

Michael M. Hoffman

101 College St PMCRT 11-311

Toronto, ON M5G 1L7

Work +1 416 581 7481

E-mail michael.hoffman@utoronto.ca

ORCID [0000-0002-4517-1562](https://orcid.org/0000-0002-4517-1562)

Current positions

Scientist

Princess Margaret Cancer Centre, Toronto, ON, 2013–present.

Assistant Professor

Department of Medical Biophysics, University of Toronto, Toronto, ON, 2014–present.

Department of Computer Science, University of Toronto, Toronto, ON, 2014–present.

Full Member, Graduate Faculty

School of Graduate Studies, University of Toronto, Toronto, ON, 2017–present.

Member

Collaborative Graduate Program in Genome Biology and Bioinformatics, University of Toronto, Toronto, ON, 2016–present.

Education

Senior Fellow

Department of Genome Sciences, University of Washington, Seattle, WA, 2008–2013.

- Mentor: Dr. William Stafford Noble, Professor
- National Human Genome Research Institute K99/R00 Pathway to Independence Award

Doctor of Philosophy, Bioinformatics

University of Cambridge (Trinity College), Cambridge, UK, 2003–2008.

- Adviser: Dr. Ewan Birney, Director, European Bioinformatics Institute
- Marshall Scholar
- National Science Foundation Graduate Research Fellow

Bachelor of Science, Biochemistry, with Highest Honors, Special Honors

Bachelor of Arts, Plan II Honors Liberal Arts, with Highest Honors

University of Texas at Austin, Austin, TX, 1998–2003.

- Adviser: Dr. Andrew D. Ellington, Professor
- College of Natural Sciences Commencement Speaker, Dean's Honored Graduate
- College of Liberal Arts Dean's Distinguished Graduate
- University Honors

Publications

Publication list also available on [Google Scholar Citations](#), [ImpactStory](#), and [PubMed](#).

Publications: journal

- J23. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, **Hoffman MM**, Xie W, Rosen GL, Lengerich BJ, Israeli J, Lanchantin J, Woloszynek S, Carpenter AE, Shrikumar A, Xu J, Cofer EM, Lavender CA, Turaga SC, Alexandari AM, Lu Z, Harris DJ, DeCaprio D, Qi Y, Kundaje A, Peng Y, Wiley LK, Segler Marwin HS, Boca SM, Swamidass SJ, Huang A, Gitter A, Greene CS. "Opportunities and obstacles for deep learning in biology and medicine." *J R Soc Interface*. 2018; 15:20170387.
- Discussed in news feature: Webb S. "Deep learning for biology." *Nature*. 2018 Feb 20; 554:555–7.
 - Blog summary: Perkel J. "TechBlog: 'Manubot' powers a crowdsourced 'deep-learning' review." Naturejobs Blog. <http://blogs.nature.com/naturejobs/2018/02/20/techblog-manubot-powers-a-crowdsourced-deep-learning-review/>
 - Recommended by Faculty of 1000.
- J22. **Denisko D, Hoffman MM**. "Classification and interaction in random forests." *Proc Natl Acad Sci U S A*. 2018; 115:1690–2.
- J21. **Chan RCW***, Libbrecht MW*, **Roberts EG**, Bilmes JA, Noble WS, **Hoffman MM**. "Segway 2.0: Gaussian mixture models and minibatch training." *Bioinformatics*. 2018; 34:669–71.
- J20. **Chicco D, Hoffman MM**. "Genome Informatics 2016." *Genome Biol*. 2017; 18:5.
- J19. **Karimzadeh M, Hoffman MM**. "Top considerations for creating bioinformatics software documentation." *Brief Bioinform*. 2017. In press.
- J18. Lundberg SM, Tu WB, Raught B, Penn LZ, **Hoffman MM**, Lee SI. "ChromNet: learning the human chromatin network from all ENCODE ChIP-seq data." *Genome Biol*. 2016; 17:82.
- Recommended by Faculty of 1000.
- J17. Franke B, Plante JF, Roscher R, Lee A, Smyth C, Hatefi A, Chen F, Gil E, Schwing A, Selvitella A, **Hoffman MM**, Grosse R, Hendricks D, Reid N. "Statistical inference, learning and models in big data." *Int Stat Rev*. 2016. 84:371–89.
- J16. **Viner C, Hoffman MM**. "Determining the epigenome using DNA alone." *Nat Methods*. 2015 Mar; 12:191–2.
- J15. Libbrecht MW, Ay F, **Hoffman MM**, Gilbert DM, Bilmes JA, Noble WS. "Joint annotation of chromatin state and chromatin conformation reveals relationships among domain types and identifies domains of cell type-specific expression." *Genome Res*. 2015; 25:544–57.
- Selected as one of top 11 papers of 2014–2015 at the Research in Computational Molecular Biology (RECOMB)/International Society for Computational Biology Conference on Regulatory and Systems Genomics 2015.
- J14. Church DM, Schneider VA, Steinberg KM, Schatz MC, Quinlan AR, Chin CS, Kitts PA, Aken B, Marth GT, **Hoffman MM**, Herrero J, Mendoza LZ, Durbin R, Flicek P. "Extending reference assembly models." *Genome Biol*. 2015; 16:13.
- J13. Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn K, Minoda A, Tolstorukov MY, Appert A, Parker SCJ, Gu T, Kundaje A, Riddle NC, Bishop E, Egelhofer TA, Hu SS, Alekseyenko AA, Rechtsteiner A, Asker D, Belsky JA, Bowman SK, Chen QB, Chen RA, Day DS, Dong Y, Dose AC, Duan X, Epstein CB, Ercan S, Feingold EA, Ferrari F, Garrigues JM, Gehlenborg N, Good PJ, Haseley P, He D, Herrmann M, **Hoffman MM**, Jeffers TE, Kharchenko PV, Kolasinska-Zwierz P, Kotwaliwale CV, Kumar N, Langley SA, Larschan EN, Latorre I, Libbrecht MW, Lin X, Park R, Pazin MJ, Pham HN, Plachetka A, Qin B, Schwartz YB, Shoshitaishvili N, Stempor P, Vielle A, Wang C, Whittle CM, Xue H, Kingston RE, Kim JH, Bernstein BE, Dernburg AF, Pirrotta V, Kuroda MI, Noble WS, Tullius TD, Kellis M, MacAlpine DM, Strome S, Elgin SCR, Liu XS, Lieb JD, Ahringer J, Karpen GH, Park PJ. "Comparative analysis

