

Mathieu Blanchette - Curriculum Vitae in a Nutshell

Career path:

- PhD in computer science (2002) from U. of Washington.
- Assistant (2003-2008) and associate (2008-now) professor at McGill's School of Computer Science since 2003.

Research topics and productivity:

- Research area: bioinformatics, with applications in genomics, epigenomics, evolution, phylogenetics, gene regulation, and proteomics.
- Author of 70 peer-reviewed journal papers (published in Nature, Nature Genetics, Genome Research, PNAS, Nucleic Acids Research, Genome Biology, etc.) and 28 peer-reviewed conference papers (note: in bioinformatics, top publication venues include both journals and peer-reviewed conferences), and 6 book chapters.
- More than 6,800 citations (Google scholar Feb. 2015).
- Gave approximately 70 invited talks in the last 10 years.

Research funding:

- Recipient of the largest NSERC Discovery grant in Canada among researchers below 40 years old in the Genes and Molecules evaluation group (904 researchers Canada-wide) and of a Discovery Accelerator Supplement (2013-2016).
- Principal investigator on grants from NSERC, Genome Quebec, Genome Canada, FQRNT, and co-investigator on several CIHR grants.
- Received more than \$300,000 in competitive operating funding in 2014.

Selected research awards:

- Chris Overton prize (2006) from the International Society for Computational Biology for outstanding accomplishment for a scientist in the early to mid stage of his career.
- Outstanding Young Computer Scientist Researcher Prize from the Canadian Association for Computer Science (2012).
- Alfred P. Sloan fellow (2007-2009).

Service to the research community:

- Member of NSERC Discovery Genes and Molecules selection group since 2013
- Member of NIC CGAT review panel (2007,2008,2011)
- Associate editor of Genome Research (2007-2009; top journal in genomics), Frontier in Computational Biology and Algorithms for Molecular Biology
- Organizer of an annual conference on computational genomics in Barbados since 2004.
- Member of the program committee of ISMB and RECOMB for most of the last 10 years.

Teaching and supervision:

- Leo Yaffe teaching award in 2008 (Faculty of Science, McGill)
- Graduated 7 Ph.D. and 18 M.Sc. students in 10 years
- Alumni from the lab are now professors at Waterloo University, Université du Québec à Montréal, Université de Sherbrooke, and U. Montpellier II (France).

Mathieu Blanchette – Complete Curriculum Vitae

Contact information:

blanchem@mcb.mcgill.ca
514-398-5209 (office)
514-398-3387 (fax)
<http://www.mcb.mcgill.ca/~blanchem>

Postal Address:

School of Computer Science
McGill University
3480 University Street
McConnell Eng. Bldg., Room 318
Montreal, Quebec, H3A 2A7
CANADA

Positions held

- 2008-now** Associate professor, School of Computer Science (McGill Centre for Bioinformatics), McGill University.
- 2003-2008** Assistant professor, School of Computer Science (McGill Centre for Bioinformatics), McGill University.
- 2002-2003** Postdoctoral researcher, Haussler's lab, University of California, Santa Cruz.
- 1999-2002** Research assistant, Dept. of Computer Science, U. of Washington.
- 1998-2000** Teaching assistant, Dept. of Computer Science, U. of Washington.
- 1995-1998** Research assistant, Dept. of Mathematics and statistics, Université de Montréal.

Education

- 2002-2003** Post-doctoral researcher, Center for Biomolecular Science and Engineering, University of California in Santa Cruz.
Advisor: David Haussler.
- 1998-2002** Ph.D. in Computer Science, University of Washington, Seattle. (June 2002).
Thesis: Algorithms for Phylogenetic Footprinting.
Advisor: Martin Tompa.
- 1997-1998** M.Sc. in Computer Science, Université de Montréal, Montréal, Canada (1998).
Thesis: Breakpoint Phylogeny (Phylogénétique basée sur les cassures du génome).
Advisor: David Sankoff.
- 1994-1997** B.Sc. in Mathematics and Computer Science, Université de Montréal, Montreal, Canada (1997).
Advisor: David Sankoff. GPA: 4.18 / 4.3.

Research Grants Currently Held

- 2013-2015 Genome Canada Bioinformatics and Computational Biology Competition**
Title: User-friendly application for high-resolution transposable element identification and characterization
PI: Mathieu Blanchette
Co-PIs: Thomas Bureau
Amount: 125,000\$/year
- 2013-2015 Genome Canada Bioinformatics and Computational Biology Competition**
Title: A Development and Deployment Platform for Citizen Science Games in biology
PI: Jérôme Waldispühl
Co-PIs: Mathieu Blanchette, Derek Ruths
Amount: 125,000\$/year
- 2013-2016 Genome Canada Bioinformatics and Computational Biology Competition**
Title: Genomics Large Data Sets and Novel Tools for Plant Biology for use International Infrastructure-Tier Data Repositories and Portals
PI: Nicholas Provar
Amount: 125,000\$/year
- 2013-2017 NSERC Accelerator**
Title: Inference and analysis of ancestral genomes
Amount: 40,000\$/year
- 2013-2017 NSERC Discovery**
Title: Inference and analysis of ancestral genomes
Amount: 69,000\$/year
- 2013-2016 FQRNT Team grant**
Title: Évolution des gènes d'ARN de transfert à l'échelle génomique
PI: Mathieu Blanchette. *Co-PIs:* Eric Lecuyer, Jerome Waldispühl.
Amount: 48,000\$/year
- 2013-2017 Canadian Foundation for Innovation – Leader Edge Fund**
Title: The Montreal CardioMet Biomarker and Drug Discovery Consortium
PI: Mark Prentki
Co-PIs: Pavel Hamet, Benoit Coulombe, Rob Sladek, Remi Rabasa-Lhoret, Mathieu Blanchette, Vincent Poitout, Shant Dersakissian, Erik Joly, Nabil Seidah.
Amount: 2 312 056 \$
- 2012-2013 NSERC Research Tools and Instrumentation**
Title: Computing infrastructure for bioinformatics web services
PI: Mathieu Blanchette
Co-PIs: Derek Ruths and Jérôme Waldispühl
Amount: 40173\$/year
- 2012-2017 CIHR CEEHRC Epigenomics platform – Data coordination Center**
Title: Integrative Epigenomic Data Coordination Centre (EDCC) at McGill
PI: Guillaume Bourque
Amount: 219,796\$/year
- 2012-2017 CIHR CEEHRC Epigenomics platform – Epigenomics Mapping Center**
Title: Epigenomics Mapping Centre (EMC) at McGill

PI: Mark Lathrop
Amount: 906,400\$/year

Research Grants Previously Held

2010-2013 **Genome Québec Health Program**

Title: Host macrophage transcriptomic responses to *M. tuberculosis* infection.

PI: John White

Co-PIs: Mathieu Blanchette and Rob Sladek

Amount: 424385\$/year

2009-2013 **Genome Canada Competition IV**

Title: Bridging comparative, population and functional genomics to identify and experimentally validate novel regulatory regions and genes for crop improvement

PI: Thomas Bureau

Co-PIs: Mathieu Blanchette, Stephen Wright, John Stinchcomb, Alan Moses, Paul Harrison, Daniel Schoen, Anwar Naseem and Ken Dewar.

Amount: 1177283\$/year

2008-2012 **CIHR Team grant**

Title: An integrative forward genetic approach to identify novel pathways in host response to infection: from mouse models to patients

PI: Silvia Vidal

Co-PIs: Mathieu Blanchette, Philippe Gros, Danielle Malo, Maya Saleh, Salman Qureshi

Amount: 828,994\$/year

2008-2013 **NSERC Discovery Grant**

Title: Algorithms for whole-genome comparative and regulatory genomics

Amount: 31000\$/year

2007-2009 **Alfred P. Sloan Research Fellowship**

Title: Computational and evolutionary molecular biology

Amount: 45,000\$

Note: Sloan fellowships are awarded yearly to approximately 100 pre-tenured researchers in North America.

