). Genome sequence of the Brown Norway Rat yields insights into mammalian evolution. *Nature* **428**:493–521, 2004.
13. Cooper GM, Brudno M, Stone ES, Dubchak I, Batzoglou S, Sidow A. Characterization of evolutionary rates and constraints in three mammalian genomes. *Genome Res.* **14**:539–48, 2004.
14. Brudno M, Poliakov A, Salamov A, Cooper GM, Sidow A, Rubin EM, Solovyev V, Batzoglou S, Dubchak I. Automated whole-genome multiple alignment of Rat, Mouse, and Human. *Genome Res.* **14**:685–692, 2004.
15. Brudno M, Steinkamp R, Morgenstern B. The CHAOS/DIALIGN WWW server for multiple sequence alignment. *Nucleic Acids Research* **32**:W41 – W44, 2004.
16. Do CB, Brudno M, Batzoglou S. ProbCons: Probabilistic consistency-based multiple alignment of amino acid sequences. *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI)* 2004.

17. **Sundararajan M**, Brudno M, Small K, Sidow A, Batzoglou S. Chaining algorithms for alignment of draft sequence. *Proceedings of the fourth Workshop on Algorithms in Bioinformatics (WABI)* 2004.
18. The ENCODE Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, **306**: 636–640, 2004.
19. Do CB, Mahabhashyam M, Brudno M, Batzoglou S. ProbCons: Probabilistic consistency-based multiple alignment of amino acid sequences. *Genome Res.* **15**: 330–340, 2005
20. Kovaleva G., Bazykin G., Brudno M., Gelfand M. Comparative genomics of transcriptional regulation in yeasts and its application to identification of a candidate alpha-isopropylmalate transporter. *Journal of Bioinformatics and Computational Biology* 4:981-998, 2006
21. Small KS, Brudno M, Hill MM, Sidow A. Extreme genomic variation in a natural population. *Proc Natl Acad Sci U S A.* 104(13):5698-703. 2007.
22. Small KS, Brudno M, Hill MM, Sidow A. A haplome alignment and reference sequence of the highly polymorphic *Ciona savignyi* genome. *Genome Biol.* 8(3):R41 2007
23. Brudno M, Poliakov A, Minovitsky S, Ratnere I, Dubchak I. Multiple whole genome alignments and novel biomedical applications at the VISTA postal. *Nucleic Acids Research* 35:W669-74 2007
24. Bazykin GA, Kondrashov FA, Brudno M, Poliakov A, Dubchak I, Kondrashov AS. Extensive parallelism in protein evolution. *Biol Direct.* 2:20 2007
- 25\* **Medvedev P**, Georgiou K, Myers EW, Brudno M. Computability and equivalence of models for genome assembly. *Proceedings of the Workshop on Algorithms in Bioinformatics (WABI)*, 2007
- 26\* **Dalca AV**, Brudno M. FRESCO: Flexible Alignment with Rectangle Scoring Schemes. *Proceedings of the Pacific Symposium on Biocomputing (PSB)* 2008
- 27\* **Medvedev P**, Brudno M. Ab initio Whole Genome Shotgun Assembly with Mated Short Reads. *Proceedings of the 12th Annual Research in Computational Biology Conference (RECOMB)*, 2008
- 28\* **Lee S, Cheran E**, Brudno M. A Robust Framework for Detecting Structural Variations in a Genome. *Proceedings of the 16th Annual International Conference on Intelligent Systems in Molecular Biology (ISMB); Bioinformatics* 24:i59-i67, 2008
- 29\* **Yanovsky V, Rumble S**, Brudno M. Read Mapping Algorithms for Single Molecule Sequencing Data. *Proceedings of the Workshop on Algorithms in Bioinformatics (WABI)* 2008.
30. Quon G, The YW, Chan ET, Hughes TR, Brudno M, Morris QD. A mixture model for the evolution of gene expression in non-homogeneous datasets. *Proceedings of The Neural Information Processing Systems Conference (NIPS)* 2008.
- 31\* **Donmez N**, Bazykin G, Brudno M, Kondrashov A. Polymorphism due to multiple amino acid substitutions at a codon site within *Ciona savignyi*. *Genetics*, 181: 685–690 2009
- 32\* Lagar-Cavilla A, **Whitney JA**, Scannel A, **Rumble SM**, Patchin P, de Lara E, Brudno M, Satyanarayanan, M. SnowFlock: Rapid Virtual Machine Cloning for Cloud Computing. *Proceedings of the European Conference in Computer Systems (Eurosys)* 2009 (**Best Paper Award**).
- 33\* Dubchak I, Poliakov A, Kislyuk A, Brudno M. Multiple Whole Genome alignment without a Reference Sequence. *Genome Res.* 19:682-689 2009.