- of metazoan chromatin organization.” *Nature*. 2014 Aug 28; 512:449–52.
- News and views summary: Muerdter F, Stark A. *Nature*. 2014 Aug 28; 512:374–5.
 - Cited 146 times.
- J12. **Hoffman MM***, Ernst J*, Wilder SP, Kundaje A, Harris RS, Libbrecht M, Giardine B, **Ellenbogen PM**, Bilmes JA, Birney E, Hardison RC, Dunham I, Kellis M, Noble WS. “Integrative annotation of chromatin elements from ENCODE data.” *Nucleic Acids Res*. 2013 Jan; 41:827–41.
- Featured Article in *Nucleic Acids Research*.
 - Cited 260 times.
- J11. **ENCODE Project Consortium**. “An integrated Encyclopedia of DNA Elements in the human genome.” *Nature*. 2012 Sep 6; 489:57–74.
- Author contribution: Lead analyst.
 - Featured on cover of *Nature*.
 - Recommended by Faculty of 1000.
 - News and views summary: Ecker JR et al. *Nature*. 2012 Sep 6; 489:52–55.
 - Cited 5,866 times.
- J10. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein B, Bickel P, Brown B, Cayting P, Chen Y, Desalvo G, Epstein C, Fisher-Aylor KI, Euskirchen G, Gerstein M, Gertz J, Hartemink AJ, **Hoffman MM**, Iyer VR, Jung YL, Karmakar S, Kellis M, Kharchenko PV, Li Q, Liu T, Liu XS, Ma L, Milosavljevic A, Myers RM, Park PJ, Pazin MJ, Perry MD, Raha D, Reddy TE, Rozowsky J, Shores N, Sidow A, Slattery M, Stamatoyonopoulos JA, Tolstorukov MY, White KP, Xi S, Farnham PJ, Lieb JF, Wold BJ, Snyder M. “ChIP-seq guidelines and practices used by the ENCODE and modENCODE consortia.” *Genome Res*. 2012 Sep; 22:1813–31.
- Recommended by Faculty of 1000.
 - Cited 730 times.
- J9. **Hoffman MM**, **Buske OJ**, Wang J, Weng Z, Bilmes JA, Noble WS. “Unsupervised pattern discovery in human chromatin structure through genomic segmentation.” *Nat Methods*. 2012 Mar 18; 9:473–6.
- Cited 299 times.
- J8. **ENCODE Project Consortium**. “A user’s guide to the Encyclopedia of DNA Elements (ENCODE) functional genomic data.” *PLOS Biol*. 2011 Apr; 9:e1001046.
- Recommended by Faculty of 1000.
 - Cited 931 times.
- J7. **Buske OJ**, **Hoffman MM**, Ponts N, Le Roch KG, Noble WS. “Exploratory analysis of genomic segmentations with Segtools.” *BMC Bioinformatics*. 2011 Oct 26; 12:415.
- J6. **Hoffman MM**, **Buske OJ**, Noble WS. “The Genomdata format for storing large-scale functional genomics data.” *Bioinformatics*. 2010 Jun 1; 26:1458–9.
- J5. Chen X, **Hoffman MM**, Bilmes JA, Hesselberth JR, Noble WS. “A dynamic Bayesian network for identifying protein-binding footprints from single molecule-based sequencing data.” *Bioinformatics*. 2010 Jun 15; 26:i334–42.
- J4. **Hoffman MM**, Birney E. “An effective model for natural selection in promoters.” *Genome Res*. 2010 May; 20:685–92.
- J3. **Hoffman MM**, Birney E. “Estimating the neutral rate of nucleotide substitution using introns.” *Mol Biol Evol*. 2007 Feb; 24:522–31.
- J2. **International Chicken Genome Sequencing Consortium**. “Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution.” *Nature*. 2004 Dec 9; 432:695–716.

- Featured on cover of *Nature*.
 - News and views summary: Schmutz J, Grimwood J. *Nature*. 2004 Dec 9; 432:679–80.
 - Cited 1,914 times.
- J1. **Hoffman MM**, Khrapov MA, Cox JC, Yao J, Tong L, Ellington AD. “AANT: the Amino Acid-Nucleotide Interaction Database.” *Nucleic Acids Res*. 2004 Jan 1; 32:D174–81.
- Cited 114 times.

Boldface indicates Michael Hoffman or one of his trainees is an author or a consortium member.

Publications: conference proceedings

- C1. Libbrecht MW, **Hoffman MM**, Noble WS, Bilmes JA. “Entropic graph-based posterior regularization.” International Conference on Machine Learning. 2015. Lille, France.