2007-2012 **CIHR Operating Grant**

Title: Proteomic analysis of the cell machinery that interprets the human genome

PI: Benoit Coulombe (IRCM)

Co-PI: Mathieu Blanchette

Amount: 177,500\$/year (Personal share: ~15%)

2005-2008 **Genome Canada Competition III**

Title: GRID: Gene Regulators in Diseases

PI: Thomas Hudson (McGill)

Co-PIs: Mathieu Blanchette, Alan Peterson, Daniel Sinnett, Constantin Polychronakos, Ken Dewar, Jacek Majewski, Damian Labuda, Bartha Knoppers, Anna Naumova, Rob Sladek, Tommi Pastinen.

Amount: 9,404,836\$ (Personal share: ~1.5%)

- 2005-2009 CIHR Operating Grant**
Title: Bioinformatics detection and experimental validation of splicing control elements in mammalian introns
PI: Jacek Majewski
Co-PI: Mathieu Blanchette
Amount: 110,384\$/year (Personal share: ~20%)
- 2005-2008 FQRNT Team Grant**
Title: Algorithms for reconstructing ancestral genomes
PI: Nadia El-Mabrouk (U. of Montreal)
Co-PIs: Mathieu Blanchette, David Sankoff, Damian Labuda,
Amount: 55,000 \$/year (Personal share: 25%)
- 2005-2008 New Opportunity Grant**
Title: Computing infrastructure for the analysis of non-coding functional regions of the human genome.
Amount: 224,000\$
- 2005-2009 CFI Infrastructure Operating Grant**
Title: Computing infrastructure for the analysis of non-coding functional regions of the human genome
Amount: 5,000\$ / year
- 2003-2007 NSERC Discovery Grant**
Title: Algorithms for comparative analysis of biological sequences
Amount: 29,000 \$ / year
Note: The grant was originally awarded for four years, but NSERC extended it by one year. I am re-applying in 2007.
- 2003 FQRNT Equipment Grant**
Title: Algorithmes pour l'analyse comparative de séquences biologiques.
Amount: 15,000 \$
- 2003-2006 FQRNT New Researcher Grant**
Title: Algorithmes pour l'analyse comparative de séquences biologiques.
Amount: 15,000 \$ / year
- 2003-2007 FQRNT Strategic Researcher**
Amount: 2,000 \$ / year + teaching load reduction
- 2003-2008 McGill startup grant**
Amount: 60,000\$
- 2005-2006 Genome Quebec Comparative and Integrative Genomics Project**
Title: Identification of human cis-regulatory modules through comparative genomics.
Amount: 18,000 \$
- 2004-2005 Genome Quebec Transcriptional Regulation Project**
Title: Computational detection of human regulatory elements
Amount: 18,000 \$

Publications – Most significant contributions

1. Comparative genomics approaches to cis-regulatory module annotation

In Blanchette et al. (*Genome Res.* 2006) we introduce a new computational approach to the prediction of cis-regulatory modules (CRM) and apply it to the complete human genome. A map of more than 100,000 putative CRM is produced and reveals a number of unexpected features, including the presence of regulatory modules at the 3' end of genes and the unexpected localization preferences of certain transcription factors. The associated public database, called PreMod (Ferretti et al. *NAR* 2007), allows researchers to identify regulatory regions for their genes of interest. We have since validated many cis-regulatory module predictions (Dufour et al. *Cell Met.* 2007, Mongin et al. *Plos One* 2011) and mapped them to their target genes (Mongin et al. *J. Comp. Biol.* 2011, Mongin et al. *BMC Evol. Genom.* 2009). We recently applied a similar approach to obtain the first comprehensive map regulatory regions in Brassicaceae (Haudry et al. *Nature Genetics* 2013).

2. Ancestral genome reconstruction

In Blanchette et al. (*Genome Res.* 2004), we introduce the very non-intuitive and exciting idea that a very large fraction of an ancestral mammalian genome can be reconstructed with a 98% base-by-base accuracy. New algorithmic techniques are presented and extensive simulations are carried to establish this number. Applications are important: study of mammalian evolution and genome evolution in general, detection of conserved regions or of regions conserved in a lineage-specific fashion. This paved the way to a number of other publications, including algorithmic development (Hickey et al. *RECOMB* 2011, Diallo et al. *Bioinformatics* 2010, Diallo et al. *J. Comp. Biol.* 2007) applications (Sadri et al. *Bioinformatics* 2011).

3. Inference and analysis of protein-protein interaction networks

We are developing tools for the reliable inference of protein-protein interaction (PPI) networks from AP-MS data (Lavallée-Adam et al. *J. Prot. Res.* 2011; *J. Prot. Res.* 2013). These tools have been key to the success of collaborative projects that have led to important collaborative biological publications (Forget et al. *Mol. Cell Prot* 2010, Jeronimo et al. *Mol. Cell* 2007). In parallel, we are developing computational approaches to extract biological knowledge out of complex networks (Lavallée-Adam et al. *J. Comp. Biol.* 2010; Kim et al. *Alg. Mol. Biol.* 2010).

4. Genetics and epigenetics of gene regulation

Genetic variation within a population translates in differences in gene expression, often with an impact on health. It can also affect DNA methylation, which in turn impact gene expression. In collaboration with Tomi Pastinen, we are developing approaches to reliably detect allelic expression in a hypothesis-free manner (Ge et al. *Nature Gen.* 2009, Wagner et al. *Plos. Comp. Biol.* 2010) from noisy experimental data. We then develop the computational approaches required for the analysis of population methylation data (Wagner et al. *Genome Biology* 2014).

5. Analysis of 3D chromatin conformation

Inside a cell's nucleus, chromatin is folded into a complex, dynamic three-dimensional conformation. In collaboration with Josée Dostie, we are developing computational approaches to robustly infer 3D structure from experimental data and investigate biological applications to the Hox clusters (Rousseau et al. *BMC Bioinformatics* 2011, Rousseau et al. *NAR* 2014), with potential applications as leukemia subtype biomarker (Rousseau et al. *Genome Biology* 2014).

Journal publications

Note: The names of the students I have supervised or co-supervised are *italicized*. Co-corresponding authors are denoted with a *.

- [70j] *Cingolani P*, Sladek R, Blanchette M. (2015) **BigDataScript: a scripting language for data pipelines**. *Bioinformatics*. 31(1):10-6.
- [69j] Fodil N, Langlais D, Moussa P, Boivin GA, Di Pietrantonio T, Radovanovic I, Dumaine A, Blanchette M, Schurr E, Gros P, Vidal SM. (2014) **Specific dysregulation of IFN γ production by natural killer cells confers susceptibility to viral infection**. *PLoS Pathogen*. 10(12):e1004511.
- [68j] Williamson RJ, Josephs EB, *Platts AE*, Hazzouri KM, Haudry A, Blanchette M, Wright SI. (2014) **Evidence for widespread positive and negative selection in coding and conserved noncoding regions of *Capsella grandiflora***. *PLoS Genetics*. 10(9):e1004622.
- [67j] *Rousseau M*, Ferraiuolo MA, Crutchley JL, Wang XQ, Miura H, Blanchette M*, Dostie J.* (2014) **Classifying leukemia types with chromatin conformation data**. *Genome Biology* 15:R60.
- [66j] *Wagner JR*, Busche S, Ge B, Kwan T, Pastinen T*, Blanchette M* (2014) **The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts**. *Genome Biol*. 15(2):R37.
- [65j] Cloutier P, *Lavallée-Adam M*, Faubert D, Blanchette M, Coulombe B. (2014). **Methylation of the DNA/RNA-binding protein Kin17 by METTL22 affects its association with chromatin**. *J Proteomics*. 100:115-24.
- [64j] *Rousseau M*, Crutchley JL, Miura H, Suderman M, Blanchette M, Dostie J. (2014) **Hox in motion: tracking HoxA conformation during differentiation**. *Nucl. Acids. Res*. 42(3):1524-40.
- [63j] Caignard G, Leiva-Torres GA, Leney-Greene M, Charbonneau B, Dumaine A, Fodil-Cornu N, Pyzik M, *Cingolani P*, Schwartzentruber J, Dupaul-Chicoine J, Guo H, Saleh M, Veillette A, Lathrop M, Blanchette M, Majewski J, Pearson A, Vidal SM. (2013) **Genome-Wide Mouse Mutagenesis Reveals CD45-Mediated T Cell Function as Critical in Protective Immunity to HSV-1**. *PLoS Pathog*. 9(9):e1003637.
- [62j] Kwak D, Kam A, Becerra D, Zhou Q, Hops Q, Zarour E, Sarmenta L, Blanchette M, Waldspühl J. (2013) **Open-Phylo: a customizable crowd-computing platform for multiple sequence alignment**. *Genome Biology*. 14:R116.
- [61j] Haudry A, *Platts AE*, Vello E, Hoen D, *Leclercq M*, Williamson R, Forczek E, Joly-