- 34\*. Patchin P, Lagar-Cavilla HA, de Lara E, Brudno M. Adding the Easy Button to the Cloud with Snowflock and MPI. *3rd Workshop on System-level Virtualization for High Performance Computing (HPCVirt)* 2009
35. Chan ET, Quon GT, Chua G, Babak T, Trochesset M, Zirngibl RA, Aubin J, Ratcliffe M, Wilde A, Brudno M, Morris QD, Hughes TR. Conservation of core gene expression invertebrate tissues. *Journal of Biology*, 8:33 2009
- 36\*. **Rumble SM**, Lacroute P, **Dalca AV**, **Fiume M**, Sidow A, Brudno M. SHRiMP: Accurate Mapping of Short Color-space Reads. *PLoS Computational Biology*, 5:5 2009
- 37\*. **Medvedev P**, Brudno M. Maximum Likelihood Genome Assembly. *Journal of Computational Biology* 16:1101-1116 2009. (Extended journal version of #27).
- 38\*. **Lee S**, Hormozdiari F, Alkan C, Brudno M. MoDIL: detecting small indels from clone-end sequencing with mixtures of distributions *Nature Methods*, 6:473-474 2009
- 39\*. **Medvedev P**, **Stanciu M**, Brudno M. Computational methods for discovering structural variation with next generation sequencing. *Nature Methods*, 6:S13–20 2009 (Review).
- 40\*. **Dalca A**, Brudno M. Genome Variation Discovery with High Throughput Sequencing Data. *Briefings in Bioinformatics*, in press 2010 (Review).
- 41\*. **Lee S**, Xing EP, Brudno M. MOGuL: Detecting Common Insertions and Deletions in a Population. *Proceedings of the 14th Annual Research in Computational Biology Conference (RECOMB)*, in press 2010

### **Book Chapters**

1. Brudno M and Dubchak I. Comparison of Long Genomic Sequences: Algorithms and Applications. In S. Aluru, ed. *Handbook of Computational Biology*, CRC Press, 2006.
2. Brudno M, LAGAN Alignment Toolkit, book chapter in *Comparative Genomics, Methods in Molecular Biology* Nicholas Bergman, ed. Published by Humana Press, 2007

### **Unrefereed Publications**

1. Brudno M. Building Genomes. *Idea&s: The Arts and Science Review*, University of Toronto. Fall 2007, p. 21
2. Brudno M, Moret B, Linder R, Warnow T; Beyond Gap Models: Alignment and Phylogeny under genome-scale events: Session Introduction. *Pacific Symposium on Biocomputing (PSB)* 13:1-2(2008)
3. Brudno M, Medvedev P, Stoye J, De La Vega, F. A Report on the 2009 SIG on Short Read Sequencing and Algorithms (Short-SIG). *Bioinformatics*, 2009

### **Invited Conference Presentations** (last 5 years; altogether 15 in 8 countries)

#### **Next Generation Sequencing Technologies**

- Canada-Taiwan Biotechnology Symposium, Toronto, Canada, November 2008
- Toronto Universities Biotechnology Society Symposium, Toronto, Canada April 2008

#### **Discovering Genome Variation with Matepair Data**

- Rendezvous Sequenage meeting, Genome Quebec, Montreal December 2008

#### **Discovering INDEL and CNV Variation in the Human Genome.**

- 2<sup>nd</sup> BeNeLux NGS User Meeting (**Keynote address**), Utrecht, Netherlands, July 2009
- MonBUG Symposium, Montreal, Canada, September 2009
- NGS2009 Conference, Barcelona, Spain, October 2009

#### **Whole Genome Alignment without a Reference Sequence**

- CRNG Quantomics Workshop, Barcelona, Spain, September 2009

### **Invited Seminar Talks** (last 5 years; of > 40 in 6 countries)

- **Discovering INDEL and CNV Variation in the Human Genome.**  
June 2009, J. Craig Venter Institute (Rockville, MD)  
November 2009, Weizmann Institute, Israel
- **Assembling Genomes with Next Generation Sequencing Data**  
September 2008, Ben Gurion University (Be'er Sheva, Israel)  
February 2008, Genelogic Inc.
- **Discovering Genome Variation with Matepair Data**  
November 2008, University of Waterloo  
February 2008, DOE Joint Genome Institutes  
February 2008, Stanford University  
February 2008, Applied Biosystems (Foster City, CA)  
February 2008, University of Maryland  
June 2008, Helicos (Cambridge, MA)
- **Chinese Postmen and Virtual Clusters**  
May 2007, Google
- **Predicting Ancestral Genome Order**  
November 2005, Moscow Seminar on Computational Biology  
November 2005, MIT-CSAIL Bioinformatics Seminar
- **Alignment of Whole Genomes: Algorithms & Tools**  
September 2005, University of Massachusetts