Publications: submitted

- S7. **Karimzadeh M**, **Hoffman MM**. “Virtual ChIP-seq: Predicting transcription factor binding by learning from the transcriptome” 2018. Preprint: <https://doi.org/10.1101/168419>
- S6. **Karimzadeh M**, Ernst C, Kundaje A, **Hoffman MM**. “Umap and Bimap: quantifying genome and methylome mappability.” 2017. Preprint: <https://doi.org/10.1101/095463>
- S5. Libbrecht MW*, Rodriguez O*, Weng Z, Bilmes JA, **Hoffman MM**, Noble WS. “A unified encyclopedia of human functional DNA elements through fully automated annotation of 164 human cell types.” 2016. Preprint: <https://doi.org/10.1101/086025>
- S4. **Roberts EG**, **Mendez M**, **Viner C**, **Karimzadeh M**, **Chan R**, Ancar R, **Chicco D**, Hesselberth JR, Kundaje A, **Hoffman MM**. “Semi-automated genome annotation using epigenomic data and Segway.” 2016. Preprint: <https://doi.org/10.1101/080382>
- S3. **Sood AJ***, **Viner C***, **Hoffman MM**. “DNAmod: the DNA modification database.” 2016. Preprint: <https://doi.org/10.1101/071712>
- S2. Carpenter A, Birney E, MacManes M, **Hoffman MM**. “Scientist-centered design: a checklist for improving grant systems and increasing scientific productivity.” 2016.
- S1. **Viner C**, Johnson J, Walker N, Shi H, Sjöberg M, Adams DJ, Ferguson-Smith AC, Bailey TL, **Hoffman MM**. “Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.” 2016. Preprint: <https://doi.org/10.1101/043794>

Publications: theses

- D1. **Hoffman MM**. *Quantifying evolution and natural selection in vertebrate noncoding sequence*. PhD thesis, University of Cambridge. 2008 Jul 8.

Intellectual property

- Segway: a dynamic Bayesian network method for segmenting genomic data. <http://pmgenomics.ca/hoffmanlab/proj/segway/>. Canadian copyright registration number 1115448.
- Segtools: exploratory data analysis of genomic segmentations. <http://pmgenomics.ca/hoffmanlab/proj/segtools/>. Canadian copyright registration number 1115535.
- Genomedata: a format for storing large-scale functional genomics data. <http://pmgenomics.ca/hoffmanlab/proj/genomedata/>. Canadian copyright registration number 1115391.
- Sunflower: a model of transcription factor binding and evolution. Canadian copyright registration number 1115330.

Recognitions

Recognitions: honors

- Canadian Institutes of Health Research New Investigator, 2017–2022
- Ontario Early Researcher Award, 2016–2021.
- Bioinformatics.ca Bioinformatics Expert, 2014.
- *Genome Technology* Young Investigator, 2011.
- Phi Beta Kappa Award of Distinction, 2003.
- Phi Beta Kappa Induction, 2003.
- Cambridge Overseas Trust Honorary Scholar, 2003.
- College of Liberal Arts Junior Fellow, 2000–2003.
- College of Natural Sciences Dean's Scholar, 1998–2003.
- College of Liberal Arts Plan II Honors Program (2% of university), 1998–2003.
- College of Natural Sciences College Scholar 2000, 2001, 2002, 2003.

Recognitions: other scholarships and awards

- Canadian Cancer Society Research Institute Junior Investigator Grant Panel Travel Award, 2014.
- Trinity College Moore, Beale Sargent and Mitchell Fund, 2005.
- The University of Texas at Austin Junior Fellows Research Grant, 2003.
- The University of Texas at Austin Undergraduate Research Fellowship, 1999–2000, 2000–2001, 2002–2003.
- Dedman Merit Scholarship, 1998–2002.
- Dorothy B. Banks Charitable Trust Scholarship, 1999–2000.
- IBM Thomas J. Watson Memorial Scholarship, 1998–2002.
- National Merit Scholarship, 1998–2002.

Funding

Funding: lead principal investigator

- F15. "Virtual ChIP-seq". Nvidia GPU Grant. In-kind, USD 4999. February 2018.
- F11. "Comprehending epigenomic changes in gene dysregulation and cancer using machine learning." Canadian Institutes of Health Research New Investigator Salary Award. 5 years, direct costs CAD 300,000. March 2017–February 2022.
- F10. "The expanded epigenetic alphabet: transcription factor binding in methylated DNA and beyond." Ontario Ministry of Research, Innovation and Science. Ontario Research Fund: Early Researcher Award. 5 years, total costs CAD 140,000. April 2016–March 2021.
- F9. "Epigenetic DNA modifications as drivers of leukemia gene expression programs." Canadian Cancer Society Research Institute. Innovation Grant. Awarded: 3 years, direct costs CAD 200,000. Actual: 3 years, direct costs CAD 199,000 (sponsor-wide cut to continuing grants). August 2015–July 2018.
- F8. "Interpreting epigenetic DNA modifications in glioblastoma stem cells." McLaughlin Centre. Accelerator Grant in Genomic Medicine and Health Informatics. 1 year, direct costs CAD 65,000. May 2015–April 2016.
- F7. "Transcription factor recognition models with modified nucleobases." Natural Sciences and Engineering Research Council of Canada. Discovery Grant. 5 years, direct costs CAD 175,000. April 2015–March 2020.
- F2. "Pattern discovery for comparative epigenomics." National Institutes of Health/National Human Genome Research Institute. Pathway to Independence Award (Parent K99/R00). Awarded: 5 years,

total costs USD 966,069. Actual: 2 years, total costs USD 206,244 (no longer eligible for R00 portion after move to Canada). September 2011–October 2013.

- F1. Graduate research fellowship. National Science Foundation. 5 years, direct costs USD 142,872. September 2003–August 2008.

Funding: co-principal investigator

- F14. “Acute Leukemia Translational Research Initiative.” Ontario Institute for Cancer Research. 2 years, total costs \$5,000,001. April 2017–March 2019. Lead principal investigators: John E. Dick, Aaron D. Schimmer.
- F13. “Deciphering and manipulating cell-specific regulatory network to produce therapeutic designer cells.” Medicine by Design, University of Toronto. Team Project Award. 3 years, direct costs CAD 314,309. September 2016–August 2019. Lead principal investigator: Jason Fish.
- F12. “Regulatory network control of neural stem cells for endogenous repair.” Medicine by Design, University of Toronto. Team Project Award. 3 years, direct costs CAD 3,053,789. September 2016–August 2019. Lead principal investigator: Gary Bader.

Funding: co-investigator

- F6. “Cancer Stem Cell Program.” Ontario Institute for Cancer Research. 2 years, total costs CAD 5,436,364. April 2015–March 2017. Lead principal investigator: John E. Dick.
- F5. “Sequence variation and DNA methylation patterning.” Natural Sciences and Engineering Research Council of Canada. Discovery Grant. 5 years, direct costs CAD 160,000. Hoffman Lab: January 2015–January 2017. Lead principal investigator: Carl P. Ernst.