- Lopez Z, Steffen J, Hazzouri KM, Dewar K, Stinchcombe JR, Schoen DJ, Wang X, Schmutz J, Town JD, Edger PP, Pires JC, Schumaker KS, Jarvis DE, Mandáková T, Lysak MA, Schranz ME, van den Bergh E, Harrison P, Moses AM*, Bureau TE*, Wright SI*, Blanchette M*. **An Atlas of over 90,000 Conserved Non-Coding Sequences Yields a Detailed Map of Crucifer Regulatory Regions.** *Nature Genetics*. 45: 891-898.
- [60j] Zhang Y, Ponty Y, Blanchette M, Lecuyer E, Waldispuhl J (2013) **SPARCS: a web server to analyze (un)structured regions in coding RNA sequences.** *Nucleic Acids Research*. W480-5.
- [59j] Leclercq M, Diallo AB, Blanchette M. **Computational prediction of the localization of microRNAs within their pre-miRNA.** *Nucleic Acids Research*. 41(15):7200-11.
- [58j] Forget D, Lacombe AA, Cloutier P, Lavallée-Adam M, Blanchette M, Coulombe B. (2013) **Nuclear import of RNA polymerase II is coupled with nucleocytoplasmic shuttling of the RNA polymerase II-associated protein 2.** *Nucleic Acids Res*. 41 (14):6881-91.
- [57j] Slotte T, Hazzouri KM, Ågren JA, Koenig D, Maumus F, Guo YL, Steige K, Platts AE, Escobar JS, Newman LK, Wang W, Mandáková T, Vello E, Smith LM, Henz SR, Steffen J, Takuno S, Brandvain Y, Coop G, Andolfatto P, Hu TT, Blanchette M, Clark RM, Quesneville H, Nordborg M, Gaut BS, Lysak MA, Jenkins J, Grimwood J, Chapman J, Prochnik S, Shu S, Rokhsar D, Schmutz J, Weigel D, Wright SI. (2013) **The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution.** *Nature Genetics*. 45(7):831-5.
- [56j] Meunier C, Van Der Kraak L, Turbide C, Groulx N, Labouba I, Cingolani P, Blanchette M, Yeretssian G, Mes-Masson AM, Saleh M, Beauchemin N, Gros P. (2013) **Positional mapping and candidate gene analysis of the mouse *ccs3* locus that regulates differential susceptibility to carcinogen-induced colorectal cancer.** *PLoS One*.8(3):e58733.
- [55j] Cloutier P, Lavallée-Adam M, Faubert D, Blanchette M, Coulombe B. (2013) **A newly uncovered group of distantly related lysine methyltransferases preferentially interact with molecular chaperones to regulate their activity.** *PLoS Genet*. 2013 Jan;9(1):e1003210.
- [54j] Lavallée-Adam M, Rousseau J, Domecq C, Bouchard A, Forget D, Faubert D, Blanchette M, Coulombe B. (2013) **Discovery of Cell Compartment Specific Protein-Protein Interactions using Affinity Purification Combined with Tandem Mass Spectrometry.** *J Proteome Res*. Jan 4;12(1):272-81.
- [53j] Blanchette M. (2012) **Exploiting ancestral mammalian genomes for the prediction of human transcription factor binding sites.** *BMC Bioinformatics*. 13 Suppl 19:S2.

- [52j] Gagnon Y., Blanchette M., El-Mabrouk, N. (2012) **A Flexible Ancestral Genome Reconstruction Method based on Gapped Adjacencies.** *BMC Bioinformatics*. 13 Suppl 19:S4.
- [51j] Bongfen S.E., Rodrigue-Gervais I.G., Berghout J., Cingolani P., Torre S., Sladek R., Blanchette M., Behr M., Gruenheid S., Vidal S., Saleh M., Gros P. (2012) **An n-ethyl-n-nitroso-urea (ENU)-induced mutation in jak3 protects against cerebral malaria but causes susceptibility to mycobacteria.** *PLoS One*. 7(2): e31012.
- [50j] *Kawrykow A., Roumanis G., Kam A., Leung C., Wu C., Zarour E., Phylo players, Sarmenta L., Blanchette M.*, Waldispühl J.** (2011) **Phylo: A citizen science approach for improving multiple sequence alignment.** *PLoS One*. 7(3): e31362. [5 citations]
- [49j] *Hickey G., Blanchette M., Carmi P. Maheshwari A. Zeh N.* (2011) **An Approximation Algorithm for the Noah's Ark Problem with Random Feature Loss.** *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 8(2), 551-556.
- [48j] *Rousseau M., Fraser J., Ferraiuolo MA., Dostie J. and Blanchette M.* (2011) **Three-dimensional modeling of chromatin structure from interaction frequency data using Markov chain Monte Carlo sampling.** *BMC Bioinformatics*. 12:414
- [47j] Hickey G and Blanchette M. (2011) **A Probabilistic Model for Sequence Alignment with Context-Sensitive Indels.** *J. Computational Biology*. 18(11):1449-64.
- [46j] *Mongin, E., Dewar, K. and Blanchette M.* (2011) **Mapping association between long-range cis-regulatory regions and their target genes using synteny.** *J. Computational Biology*. 18(9):1115-30
- [45j] *Mongin, E., Auer. T.O., Bourrat. F., Gruhl F., Dewar, K., Blanchette, M., Wittbrodt, J. and Ettwiller, L.* (2011) **Combining computational prediction of cis-regulatory elements with a new enhancer assay to efficiently label neuronal structures in the medaka fish.** *PLoS One*. 2011;6(5):e19747.
- [44j] *Sadri, J., Diallo, A.B. and Blanchette M.* (2011) **Predicting site-specific human selective pressure using evolutionary signatures.** *Bioinformatics (Proceedings of ISMB 2011)*. 27(13):i266-i274.
- [43j] *Lavallée-Adam, M., Cloutier, P., Coulombe, B., and Blanchette, M.* (2011) **Modeling contaminants in AP-MS/MS experiments.** *J Proteome Research*, v. 10, n. 2, pp. 886-895.
- [42j] *Diallo, A., Makarenkov, V., and Blanchette, M.* (2010) **Ancestors 1.0: a web server for ancestral sequence reconstruction.** *Bioinformatics*, v. 26(1), pp. 130-1.
- [41j] *Ferraiuolo, M., Rousseau, M., Miyamoto, C., Shenker, S., Wang, X., Nadler, M.,*