### **Lectures at Summer Schools & Special Courses (last 5 years):**

June 2005, Lawrence Berkeley Laboratory, California; PGA Educational Workshop  
November 2005, University of Utrecht, Netherlands; Mining the Genome PhD Course  
February 2006, Bioinformatics.ca Workshop, Toronto Canada; Course on Genomics  
July 2007, Cluster in Biomedicine, Trieste, Italy Summer School on Bioinformatics  
July 2008, Bioinformatics.ca Workshop, Toronto Canada; Course on Sequencing  
July 2009, Bioinformatics.ca Workshop, Toronto Canada; Course on Sequencing

### **Student supervision**

#### Graduate

Marc Fiume (BSc 2008-2009, MSc 2009-present)  
Paul Medvedev (PhD 2006-present, co-supervised with Allan Borodin)  
Nilgun Donmez (MSc, 2006-2008, PhD 2008-present)  
Joseph Whitney (PhD 2007-present)  
Vladimir Yanovsky (PhD 2007-present)

#### Undergraduate

Michael Dzamba (BSc in Bioinformatics, 05/2009-present)

#### Alumni

Elango Cheran (MSc, 2006-2008), now Programmer at Sick Children's Hospital (Toronto)  
Stephen Rumble (BSc 2007-2008), now PhD student at Stanford University  
Seunghak Lee (MSc, 2007-2009), now PhD student at Carnegie Mellon University  
Adrian Dalca (BSc 2006-2008, MSc 2008-2009) now PhD student at MIT

### Previous Undergraduate Research Supervision

Michael Andruschenko (BSc in Statistics, 05/2007-09/2007)

Botond Ballo (BSc in Math and Computer Science, 05/2009-09/2009)

Taya Santare (BSc in Math and Computer Science, 05/2008-05/2009)

### Teaching

#### Undergraduate:

CS61B: Data Structures and Advanced Programming (Summer 2000, UC Berkeley)

BCB 410: Applied Bioinformatics (Fall 2007)

CSC 373: Algorithms (Spring 2009, Spring 2010)

#### Graduate:

CSC 2431: Topics in Computational Biology (Spring 2008, Spring 2010)

CSC 2417: Algorithms in Genomics (Spring 2006, Fall 2006, Spring 2009)

### Professional Activities

**Associate Editor** of the Algorithms for Molecular Biology journal

#### **Conference Organizer**

- HiTSeq: High Throughput Sequencing SIG (at Intelligent Systems for Molecular Biology conference); 2008 (Toronto, 120 attendees), 2009 (Stockholm, 200 attendees), 2010 (Boston)

#### **Session Organizer or Chair**

- Pacific Symposium for Biocomputing (PSB) 2008, Session on Alignment and Evolution
- CSHL/Wellcome Trust Genome Informatics 2009, Session on Next Gen Sequencing
- Pacific Symposium for Biocomputing (PSB) 2010, Session on Personal Genomics
- Intelligent Systems for Molecular Biology (ISMB) 2010, Sequence Analysis Track Chair

**Program Committees:** ISMB 2008, 2009, 2010 (Track Chair), RECOMB 2009, 2010, COCOON 2009, BIODDD 2008, Recomb Comparative Genomics 2006, 2010.

**Guest associate editor** for the PLoS Computational Biology journal

**Referee** for Science, Nature Methods, Genome Research, PLoS Genetics, PLoS Computational Biology, Genetics, Molecular Biology and Evolution, Molecular and Systems Biology, RECOMB, ISMB, Bioinformatics Journal, BMC Bioinformatics, BMC Genomics, Journal of Bioinformatics and Computational Biology, ACM/IEEE Transactions on Computational Biology and Bioinformatics, IEEE CSB, SODA, FOCS, IEEE Transactions on Computers.

**Grant reviewer** for National Institutes of Health (NIH, USA), NSERC, Health Canada, Netherlands Genomics Institute, French National Research Agency, Canada Research Chairs program.