Presentations

Presentations: invited/keynote

- PA35. Reading between the genes: interpreting noncoding DNA in high throughput. Pacific Symposium on Biocomputing (PSB). Kohala Coast, HI. 3–7 January 2019
- PA34. Cold Spring Harbor Laboratory Meeting on Biological Data Sciences. Cold Spring Harbor, NY. 7–10 November 2018.
- PA33. Department of Computational Medicine and Bioinformatics Seminar Series. University of Michigan. Ann Arbor, MI. 24 October 2018.
- PA32. Fifth Canadian Conference on Epigenetics. Epigenetics Canada. Estérel, QC. 30 September–3 October 2018.
- PA31. “Virtual ChIP-seq: predicting transcription factor binding by learning from the transcriptome.” Symposium on Advances in Genomics, Epidemiology, and Statistics (SAGES). Philadelphia, PA. 1 June 2018.
- PA30. “Understanding epigenetics—and what it means for cancer.” Barbados Workshop on Cancer Epigenetics. Bellairs Research Institute. Holetown, Barbados. 8 January 2018.
- PA29. “Identifying transcription factor binding using open chromatin, transcriptome, and methylation data.” Research Program in Quantitative Sciences Seminar Series. Division of Biostatistics and Bioinformatics. Sidney Kimmel Comprehensive Cancer Center. Johns Hopkins University. Baltimore, MD. 14 December 2017.
- PA28. “Identifying transcription factor binding using open chromatin, transcriptome, and methylation data.”. Research Seminar Series. British Columbia Cancer Agency. Vancouver, BC. 2 October 2017.
- PA27. “Novel inferences from Hi-C data with protein-coding gene data.” Great Lakes Bioinformatics Confer-

- ence. Chicago, IL. 16–17 May 2017.
- PA26. “Computational predictive models and 3D genome organization.” Workshop on the Physical Basis of Functional Genome Organization. Bellairs Research Institute. Holetown, Barbados. 15 April 2017.
- PA25. “Automated genome annotation and an expanded epigenetic alphabet.” Microsoft Research New England. Cambridge, MA. 27 March 2017.
- PA24. “Automated genome annotation and an expanded epigenetic alphabet.” HudsonAlpha Institute for Biotechnology. Huntsville, AL. 8 March 2017.
- PA23. “Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.” Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting. Nashville, TN. 15 November 2016.
- PA22. “Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.” Cold Spring Harbor Laboratory Meeting on Biological Data Sciences. Cold Spring Harbor, NY. 27 October 2016.
- PA21. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Wellcome Trust Sanger Institute. Hinxton, England. 19 September 2016.
- PA20. “Semi-automated human genome annotation using chromatin data.” Intelligent Systems for Molecular Biology (ISMB). Orlando, FL. 12 July 2016.
- PA19. “Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.” Bioinformatics Italian Society (BITS) Annual Meeting. Salerno, Italy. 17 June 2016.
- PA18. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Early Career Investigators Meeting on Quantitative Problems in Human Genetics and Health. Banff International Research Station. Banff, AB, Canada. 11 January 2016.
- PA17. “Semi-automated human genome annotation using chromatin data.” Institute for Operations Research and the Management Sciences (INFORMS) Annual Meeting. Philadelphia, PA. 2 November 2015.
- PA16. “Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet.” Cold Spring Harbor Laboratory Meeting on Genome Informatics. Cold Spring Harbor, NY. 29 October 2015.
- PA15. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Department of Biology. New York University. New York, NY. 28 October 2015.
- PA14. “Transcription factor binding motifs in an expanded epigenetic alphabet.” Taiwan-Canada Joint Workshop on Epigenetics. Canadian Human and Statistical Genetics Meeting. Vancouver, BC. 21 April 2015.
- PA13. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Department of Biochemistry and Molecular Genetics. University of Colorado Anschutz Medical Campus. Denver, CO. 3 April 2015.
- PA12. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Bioinformatics Club. University of Waterloo. Waterloo, ON. 24 March 2015.
- PA11. “Semi-automated genome annotation and an expanded epigenetic alphabet.” Understanding non-coding DNA through intra and inter-species epigenomic variation. Bellairs Research Institute. Holetown, Barbados. 9 January 2015.
- PA10. “Semi-automated genome annotation and an expanded epigenetic alphabet.” McGill University. Montreal, QC. 21 November 2014.
- PA9. “Semi-automated annotation of the human genome using chromatin and RNA-seq data.” Major Discoveries in Biology. Undergraduate Program in Genomic Sciences. National Autonomous University of Mexico. Mexico City, Mexico (videoconference). 5 March 2014.
- PA8. “Semi-automated Genome Annotation with Segway.” Institute of Electrical and Electronics Engineers

International Workshop on Genomic Signal Processing and Statistics 2013. Houston, TX. 17 November 2013.

- PA7. "Unsupervised pattern discovery in human chromatin structure through genomic segmentation." Workshop on Epigenomics and Cell Function 2013. Association for Computing Machinery Conference on Bioinformatics, Computational Biology and Biomedicine. Washington, DC. 22 September 2013.
- PA6. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Genentech. South San Francisco, CA. 12 November 2012.
- PA5. "Unsupervised pattern discovery in human chromatin data." Computational Biology Center Guest Seminar. Memorial Sloan-Kettering Cancer Center. New York, New York. 18 May 2011.
- PA4. "Segway: finding patterns in chromatin data." Center for Bioinformatics Research Special Talk, Indiana University. Bloomington, Indiana. 18 November 2010.
- PA3. "Finding the patterns in chromatin data." Fifth Barbados Workshop on Gene Regulation: The role of chromatin in 3D structure. Holetown, Barbados. 21 April 2010.
- PA2. "Properties of natural selection in mammalian promoters." Weizmann UK Symposium: Biological complexity: from models to systems. London, England. Summer 2008.
- PA1. "Predicting Selection in Promoters by Simulating the Effects of Mutations." Waterman Seminar, Leibniz Institute of Plant Genetics and Crop Plant Research. Gatersleben, Germany. Autumn 2007.