- Blanchette, M., and Dostie, J. (2010) **The three-dimensional architecture of Hox cluster silencing.** *Nucleic Acids Research.*, v. 38, n. (21), pp. 7472-84.
- [40j] Forget, D., Lacombe, A., Cloutier, P., Al-Khoury, R., Bouchard, A., Lavallée-Adam, M., Faubert, D., Jeronimo, C., Blanchette, M., and Coulombe, B. (2010) **The protein interaction network of the human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II.** *Molecular Cell Proteomics*, v. 9, n. 12, pp. 2827-39.
- [39j] Wagner, J., Ge, B., Pokholok, D., Gunderson, K., Pastinen, T., and Blanchette, M. (2010) **Computational analysis of whole-genome differential allelic expression data in human.** *PLoS Computational Biology*, v. 6(7), p. e1000849.
- [38j] Lavallée-Adam, M., Coulombe, B., and Blanchette, M. (2010) **Detection of locally over-represented GO terms in protein-protein interaction networks.** *Journal of computational biology.*, v. 17(3), pp. 443-57.
- [37j] Hickey, G., Blanchette, M., Carmi, P., Maheshwari, A., and N., Z. (2010) **An approximation algorithm for the Noah's ark problem with random feature loss.** *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, v. 8, n. 2, pp. 551-6.
- [36j] Kim, E., Sabharwal, A., Vetta, A., and Blanchette, M. (2010) **Predicting direct protein interactions from affinity purification mass spectrometry data.** *Algorithms in Molecular Biology*, v. 5, n. 34, pp. 1-17.
- [35j] Fraser J, Rousseau M, Shenker S, Ferraiuolo MA, Hayashizaki Y, Blanchette M, Dostie J. (2009) **Chromatin conformation signatures of cellular differentiation.** *Genome Biology*. 10(4):R37
- [34j] Cloutier, P., Al-Khoury, R., Lavallée-Adam, M., Faubert, D., Jiang, H., Poitras, C., Bouchard, A., Forget, D., Blanchette, M., and Coulombe, B. (2009) **High-resolution mapping of the protein interaction network for the human transcription machinery and affinity purification of RNA polymerase II-associated complexes.** *Methods*. 48(4):381-6
- [33j] Mongin E, Dewar K, Blanchette M. (2009) **Long-range regulation is a major driving force in maintaining genome integrity.** *BMC Evol Biol*. 9(1):203.
- [32j] Hoberman R, Dias J, Ge B, Harmsen E, Mayhew M, Verlaan DJ, Kwan T, Dewar K, Blanchette M, Pastinen T. (2009) **A probabilistic approach for SNP discovery in high-throughput human resequencing data.** *Genome Research*. 19(9):1542-52.
- [31j] Ge B, Pokholok D.K., Kwan T., Grundberg E., Morcos L., Verlaan D.J., Le J., Koka V., Lam K.C.L., Gagné V., Dias J., Hoberman R., Montpetit A., Joly M.M., Harvey

- E.J., Sinnett D., Beaulieu P., Hamon R., Graziani A., Dewar K., Harmsen E., Majewski J., Göring H.H.H, Naumova A.K., Blanchette M., Gunderson K.L. and Pastinen T. (2009) **Global patterns of cis-variation in human cells revealed by high-density allelic expression analysis.** *Nature Genetics*. 41(11):1216-22.
- [30j] *Diallo A.B.*, Badescu D., Blanchette M., Makarenkov V. (2009) **A whole genome study and identification of specific carcinogenic regions of the human papilloma viruses.** *Journal of Computational Biology*. 16(10):1461-73.
- [29j] *Bertrand D.*, Blanchette M., El-Mabrouk N. (2009) **Genetic map refinement using a comparative genomic approach.** *Journal of Computational Biology*. 16(10):1475-86.
- [28j] *Fauteux F.*, Blanchette M. and Strömviik M.V. (2008) **Seeder: Discriminative Seeding DNA Motif Discovery.** *Bioinformatics*. 24(20):2303-7.
- [27j] *Smith M.*, Blanchette M., Papadopoulou B. (2008) **Improving the prediction of mRNA extremities in the parasitic protozoan Leishmania.** *BMC Bioinformatics*. 9:158.
- [26j] *Chen, X.*, Blanchette, M. (2007) **Prediction of tissue-specific cis-regulatory modules using bayesian networks and regression trees.** *BMC Bioinformatics*. 8 Suppl 10:S2.
- [25j] *Diallo, A.B.*, Makarenkov, V., Blanchette, M. (2007) **Exact and heuristic algorithms for the indel maximum likelihood problem.** *Journal of Computational Biology*. 14(4): 446-461.
- [24j] *Blin, G.*, *Blais, E.*, Hermelin, D., Guillon, P., Blanchette, M., El-Mabrouk, N. (2007) **Gene maps linearization using genomic rearrangement distances.** *Journal of Computational Biology*. 14(4): 394-407.
- [23j] *Jeronimo, C.*, Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Thérien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D. and Coulombe, B. (2007) **Analysis of a mammalian transcription machinery interaction network reveals the identity of the 7SK capping enzyme.** *Molecular Cell*, 27(2):262-274.
- [22j] *Dufour, C.R.*, Wilson, B.J., Huss, J.M., Kelly, D.P., Alaynick, W.A., Downes, M., Evans, R.M., Blanchette, M., Giguère, V. (2007) **Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma.** *Cell Metabolism*. 5(5):345-56.
- [21j] *Ferretti, V.*, Poitras, C., Bergeron, D., Coulombe, C., Robert, F., and Blanchette, M. (2007) **PReMod: a database of genome-wide mammalian cis-regulatory module predictions.** *Nucl. Acids Res*. 35: D122-D126.
- [20j] *Chen, H.* and Blanchette, M. (2006) **Detecting non-coding selective pressure in coding regions.** *BMC Evolutionary Biology*, Feb 8;7 Suppl 1:S9.

- [19j] Blanchette, M., Bataille, A.R., *Chen, X.*, Poitras, C., Laganière, J., Lefèbvre, C., Deblois, G., Giguère, V., Ferretti, V., Bergeron, D., Coulombe, B. Robert. (2006) **Genome-wide computational prediction of transcriptional regulatory modules reveals new insights into human gene expression.** *Genome Research* 16(5):656-68.
- [18j] *Chindelevitch, L.*, *Li, Z.* *Blais, E.*, and Blanchette, M.(2006) **On the inference of parsimonious indel scenarios.** *Journal of Bioinformatics and Computational Biology*, 4(3):721-44.
- [17j] *Fang, F.* and Blanchette, M.(2006) **FootPrinter3: Phylogenetic footprinting in partially alignable sequences.** *Nucleic Acids Research*, 34:W617-W620.
- [16j] Ma, J., Zhang, L., Suh, B., Raney, B., Burhans, R., Kent, W.J., Blanchette, M., Haussler, D., and Miller, W. (2006) **Reconstructing contiguous regions of an ancestral genome.** *Genome Research*, 16(12): 1557-65.
- [15j] Warr, M., *Acoca, S.*, Liu, Z., Germain, M., Watson, M., Wing, S., Blanchette, M. and Shore, G. (2005) **BH3-ligand regulates access of MCL-1 to its E3 ligase.** *FEBS Letters* 579(25):5603-8.
- [14j] Guillemette, B., Bataille, A.R., Gévry, N., Adam, M., Blanchette, M., Robert, F., Gaudreau, L. (2005) **Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning.** *PLoS Biology* 3(12): 2100-10
- [13j] Sinha, S., Blanchette, M., Tompa, M. (2004) **PhyME: A probabilistic algorithm for finding motifs in sets of orthologous sequences.** *BMC Bioinformatics*, 5:170.
- [12j] Blanchette, M., Green, E., Miller, W., Haussler, D. (2004) **Reconstructing large regions of an ancestral mammalian genome in silico.** *Genome Research* 14(12):2412-23.
- [11j] Shnyreva, M., Weaver ,W.M., Blanchette, M., Taylor, S.L., Tompa, M., Fitzpatrick, D.R., Wilson, C.B. (2004). **Evolutionarily conserved sequence elements that positively regulate IFN-gamma expression in T cells.** *Proc Natl Acad Sci U S A.* 101(34):12622-7.
- [10j] Blanchette, M., Kent, J. Riemer, C., Elnitski, L., Smit, A.F.A, Roskin, K., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E., Haussler, D., Miller, W. (2004) **Aligning multiple genomic sequences with the threaded blockset aligner.** *Genome Research*:14(4):708-15.
- [9j] Margulies, E., Blanchette, M., NISC Comparative Sequencing Program, Haussler, D., Green, E. (2003) **Identification and characterization of multi-species conserved sequences.** *Genome Research*, 13(12), 2507-18.