Presentations: conference, selected from abstracts

- PC19. "Inference of transcription factor binding sites in new cell types from open chromatin and gene expression data." Topic-contributed session on statistical learning for epigenomics data. Joint Statistical Meetings. Vancouver, BC. 28 July–2 August 2018
- PC18. "Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet." Keystone Conference on DNA and RNA Methylation. Vancouver, BC. 22 January 2018
- PC17. "Modeling methyl-sensitive transcription factor motifs with an expanded epigenetic alphabet." Canadian Cancer Research Conference. Vancouver, BC. 7 November 2017
- PC16. "Transcription factor expression and its effects on binding site occupancy and motif preference." Cold Spring Harbor Laboratory/Wellcome Trust Conference on Genome Informatics. Hinxton, England. 19 September 2016.
- PC15. "Segway: semi-automated genome annotation." Bioinformatics Open Source Conference. Dublin, Ireland. 10 July 2015.
- PC14. "Transcription factor binding in an expanded epigenetic alphabet." World Congress on Medical Physics. International Union for Physical and Engineering Sciences. Toronto, ON. 11 June 2015.
- PC13. "Transcription factor binding in an expanded epigenetic alphabet." Keystone Symposium on DNA Methylation. Keystone Symposia on Molecular and Cellular Biology. Keystone, CO. 1 April 2015.
- PC12. "Transcription factor motif discovery and prediction in an expanded epigenetic alphabet." Cold Spring Harbor Laboratory/Wellcome Trust Conference on Genome Informatics. Cambridge, England. 22 September 2014.
- PC11. "One genome, many epigenomes, machine learning." Workshop on Big Data for Health, University of Toronto. Toronto, ON. 4 July 2014.
- PC10. "Semi-automated annotation of functional elements with whole genome RNA-seq." Machine Learning in Computational Biology 2012. Neural Information Processing Systems Conference 2012. Stateline, NV. 7 December 2012.
- PC9. "Functional genomics and machine learning." SynBioCon 2012. Austin, TX. 26 May 2012.
- PC8. "Semi-automated annotation of the human genome using whole-genome RNA-seq." ENCODE and

modENCODE AWG/PI Meeting. Cambridge, MA. 22 May 2012.

- PC7. "Discovery of functional elements using whole-genome RNA-seq." The Biology of Genomes. Cold Spring Harbor, NY. 9 May 2012.
- PC6. "Chromosomal features: large-scale integration." ENCODE AWG Workshop. Stanford, CA. 7 March 2011.
- PC5. "Large-scale integration group update." ENCODE AWG Workshop. Barcelona, Spain. 20 July 2010.
- PC4. "Large-scale behavior task group." ENCODE AWG Workshop. Rockville, MD. 8 March 2010.
- PC3. "Segway: a dynamic Bayesian network for genomic segmentation." ENCODE and modENCODE Consortia Meeting. Bethesda, MD. Spring 2009.
- PC2. "ENCODE and Segway." ENCODE AWG Workshop. Rockville, MD. December 2008.
- PC1. "Poly: rapid development of embarrassingly parallelizable applications." Bioinformatics Open Source Conference. Glasgow, Scotland. Summer 2004.

Presentations: invited panelist

- PE1. "Genomic computing challenges and perspectives." Intelligent Systems for Molecular Biology (ISMB). Orlando, FL. 12 July 2016.

Presentations: other teaching

- PG20. "Machine learning in epigenomics." Guest seminar for Epigenetics. Department of Medical Biophysics. University of Toronto. 2 February 2017.
- PG19. "Data integration and epigenomics." Advanced Sequencing Technologies and Applications. Cold Spring Harbor Laboratory. 13 November 2017.
- PG18. "Biophysics Seminar," MBP 1015Y. Department of Medical Biophysics. University of Toronto. 27 September 2017–2018.
- PG17. "Cancer epigenomics." Guest seminar for Quantitative Cancer Genomics. Department of Medical Biophysics. University of Toronto. 2 February 2017.
- PG16. "Semi-automated human genome annotation using chromatin data." Guest lecture for CSC 2417, Department of Computer Science, University of Toronto. 16 November 2016.
- PG15. "Data integration and epigenomics." Advanced Sequencing Technologies and Applications. Cold Spring Harbor Laboratory. 14 November 2016.
- PG14. "Biophysics Seminar," MBP 1015Y. Department of Medical Biophysics. University of Toronto. 2016–2017.
- PG13. "Epigenomics." Summer Student Program. Department of Medical Biophysics. University of Toronto. 30 May 2016.
- PG12. "Gene Regulation Network Analysis." High-throughput Biology: From Sequence to Networks. Canadian Bioinformatics Workshop. 15 June 2016.
- PG11. "Biophysics Seminar," MBP 1015Y. Department of Medical Biophysics. University of Toronto. 2015–2016.
- PG10. "Epigenomics." Summer Student Program. Department of Medical Biophysics. University of Toronto. 8 June 2015.
- PG9. "Gene Regulation Network Analysis." Pathway and Network Analysis of -omic Data. Canadian Bioinformatics Workshop. 3 June 2015.
- PG8. "Biophysics Seminar," MBP 1015Y. Department of Medical Biophysics. University of Toronto. 15 October 2014–2015.
- PG7. "Gene Regulation Network Analysis." High-throughput Biology: From Sequence to Networks. Cana-

dian Bioinformatics Workshop. 3 May 2015.

- PG6. "Transcription Factor Regulatory Analysis." Pathway and Network Analysis of omics Data. Canadian Bioinformatics Workshop, 4 June 2014.
- PG5. "Introduction to ENCODE" and "The Segway annotation of the ENCODE Data." Iowa Institute of Human Genetics Bioinformatics Short Course. Taught two one-hour lectures. University of Iowa. 30–31 July 2013.
- PG4. "Genes and gambling: using probability to make sense of biology." Guest lecture for undergraduate computer science course, Bellevue College. Winter 2010.
- PG3. "Introduction to Python." Taught two-day seminar. European Molecular Biology Laboratory. Autumn 2005.
- PG2. "EMBL Predoctoral Bioinformatics Workshop." Organizing committee member and teaching assistant for three-day course. European Molecular Biology Laboratory. Autumn 2004.
- PG1. "Biological sequence analysis." Discussion leader for eight-week seminar. The University of Texas at Austin Society for Computational Biology. Summer 2003.

Presentations: other seminars

- PH19. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Genetics. University of Pennsylvania. Philadelphia, PA. 25 July 2013.
- PH18. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Ontario Cancer Institute. Toronto, ON. 4 July 2013.
- PH17. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Institute for Human Genetics. University of California, San Francisco. San Francisco, CA. 21 March 2013.
- PH16. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Analytical and Translational Genetics Unit. Massachusetts General Hospital. 24 April 2013.
- PH15. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Electrical and Computer Engineering. University of Texas at Austin. 22 April 2013.
- PH14. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." The Jackson Laboratory. Bar Harbor, ME. 11 March 2013.
- PH13. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." McKusick-Nathans Institute of Genetic Medicine. Johns Hopkins University. Baltimore, MD. 7 March 2013.
- PH12. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Human Genetics. University of Utah. Salt Lake City, UT. 25 February 2013.
- PH11. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." National Eye Institute. National Institutes of Health. Bethesda, MD. 30 January 2013.
- PH10. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Molecular and Medical Genetics. Oregon Health and Science University. Portland, OR. 23 January 2013.
- PH9. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Human Genetics. Emory University. Atlanta, GA. 14 January 2013.
- PH8. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." McDermott Center for Human Growth and Development/Center for Human Genetics. University of Texas Southwestern Medical Center. Dallas, TX. 7 January 2013.
- PH7. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Earl Stadtman Symposium on Genetics, Genomics, and Systems Biology. Division of Intramural Research. National Institutes of Health. Bethesda, MD. 17 December 2012.

- PH6. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Program in Bioinformatics and Integrative Biology. University of Massachusetts Medical School. Worcester, MA. 10 December 2012.
- PH5. "Semi-automated annotation of the human genome using chromatin and RNA-seq data." Department of Pediatrics. University of California, San Diego. 4 June 2012.
- PH4. "Unsupervised and semi-supervised pattern discovery in human chromatin structure." Section of Genetic Medicine. University of Chicago. Chicago, IL. 23 April 2012.
- PH3. "Simultaneous segmentation of multiple functional genomics data sets with heterogeneous patterns of missing data." Department of Computer Science. Colorado State University. Fort Collins, CO. 10 April 2012.
- PH2. "Properties of natural selection in mammalian promoters." Seminar, Broad Institute of MIT and Harvard. Cambridge, MA. 9 April 2008.
- PH1. "Properties of natural selection in mammalian promoters." Seminar, Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School. Worcester, MA. 25 March 2008.

Professional activities

Professional activities: society leadership positions

- **Co-chair**, Ontario Region, Association of Canadian Early Career Health Researchers, 2016–present.
- **Secretary/Treasurer**, International Society for Computational Biology Student Council, 2004. **Member**, 2003–2004.

Professional activities: conference organizer

- International Conference on Epigenetics and Bioengineering (ICEB 2018). American Institute of Chemical Engineers. San Francisco, CA. 4–6 October 2018.
- Early Career Investigators Meeting on Quantitative Problems in Human Genetics and Health. Banff International Research Station. Banff, AB. 10–15 January 2016.

Professional activities: advisory

- **Member**, Reference group for the Chief Scientist, Ontario Ministry of Research, Innovation, and Science, 2017–present.
- **Invited Participant**, Ontario Ministry of Research and Innovation workshop: Ontario Open Access: Accelerating Science, 2016.
- **Invited Participant**, National Human Genome Research Institute planning workshop: Planning the future of genomics: foundational research and applications in genomic medicine, 2010.

Professional activities: funding review

- **Member**, College of Reviewers, Canadian Institutes of Health Research, 2017–present.
- **Member**, Genomics Peer Review Committee, Canadian Institutes of Health Research, 2017. **Teleconference member**, 2015.
- **Member**, Project Grant Competition Stage 2 Peer Review Cluster 13, Canadian Institutes for Health Research, 2017.
- **Member**, Project Grant Competition Stage 1 Peer Review Clusters 5/13/15/25, Canadian Institutes for Health Research, 2016–2017.
- **Member**, Program Committee, Idea DREAM Challenge, 2016–2017.
- **Member**, Institute Community Support Travel Award Peer Review Committee, Institute for Cancer Research, Canadian Institutes for Health Research, 2016.

- **Member**, Institute Community Support Travel Award for International Agency for Research on Cancer (IARC) Conference Peer Review Committee, Institute for Cancer Research, Canadian Institutes for Health Research, 2016.
- **Member**, Review Committee, Collaborative Personalized Cancer Medicine Team Grant, Princess Margaret Cancer Centre, 2015.
- *Ad hoc* external review: United Kingdom Medical Research Council, French Institute for Bioinformatics, Natural Sciences and Engineering Research Council of Canada, Canadian Statistical Sciences Institute.