- [8j] Thomas JW, Touchman JW, Blakesley RW, Bouffard GG, Beckstrom-Sternberg SM, Margulies EH, Blanchette M, Siepel AC, Thomas PJ, McDowell JC, Maskeri B, Hansen NF, Schwartz MS, Weber RJ, Kent WJ, Karolchik D, Bruen TC, Bevan R, Cutler DJ, Schwartz S, Elnitski L, Idol JR, Prasad AB, Lee-Lin SQ, Maduro VV, Summers TJ, Portnoy ME, Dietrich NL, Akhter N, Ayele K, Benjamin B, Cariaga K, Brinkley CP, Brooks SY, Granite S, Guan X, Gupta J, Haghighi P, Ho SL, Huang MC, Karlins E, Laric PL, Legaspi R, Lim MJ, Maduro QL, Masiello CA, Mastrian SD, McCloskey JC, Pearson R, Stantripop S, Tiongson EE, Tran JT, Tsurgeon C, Vogt JL, Walker MA, Wetherby KD, Wiggins LS, Young AC, Zhang LH, Osoegawa K, Zhu B, Zhao B, Shu CL, De Jong PJ, Lawrence CE, Smit AF, Chakravarti A, Haussler D, Green P, Miller W, Green ED. (2003) **Multi-species sequencing of targeted genomic regions: Comparative studies of genome structure, function, and evolution.** *Nature*, 424(6950):788-793.
- [7j] Blanchette, M. and Tompa, M. (2003) **FootPrinter: a program designed for phylogenetic footprinting.** *Nucleic Acids Research*. Jul 1;31(13):3840-2.
- [6j] Blanchette, M. and Tompa, M. (2002) **Discovery of regulatory elements by a computational method for phylogenetic footprinting.** *Genome Research*, 12(5), 739-748.
- [5j] Blanchette, M., Schwikowski, B., and Tompa, M. (2002) **Algorithms for phylogenetic footprinting.** *Journal of Computational Biology*, vol. 9, no. 2, 211-223. [Extended version of (10c) and (7c)]
- [4j] Blanchette, M., Kunisawa, T. and Sankoff, D. (1999). **Gene order breakpoint evidence in animal mitochondrial phylogeny.** *Journal of Molecular Evolution* 49, pp. 193-203.
- [3j] Sankoff, D. and Blanchette, M. (1999). **Phylogenetic invariants for Genome Rearrangement.** *Journal of Computational Biology* 6, 431-445.
- [2j] Sankoff, D. and Blanchette, M. (1998). **Multiple genome rearrangement and breakpoint phylogeny.** *Journal of Computational Biology* 5, 555-570. [Extended version of (3c)]
- [1j] Blanchette, M., Kunisawa, T. and Sankoff, D.(1996). **Parametric genome rearrangement.** *Gene* 172, GC pp. 11-17.

Peer-reviewed conference papers

- [28c] Waldispuhl J, Blanchette M. (2014) **Phylo and Open-Phylo: A Human-Computing Platform for Comparative Genomics** . In *Proceedings of the Second AAAI Conference on Human Computation and Crowdsourcing*.

- [27c] Hickey G. and Blanchette M. (2011) **A Probabilistic Model For Sequence Alignment with Context-Sensitive Indels**. In *Proceedings of the Annual International Conference on Research in Computational Molecular Biology (RECOMB) 2011*.
- [26c] Hickey G. and Blanchette M. (2010) **A Practical Algorithm for Estimation of the Maximum Likelihood Ancestral Reconstruction Expected Error**. In *Proceedings Pacific Symposium on Biotechnologies*. pp. 31-42.
- [25c] Lavallée-Adam M, Coulombe B, Blanchette M. (2009) **Detection of locally over-represented GO terms in protein-protein interaction networks**. In *Proceedings of the Annual International Conference on Research in Computational Molecular Biology (RECOMB)*. pp. 302-320
- [24c] Badescu D, Diallo A, Blanchette M and Makarenkov V. (2008) **Studying the Evolution of Human Papillomavirus Genomes**. In *Proceedings of RECOMB Comparative Genomics Satellite Conference*. pp 128 – 142
- [23c] Bertrand D, Blanchette M and El-Mabrouk N. (2008) **A phylogenetic Approach for Gene Mapping**. In *Proceedings of RECOMB Comparative Genomics Satellite Conference*. pp 198 – 210.
- [22c] Chen, X., and Blanchette, M. (2006) **Prediction of tissue-specific cis-regulatory modules using bayesian networks and probability trees**. In *Proceedings of the NIPS workshop on new problems and methods in computational biology*.
- [21c] Diallo, A.B., Makarenkov, V., Blanchette, M., and Lapointe, F.J. (2006) **A new efficient method for assessing missing nucleotides in DNA sequences in the framework of a generic evolutionary model**. In *Proceedings of IFCS 2006 Conference Data Science and Classification*, Springer Verlag, pp. 333-341.
- [20c] Diallo, A.B., Makarenkov, V., and Blanchette, M. (2006) **Finding maximum likelihood indel scenarios**. In *Proceeding of the fourth RECOMB satellite conference on Comparative Genomics*, pp. 171-185.
- [19c] Blais, E. and Blanchette, M. (2006) **Common substrings in random strings**. In *Proceedings of Combinatorial Pattern Matching*, pp. 129-140.
- [18c] Gaul, E. and Blanchette, M. (2006) **Ordering partially assembled genomes using gene arrangements**. In *Proceedings of the Fourth RECOMB Satellite conference on Comparative Genomics*, pp. 113-128.
- [17c] Blin, G., Blais, E., Guillon, P., Blanchette, M., and El-Mabrouk, N. (2006) **Inferring gene orders from gene maps using the breakpoint distance**. In *Proceedings of the Fourth RECOMB Satellite conference on Comparative Genomics*, pp 99-112.