Professional activities: manuscript review

- **Academic Editor**, PeerJ, 2017–present.
- **Affiliate**, bioRxiv, Cold Spring Harbor Laboratory Press, 2015–present.
- **Member**, Program Committee, Workshop on Machine Learning in Computational Biology (MLCB), Conference on Neural Information Processing Systems (NIPS), 2017.
- **Member**, Program Committee, Research in Computational Molecular Biology Satellite Workshop on Massively Parallel Sequencing (RECOMB-seq), 2016–2017.
- **Member**, Program Committee, International Joint Conference on Artificial Intelligence (IJCAI), 2015–2016.
- **Member**, Program Committee, Great Lakes Bioinformatics Conference/Canadian Computational Biology Conference (GLBIO/CCBC), 2015–2016.
- **Member**, Program Committee, Research in Computational Molecular Biology (RECOMB), 2014–2015.
- *Ad hoc* review for journals: Nature Methods, Genome Biology, Proceedings of the National Academy of Sciences of the United States of America, PLOS Computational Biology, Nucleic Acids Research, Scientific Reports, Bioinformatics, PLoS ONE, BMC Genomics, BMC Bioinformatics, Journal of Machine Learning Research, Database, IEEE Transactions on Computational Biology and Bioinformatics, Artificial Intelligence in Medicine, Computational and Structural Biotechnology Journal.
- *Ad hoc* review for conferences, with refereed proceedings: Intelligent Systems in Molecular Biology/European Conferences on Computational Biology (ISMB/ECCB), Research in Computational Molecular Biology (RECOMB).

Professional activities: consortium leadership positions

- **Chair**, Large-scale Integration Task Group, ENCODE Analysis Working Group, 2010–2012.

Professional activities: institutional leadership positions and committees

- **Co-chair**, Bioinformatics Subcommittee, Cancer Genomics Program, Princess Margaret Cancer Centre, 2013–present.
- **Member**, Research Space Committee, Princess Margaret Cancer Centre, 2016–present.
- **Member**, Director of Research Search Committee, Princess Margaret Cancer Centre, 2016–present.
- **Member**, Legacy Working Group, Medicine by Design, University of Toronto, 2017–present.
- **Co-organizer**, Princess Margaret Cancer Centre Research Retreat, 2015–2016.
- **Member**, Till and McCullough Paper of the Year Selection Committee, Princess Margaret Cancer Centre, 2016–2017.
- **Member**, Graduate Admissions Committee, Department of Computer Science, University of Toronto, 2014–2015.
- **Member**, Princess Margaret Genomics Centre Advisory Board, 2013–2015.
- **Chair**, Cambridge University Student Pugwash Society, 2006–2007. **Secretary**, 2007–2008.
- **Representative**, EMBL–European Bioinformatics Institute Predoctoral Fellow Association, 2004–2005.

- **President**, Texas Student Publications Board of Operating Trustees, 2002–2003. **Vice President**, 2001–2002. **Member**, 2000–2003.
- **Member**, College of Natural Sciences Dean’s Scholars Committee, 2000–2003.
- **Founder** and **Co-President**, University of Texas at Austin Barbecue Club, 1999–2003.
- **Chair**, Texas Student Publications Handbook Revision Committee, 2002.
- **Chair**, Texas Student Publications Election Committee, 2001–2002.
- **Member**, University of Texas at Austin Information Technology Coordinating Council, 1999–2002.
- **Member**, University of Texas at Austin General Faculty Admissions and Registration Committee, 1999–2000.

Professional activities: student committees and examinations

- **Supervisory Committee** for PhD students: Department of Electrical and Computer Engineering, University of Toronto (Michael Leung).
- **Supervisory Committee** for PhD students: Department of Computer Science, University of Toronto (Tae-Hyung Simon Kim).
- **Supervisory Committee** for MSc students: Department of Medical Biophysics, University of Toronto (Leslie Oldfield, James Hawley).
- **Examiner** for PhD students: Department of Medical Biophysics, University of Toronto (Sahar Ghanavati).
- **Examiner** for PhD students: Institute for Medical Science, University of Toronto (Sarah Gagliano).
- **Examiner** for MSc students: Department of Medical Biophysics, University of Toronto (Andrew Weatherbee).
- **Examiner** for MSc students: Department of Molecular Genetics, University of Toronto (Lina Antounians).
- **Reviewer** for MSc theses: Department of Computer Science, University of Toronto (Aryan Arbabi).
- **Host**, Gene Researcher for a Week, 2017.
- **Judge**, Poster Session, James Lepock Memorial Student Symposium, Department of Medical Biophysics, University of Toronto, 2016, 2017.
- **Judge**, Poster Session, Annual Symposium, Medicine by Design, University of Toronto, 2016.
- **Judge**, Biohackathon, Internationally Genetically Engineered Machine (iGEM) Club, University of Toronto, 2016.
- **Judge**, Summer Student Poster Day, Department of Medical Biophysics, University of Toronto, 2014–2016.
- **Mentor**, National Resource for Network Biology, Google Summer of Code, 2014.

Professional activities: memberships

- **Member**, Statistical Society of Canada, 2018.
- **Member**, Global Alliance for Genomics and Health, 2014–present.
- **Member**, Association for Computing Machinery Special Interest Group on Bioinformatics, 2013–2014.
- **Member**, National Postdoctoral Association, 2011–2013.
- **Member**, Phi Beta Kappa, 2003–present.
- **Member**, International Society for Computational Biology, 2003–2004, 2010–2011, 2015–2016.
- **Junior Member**, Isaac Newton Institute for Mathematical Sciences, 2006–2013.
- **Member**, American Association for the Advancement of Science, 2007–2008.
- **Member**, Genetics Society, 2006–2007.

Prior positions

Associate Member, Graduate Faculty

School of Graduate Studies, University of Toronto, Toronto, ON, 2014–2017.

Senior Fellow

Department of Genome Sciences, University of Washington, 2008–2013.

Predoctoral Fellow

EMBL–European Bioinformatics Institute, 2003–2008.

Research Assistant

Department of Chemistry and Biochemistry, The University of Texas at Austin, 1999–2003.

Undergraduate Research Program Participant

Cold Spring Harbor Laboratory, 2001.

Undergraduate Research Fellow

Plant Biochemistry Research Training Center, Washington State University, 2000.