- [16c] Bejerano, G., Haussler, D. and Blanchette, M. (2004) **Into the heart of darkness: large scale clustering of human non-coding DNA.** In *proceedings of ISMB 2004. Bioinformatics:20 Suppl 1:140-148.*
- [15c] Bergeron, A., Blanchette, M., Chateau, A., Chauve, C. (2004) **Intervalles conservés et Neighbor-Joining.** *Rencontres de la Société Francophone de Classification 2004.*
- [14c] Bergeron, A., Blanchette, M., Chateau, A., Chauve, C. (2004) **Reconstructing ancestral gene orders using conserved intervals.** *Workshop on Algorithms for Bioinformatics 2004*, pp. 14-25.
- [13c] Prakash, A., Blanchette, M., Sinha, S. and Tompa, M. (2004) **Motif discovery in heterogeneous sequence data.** *Pacific Symposium on Biocomputing 2004*, pp. 348-359.
- [12c] Blanchette, M., Kwong, S., Tompa, M. (2003) **An empirical comparison of tools for phylogenetic footprinting.** In *Proceedings of Bioinformatics and BioEngineering (BIBE)*, pp. 69-78.
- [11c] Blanchette, M. (2003). **A comparative analysis method for detecting binding sites in coding regions.** In *Proceedings of the Seventh Annual International Conference on Computational Molecular Biology (RECOMB 2003)*, 57-66.
- [10c] Blanchette, M. (2001). **Algorithms for phylogenetic footprinting.** In *Proceedings of the Fifth Annual International Conference on Computational Molecular Biology (RECOMB 2001)*, ACM Press, pp. 49-58. (Best student paper award).
- [9c] Blanchette, M. and Sinha, S. (2001). **Separating real motifs from their artifacts.** In *Proceedings of the Ninth International Conference on Intelligent Systems for Molecular Biology (ISMB 2001)*, *Bioinformatics.17*: S30-S38.
- [8c] Blanchette, M. (2001). **Evolutionary puzzles: An introduction to genome rearrangement.** In *Proceedings of the International Conference on Computational Science (ICCS 2001)*, Springer-Verlag, pp. 1003-1001.
- [7c] Blanchette, M., Schwikowski, B. and Tompa, M. (2000). **An exact algorithm to identify motifs in orthologous sequences from multiple species.** In *Proceedings of the Eighth International Conference on Intelligent Systems for Molecular Biology (ISMB 2000)*, AAAI Press, pp. 37-45.
- [6c] Sankoff, D. and Blanchette, M. (1999). **Probability models for genome rearrangement and linear invariants for phylogenetic inference.** In *Proceedings of the Third Annual International Conference on Computational Molecular Biology (RECOMB 99)*, ACM Press, pp. 302-309.
- [5c] Sankoff, D., and Blanchette, M. (1999). **Phylogenetic invariants for metazoan mitochondrial genome evolution.** In *Proceedings the Ninth Genome Informatics*

Conference (GIW 98), Universal Academy Press, pp. 22-23.

- [4c] Sankoff, D. and Blanchette, M. (1999). **Comparative genomics via phylogenetic invariants for Jukes-Cantor semigroups.** In *Proceedings of the International Conference on Stochastic Models Conference*, Canadian Mathematical Society Press.
- [3c] Sankoff, D. and Blanchette, M. (1998). **Multiple genome rearrangement.** In *Proceedings of the Second Annual International Conference on Computational Molecular Biology (RECOMB 98)*, ACM Press, pp. 243-247.
- [2c] Blanchette, M., Bourque, G. and Sankoff, D. (1997). **Breakpoint phylogenies.** In *Proceedings of the Eighth Genome Informatics Conference (GIW 1997)*, Universal Academy Press, pp.5-34.
- [1c] Sankoff, D. and Blanchette, M. (1997). **The median problem for breakpoints in comparative genomics.** In *Proceedings of the Third Annual Conference on Computing and Combinatorics (COCOON '97)*, Lecture Notes in Computer Science 1276, Springer Verlag, pp 251-263.

Conferences with abstract submission

- [5a] Blanchette M. (2014). **An Atlas of over 90,000 Conserved Noncoding Sequences Provides Insight into Crucifer Regulatory Regions.** *Plant and Animal Genome XXII Conference.*
- [4a] Blanchette, M., Giguère, V., Bergeron, D., Coulombe, B., and Robert, F. (2006) **Genome-wide detection of human cis-regulatory modules.** Cold Spring Harbor Symposium on Systems Biology.
- [3a] Bejerano, G., Mahadevan, S., Ratitch, B., *Blais, E.*, Precup, D., Haussler, D., Blanchette, M., (2004) **Similarity measures and clustering for regulatory modules.** *Cold Spring Harbor Labs Symposium on Comparative Genomics.*
- [2a] Bejerano, G., Haussler, D., Blanchette, M. (2005) **Computational analysis of the regulatory potential of non-coding regions conserved in vertebrates.** *Cold Spring Harbor Labs Symposium on Systems Biology (Transcriptional regulation).*
- [1a] Blanchette, M., Margulies, E., Green, E., Miller, W., and Haussler, W. (2002) **Computational detection of conserved regions in vertebrate genomes.** *AGBT 2002*, Marco Island, Florida.

Book chapters, reviews, and non-refereed publications

- [6b] *Diallo A.B.*, Blanchette M., Badunescu D., Makarenkov, V. (2011) **Classification of the human papilloma viruses.** In *Classification and Multivariate Analysis for Complex*

- Data Structures, Fichet, Picolo, Verde, Richi editors.* Springer-Verlag, Berlin. 457-465.
- [5b] Coulombe B, Blanchette M, Jeronimo C. (2008) **Steps towards a repertoire of comprehensive maps of human protein interaction networks: the Human Proteotheque Initiative (HuPI).** *Biochem Cell Biol.* 2008 Apr;86(2):149-56.
- [4b] Blanchette, M. (2007) **Substring parsimony.** Chapter in *Encyclopedia of Algorithms*, Ming-Yang Kao editor, in press.
- [3b] Blanchette, M. (2007) **Computation and analysis of genomic multi-sequence alignments.** *Annual Reviews of Genomics and Human Genetics 2007.* [22 citations]
- [2b] Philippe, H. and Blanchette, M. (2007) **Overview of the first phylogenomics conference.** *BMC Evol Biol.* 8;7 Suppl 1:S1.
- [1b] Blanchette, M., Diallo, A.B., Green, E., Miller, W., and Haussler, D. (2007) **Computational reconstruction of ancestral DNA sequences.** Chapter in *Handbook of Phylogenomics*, Humana Press. [9 citations]

Invited Talks

Keynote addresses

- “Studying the function of genomic regions using machine learning and comparative genomics”
- Pattern Recognition Iranian Conference (PRIA 2013), Iran (by teleconferencing), March 2013.
- “Ancestral Mammalian Genome Reconstruction and its Uses toward Annotating the Human Genome”
- Genome Quebec Bioinformatics Symposium, Montreal, December 2011.
 - JOBIM conference, Paris, France, June 2011.
 - VanBUG seminar, Vancouver, Canada, January 2011.
- “Whole-genome comparative and regulatory genomics”
- Chris Overton address, ISMB 2006, Fortaleza, Brazil, August 2006.
- “Detection and characterization of non-coding conserved regions in the human genome”
- JOBIM conference, Montreal, Canada, June 2004.
- “A Visit to the Mammalian Zoo: learning about our genome and that of our ancestors”
- CompBioNets 2004, Recife, Brazil, December 2004.

Other invited lectures

- “Deconvoluting the methylome”
- Barbados workshop on epigenetics and gene regulation, Barbados, January 2015.
- “An Atlas of over 90,000 Conserved Non-Coding Sequences Yields a Detailed Map of Crucifer Regulatory Regions.”
- Plant and Animal Genomes, San Diego, January 2014
- “Identification of ancient transposable elements using inferred ancestral mammalian genomes”
- Barbados Workshop on transposable element annotation, April 2014
- “Using Genome 10K data for ancestral genome inference”
- Genome 10K meeting, Fort Lauderdale, April 2013
- “Inference and analysis of protein interaction networks from AP-MS data”
- IMA Annual Program Year Workshop: Network Links: Connecting Social, Communication and Biological Network Analysis, Minneapolis, February 2012
- “Ancestral Mammalian Genome Reconstruction and its Uses toward Annotating the Human Genome”
- Cornell University Bioinformatics Symposium, March 2014.
 - Janelia Farm Conference on Probabilistic Models in Bioinformatics, Washington, March 2013
 - U. of Southern California Bioinformatics Symposium, March 2013
 - Barbados Workshop on Gene Regulation and Evolution, Barbados, April 2013
 - Genome Informatics 2012, Cambridge, UK.
 - McGill-Toronto Bioinformatics retreat 2012.
 - X Genomics Symposium CRG Barcelona 2011
 - New Zealand Annual Phylogenetic Conference, February 2011, Leigh, New Zealand.