Trainees and staff supervised

Trainees and staff supervised: current

- T29. **Liu, Yufang** (“Flora”). BSc Student, Department of Computer Science, University of Toronto, 2018–present.
- T28. **McNeil, Matthew**. BSc Student, Department of Biochemistry, University of Toronto, 2018–present.
- T27. **Xu, Winnie**. BSc Student, Department of Immunology, University of Toronto, 2018–present.
- T26. **Asenjo Ponce de León, Marc**. Bachelor’s Student, Barcelona School of Informatics, Universitat Politècnica de Catalunya, 2018–present.
- T24. **Cao, Chang**. BSc Student, Department of Statistical Sciences, University of Toronto, 2017–present.
- T19. **Nguyen, Francis**. MSc Student, Department of Medical Biophysics, University of Toronto, 2016–present.
- T17. **Denisko, Danielle**. MSc Student, Department of Physics, University of Toronto, 2016–present.
- T16. **Chan, Rachel**. BSc Student, Science Co-op Program, University of British Columbia, 2016–present.
- T15. **Mendez, Mickaël**. PhD Student, Department of Computer Science, University of Toronto. 2016–present.
- T13. **Sood, Ankur Jai**. BESC (Honors) Student, Mechatronic Systems Engineering, University of Western Ontario. 2015–present.
- T11. **Karimzadeh, Mehran R**. PhD Student, Department of Medical Biophysics, University of Toronto, 2015–present.
- T10. **Chicco, Davide**. Postdoctoral Fellow, University Health Network, 2014–present.
- T9. **Viner, Coby**. MSc Student, Department of Computer Science, University of Toronto, 2014–present.
- T8. **Roberts, Eric**. Bioinformatics Programmer, University Health Network, 2014–present.

Trainees and staff supervised: past

- T25. **Subasri, Vallijah**. MSc Student, Department of Medical Biophysics, University of Toronto, 2017–present.
- T23. **Smith, Ian**. MSc Student, Department of Medical Biophysics, University of Toronto, 2017.
Present position: MSc Student, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.
- T22. **Schonbach, Maya**. MSc Student, Department of Medical Biophysics, University of Toronto, 2016.
- T21. **Houlahan, Katie**. MSc Student, Department of Medical Biophysics, University of Toronto, 2016.
Present position: MSc Student, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.

- T20. **Hawley, James**. MSc Student, Department of Medical Biophysics, University of Toronto, 2016.
Present position: MSc Student, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.
- T18. **Bi, Haixin (“Sarah”)**. Post-BSc Summer Student, Princess Margaret Cancer Centre, 2016.
Present position: PhD Student, Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA, USA.
- T14. **Madani Tonekaboni, S. Ali**. PhD Student, Department of Medical Biophysics, University of Toronto, 2015.
Present position: PhD Student, Department of Medical Biophysics, University of Toronto, Toronto, ON, Canada.
- T12. **Zeng, Xing**. BSc (Honors) Student, Undergraduate Toronto Research Experience in Computer Science, Department of Computer Science, University of Toronto. 2015–2016.
Present position: MSc Student, Department of Computer Science, University of British Columbia, Vancouver, BC, Canada.
- T7. **Shaw, Adam J**. Student, Google Summer of Code, 2014.
Present position: Software Engineer, Google, San Bruno, CA, USA.
- T6. **Petwe, Harshad S**. BS Student, Department of Computer Science and Engineering, University of Washington, 2012.
- T5. **Ellenbogen, Paul**. BS Student, Department of Computer Science and Engineering, University of Washington, 2011–2013.
Present position: PhD Student, Department of Computer Science, Princeton University, Princeton, NJ, USA.
- T4. **Sahu, Avinash D**. MS Student, School of Communication and Computer Science, École Polytechnic Fédérale de Lausanne, 2011.
Present position: Postdoctoral Fellow, Department of Computer Science, University of Maryland, College Park, MD, USA.
- T3. **Staples, Jeffrey**. PhD Student, Department of Genome Sciences, University of Washington, 2011.
Present position: Bioinformatics R&D Developer, Regeneron Genetics, Tarrytown, NY, USA.
- T2. **Lemus Vergara, Tzitziki J**. PhD Student, Department of Genome Sciences, University of Washington, 2009.
Present position: Postdoctoral fellow, Department of Human Genetics and Biological Chemistry, University of California, Los Angeles, CA, USA, 2016.
- T1. **Buske, Orion J**. BS (Honors) Student, Department of Computer Science and Engineering, University of Washington, 2009–2010.
Present position: PhD Student, Department of Computer Science, University of Toronto, Toronto, ON, Canada.

Media coverage

Media coverage: interviews

- [“Human Longevity Study Sparks Questions About Face Prediction Claims, Data Sharing, Role of Preprints.”](#) GenomeWeb. 8 September 2017.
- [“An interview with Michael Hoffman.”](#) *EMBL Australia Bioinformatics Resource*. 3 March 2017.
- [“Expanding the DNA alphabet to understand cancer.”](#) *OICR News*. 1 December 2016.
- [“Attack of the data suckers.”](#) *ASH Clinical News*. 27 April 2016.
- [“Biden’s cancer bid exposes rift among researchers.”](#) *Politico*. 31 January 2016.

- “NEJM Editor Backtracks on Data-Sharing ‘Parasites’ Editorial.” *Medscape Medical News*. 26 January 2016.
- “Strength In Numbers: Finding And Developing Bioinformaticians.” *Front Line Genomics*. 13 October 2014.
- “Kelso: For barbecue, this Texan-turned-Canadian goes to great lengths.” *Austin American-Statesman*. 27 July 2014.
- “As Biomedical Researchers Face Tough Job Market, Experts Offer Advice and Propose Changes.” *GenomeWeb*. 5 June 2014.
- “101 questions with a bioinformatician #4: Michael Hoffman.” *ACGT*. 20 April 2014.
- “Meet Michael Hoffman.” *Princess Margaret Cancer Foundation Blog*. 16 January 2014.

Media coverage: meeting reports

- “Genome Informatics 2014.” *Genome Biology*. 22 November 2014.

Citizenship and immigration

- United States of America: citizen.
- Canada: permanent resident.