- Mathématiques, Evolution et Genomes, June 2011, Université de Provence, Marseille, France.
 - Bioinformatics seminar, Laboratoire d'informatique, robotique, et microelectronique de Montpellier, April 2011.
 - X CRG Annual Symposium "Computational Biology of Molecular Sequences", Centre for Regulatory Genomics, Barcelona, Spain, October 2011
- “Ancestral mammalian genomes reconstruction and its applications “
- Evolutionary and Ecological Functional Genomics Gordon Conference, Newport, RI 2008
- “Whole-genome comparative and regulatory genomics”
- Université Laval, March 2008
 - Duke University, March 2008
 - Université de Sherbrooke, Département de Biologie, May 2007.
 - Cold Spring Harbor Laboratories, Bioinformatics Seminar, March 2007.
 - Cornell University, Physiology, Biophysics and Systems Biology seminar, January 2007.
 - NRC - Biotechnology Research Institute Seminar, November 2006.
 - University of California in Santa Cruz, September 2006.
 - McGill Centre for Bioinformatics seminar, September 2006.
 - Lawrence Berkeley National Labs, August 2006.
- “Detection and evolution of regulatory modules in Medaka fish”
- Bellairs Research Institute, April 2007.
- “Human cis-regulatory module prediction and analysis”
- First IRCM Meeting on Systems Biology, March 2007.
- “Genome-wide detection of human cis-regulatory modules”
- Cold Spring Harbor System Biology Symposium, March 2006.
 - INSERM workshop on Identification of non-coding functional regions in genomes, La Londe les Maures, France, April 2006.
 - Fourth Bertinoro Computational Biology Meeting, Italy, June 2006.
 - Penn State 24th summer symposium in molecular biology, July 2005.
 - Pennsylvania State University, Department of Biology, November 2005.
- “Reconstructing large regions of an ancestral mammalian genome in silico”
- Lausanne Genomics Days, EPFL (Switzerland), October 2006.
 - University of California in Merced, August 2006.
 - Université de Montréal, February 2006.
 - University of Vermont, February 2006.
 - Boston University Bioinformatics Seminar, February 2005.
 - RECOMB satellite meeting on comparative genomics, Bertinoro, Italy, October 2004.
 - Bioinformatics Seminar, Boston University, February 2005.
 - Génome Québec meeting on evolution, Bromont, Québec, July 2004.
- “Annotation of non-coding functional regions of the human genome”
- Institut Suisse de Recherche Expérimentale contre le Cancer (ISREC), Lausanne, October 2006.
- “Bioinformatics, genomics, and why they matter to Panama”
- Technological University of Panama, September 2006.
- “The dark matter of the human genome.”

- Medical Genetics Seminar Series, University of Toronto, April 2005.
- “New algorithms for the detection of functional regions in DNA sequences”
 - 24th Summer Symposium, The Pennsylvania State University, July 2005.
- “Detection and characterization of non-coding conserved regions in the human genome”
 - McGill Centre for nonlinear dynamics, McGill University, October 2003.
 - Bioinformatics seminar, Université de Montréal, January 2004.
 - Applied mathematics seminar, McGill University, February 2004.
 - Bioinformatics seminar, UQAM, December 2003.
- “Computational prediction of regulatory modules”
 - Congrès de l’ACFAS 2004, Montréal, July 2004.
- “Reconstructing ancestral gene orders using conserved intervals”
 - Mini-symposium at the SIAM conference on applied mathematics, Nashville, USA, 2004.
- “Algorithms for phylogenetic footprinting”
 - Bioinformatics Workshop, Lyon, France, April 2002.
 - Dept. of Computer Science, University of California in San Diego, January 2002.
 - Dept. of Computer Science, University of Toronto, February 2002.
 - Dept. of Computer Science, University of Waterloo, February 2002.
- “Computational detection of conserved regions in vertebrate genomes”
 - AGBT 2002, Marco Island, Florida, USA, February 2002.

Awards and Scholarships

- 2011-2013** Richard Tomlinson Science award (50,000)
- 2012** CACS/AIC Outstanding Young Computer Scientist Researcher Prize
- 2008** Leo Yaffe teaching award (Faculty of Science)
- 2007** Alfred P. Sloan Fellowship, 2007-2009
- 2006** Chris Overton award from the International Society of Computational Biology.
- 2003** FQRNT Strategic Researcher
- 2002** William Chan Memorial Award for best doctoral thesis, University of Washington, 2002.
- 2001** Best student paper award, RECOMB 2001, for "Algorithms for Phylogenetic Footprinting".
- 1997-2001** Natural Sciences and Engineering Research Council of Canada Scholarship for Masters and Ph.D. studies.
- 2001-2002** Fonds pour la formation de chercheurs et l'aide à la recherche du Québec scholarship for Ph.D. studies.

Editorial boards

- Frontiers in Computational Biology (2012 - now)
- Genome Research (2007-2009)
- Journal for Algorithms and Molecular Biology (2009 – now)

Program committee work

Conference	'04	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14
RECOMB (Research in Computational Molecular Biology)			X	X	X	X		X			
ISMB (Intelligent Systems for Molecular Biology)		X	X	X	X	X	X	X	X	X	X
ECCB (European Conference on Computational Biology)		X		X	X		X		X		X
WABI (Workshop on Algorithms for Bioinformatics)		X		X			X		X		
RECOMB Satellite conference on gene regulation	X	X	X			X			X		
RECOMB Satellite conference on comparative genomics		X	X	X	X	X	X	X			
NIPS workshop on new problems and methods in comp. biology			X	X		X	X				
ACM Conf. on Bioinformatics, Comp. Biol. and Biomed. Inform.										X	
Asia Pacific Bioinformatics Conference										X	

Conference and workshop organization

Date	Conference (co-organizer(s))	Location	Participants
April 2014	Barbados workshop on Transposable Element Annotation (with T. Bureau)	Bellairs Research Institute, Barbados	20
Sept 2013	Workshop on Mathematics of Sequence Evolution: Biological Models and Applications (with H. Philippe)	Montreal	80
April 2013	Barbados Workshop on Computational and Evolutionary and Gene Regulation (with Saurabh Sinha)	Bellairs Research Institute, Barbados	30
April 2012	Barbados Workshop on Control in Biological Systems	Bellairs Research Institute, Barbados	30
April 2010	6 th Barbados Workshop: Chromatin structure (with Josée Dostie)	Bellairs Research Institute, Barbados	27
April 2009	5 th Barbados Workshop: Protein-protein interaction networks (with Benoit Coulombe)	Bellairs Research Institute, Barbados	23
April 2008	4 th Barbados Workshop on computational gene regulation: genetic polymorphism (with Tomi Pastinen and Jacek Majewski)	Bellairs Research Institute, Barbados	18
April 2007	3 rd Barbados Workshop on computational gene regulation: Development (with Gill Bejerano)	Bellairs Research Institute, Barbados	19

Summary of Teaching Activities

Graduate courses

W2006, W2007; F2012; W2014	COMP 680: Mining Biological Sequences
W2006, W2007, W2008:	COMP 618: Functional Genomics (with M. Hallett)
F2007, F2008, F2009, F2012:	COMP 561: Computational Biology Meth. and Res.
W2004, W2005	COMP 766: Special Topics in Computer Science: Mining Biological Sequences

Undergraduate courses

F2003-2009, F2013, W2012-2013:	COMP 250: Introduction to Computer Science
F2007-2009, 2012:	COMP 462: Computational Biology Meth. (taught jointly with COMP 561)

Student Supervision

Post-doctoral students

Name	Period	Co-supervisor	Project title	Current position
Glenn Hickey	2013- now		PIATEA – Integrative annotation of transposable elements	Ongoing
Douglas Hoen	2013- now	Thomas Bureau	PIATEA – Integrative annotation of transposable elements	Ongoing
Krister Swenson	2012- 2014	Nadia El-Mabrouk	Algorithms for inference of biologically-realistic genome rearrangement scenarios	Assistant professor, Université Montpellier 2, France
Javad Sadri	2008 – 2011		Machine learning approaches to the prediction of current selective pressure	Assistant professor Birjand University, Iran
Rose Hoberman	2007 – 2009	Ken Dewar (McGill)	Impact of human polymorphisms on gene regulation	Project coordinator, Max Planck Institute
Pierre-Étienne Jacques	2006 – 2006	François Robert (IRCM)	Chromatin structure and histone variants	Assistant professor, Sherbrooke University
Annie Château	2005 – 2006	Anne Bergeron (UQAM)	Inference of ancestral gene orders	Professor, LIRMM, Montpellier, France

Ph.D. students

Name	Period	Co-supervisor	Project title	Current position
Ayrin Ahia-Tabibi	2015-now		Detection of ancient evolution events	In progress
Christopher Cameron	2013-now	Josée Dostie	Modeling of chromatin 3D structure	In progress
Rola Dali	2014-now		Chromatin conformation in cancer	In progress
Mickael Leclercq	2011- now	Abdoulate Baniré Diallo (UQAM)	Prediction of plant small RNA genes	In progress
Pablo Cingolani	2011 - now	Rob Sladek	Computational analysis of population genetics data for diabetes studies	In progress
James Wagner	2008 – 2014	Tomi Pastinen	Hidden Markov Models for Allelic Imbalance	Epiq (medical software)
Mathieu Lavallée-Adam	2008 – 2013	Benoit Coulombe	Computational annotation of protein-protein interaction networks	Postdoc, Scripps Institute
Mathieu Rousseau	2008 – 2013	Josée Dostie	Computational analysis of 5C chromosome confo.	Medical School student, McGill.
Ethan Kim	2007 - 2012	Adrian Vetta	Approximation algo. for protein-protein interaction network analysis	Google
Glenn Hickey	2008 – 2011		Ancestral sequence reconstruction	Postdoc, Haussler Lab, UCSC
Abdoulaye Baniré Diallo	2005 – 2009	Vladimir Makarenkov (UQAM)	Maximum likelihood indel scenarios and their applications	Asst. prof at Université du Québec à Montréal
Emmanuel Mongin	2005 – 2009	Ken Dewar	Prediction and analysis of human cis-regulatory modules	Science policy, Canadian government.

Masters students

Name	Period	Co-supervisor	Thesis title	Current position
Faizy Ahsan	2014-now	Doina Precup	Predicting enhancer cell-type specificity	In progress
Jimmy Hsin-Chia Chao	2014-now		Machine learning approaches to identify evolution signatures of regulatory regions	In progress

Ayrin Ahia-Tabibi	2013 - 2014		Identification of ancient transposable elements	PhD student
Masoud Karzand	2011 - 2013	Guillaume Bourque	Read mapping in repeat-rich genomes	Montreal Software Company
Isabelle Roy	2010		Local enrichment of transcription factor binding sites along the human genome	On medical leave since 2011
Victor Parmar	2007 - 2009		Binding site turnover	Goldman Sachs
Javier Sanchez Galan	2007 - 2010		Fast computation of transcription factor binding site over-representation on a genome-wide scale	Ph.D.student, Exp. med, McGill
Azin Sayad	2005 - 2008		Use of chromatin structure for transcription factor binding site prediction.	Eli Lilly Pharmaceuticals, Singapore
Haisu Zheng	2006 - 2009		Upgrades to the FootPrinter package	Software engineer, Calgary
Pablo Cingolani	2006 - 2008	Michael Hallett	Minimal reports of gene ontology enrichment	Research assistant, McGill
Faiyaz Al Zamal	2006 - 2007	Theodore Perkins	Modeling the dynamics of Drosophila's gap gene system	Ph.D. student, McGill
Michael Mayhew	2006 - 2008		Analysis of coding regions under non-coding selection	Ph.D. student, Duke University
Martin Smith (U. of Montreal)	2005 - 2007	B. Papadopoulou (U. Laval) and G. Burger (U. of Montreal)	Bioinformatics analysis of post-transcriptional regulation in <i>Leishmania</i>	Bioinfo. Res. Officer, Garvan Inst. of Med. Res.
Ian Reid	2005 - 2007		Transcription factor binding site turnover in mammals	Retired
Louis-François Handfield	2006 - 2007		Sequence-based clustering of cis-regulatory modules	Ph.D. student, U. of Toronto
Eric Blais	2004 - 2006		Common substrings in random strings	Postdoc, MIT
Qian Yang	2004 - 2005		Algorithms for RNA multiple alignment and structure prediction	Software engineer in China
Fei Fang	2004 - 2005		Algorithms for phylogenetic footprinting in semi-aligned data	Software engineer, Nuance (speech recognition)
Xiaoyu Chen	2004 -		Prediction of regulatory	Researcher,

	2006		modules with sequence and expression data	Illumina
Hui Chen	2004 – 2005		A Comparative analysis method for predicting binding sites in coding regions	Software developer, HydroOne, Toronto
Stéphane Acoca (Biochemistry)	2003 – 2005	Gordon Shore	New bioinformatics approaches to protein homology detection, with applications to apoptosis-related proteins.	Ph.D. student, Biotechnology Research Institute

Undergraduate students

Name	Period	Context	Project title
Matthew Trisic	2014	Summer work-study	Allele-specific Hi-C data analysis
Willie Chang	2014	Volunteering	Deconvolution of Chip-Seq data
Zhaohui Su	2014	Summer work-study	Spatial constraints on transcription factor binding site pairs
Rory Bokser	2013	Summer work-study	Parallel de novo genome assembly
Arvand Malekzad	2012	COMP 396	Detection of carry-over events in mass spectrometry data
Ming Yang Zhou	2012, 2013	COMP 396	Computational inference of chromatin 3D structure from 5C data
Yang Zhou	2012	NSERC USRA	Comparison of mammalian and arthropod coding regions under noncoding selection
Gary Roumanis	2010	NSERC USRA (with J. Waldispuhl)	Human computing for multiple sequence alignment
Alex Kawrykow	2010	NSERC USRA (with J. Waldispuhl)	A divide-and-conquer architecture for distributed multiple sequence alignment
Yang Li	2009	NSERC USRA	Mapping genes in the Naked Mole Rat
Jesse Doherty	2007	NSERC USRA	Prediction of SNPs affecting RNA stability
Omar Abdalbaki	2006	COMP 396	Parallelization of the CRUNCS search algorithm
David Azose	2006	COMP 396	Using cluster computing to identify CRUNCS
Mathieu Lavallée-Adam	2007	COMP 396 and CIHR project	Analysis of human protein-protein interaction networks
Olivier Hébert	2004	COMP 400 (honors project)	Computational prediction of DNA-protein affinities.
Oana Sandu	2005	NSERC USRA	Ordering constraints in motif discovery.
Zhentaο Li	2005 – 2006	NSERC USRA	Computational complexity of indel reconstruction
Leonid	2004 –	NSERC USRA project	Reconstructing parsimonious indel

Chindelevitch	2006	+ continuation	scenarios
Andrew Carbonetto	2004	COMP 400 (honors project)	Generating realistic biological simulated sequences

Service to the McGill Community

University committees

1. Work group for the creation of a graduate program in Quantitative Life Sciences
2. Search committee for the new Dean of Graduate Studies (2006)
3. Search committee for the new Director of the Bellairs Institute (2006)

School of Computer Science committees

2012-now Head of Awards Committee

2005-2011 Head of Masters Committee

2003-now Member of the Bioinformatics Committee

2007-2010 Member of the Hiring Committee