Lee Zamparo

Research Scholar, Memorial Sloan Kettering Cancer Center

Education

2008–2015	**PhD, Computer Science**, *University of Toronto*, Toronto, ON, Canada.
	Research Interests: Machine learning, computational biology: chromatin structure, gene regulation, epigenomics.
	Courses: Topics in Machine Learning, Machine Learning, Graph Theory, Machine learning for Computational Biology, Topics in Software Engineering
2012	**SciNet High Performance Computing Summer School**, *University of Toronto*, Toronto, ON, Canada.
2005–2008	**MSc, Computer Science**, *Concordia University*, Montreal, QC, Canada.
	Courses: Topics in Bioinformatics, Data-base systems, Readings in Evolutionary Algorithms, Mining Biological Sequences, Knowledge Representation in Bioinformatics
1999–2004	**BSc in Computer Science**, *University of Toronto*, Toronto, ON, Canada.
	Specialist in Computer Science and Statistics

Technical skills

Languages **(fluent)**	Python, R	Software tools	Git, SVN, Nose, Docker
Languages **(familiar)**	Java, C++	Numerical tools	Numpy/Scipy
GPGPU tools	Theano, Lasagne, PyTorch, CUDA, Torch		

Research Experience

Apr 2015 - present	**Post Doctoral Research Scholar**, *Memorial Sloan Kettering Cancer Center*, New York, NY.

- Working on multi-task deep learning models for regulatory genomics (in-development)
- Working on DNA sequence embedding for TF-DNA binding prediction **(submitted)**
- Developed statistical model for calling significant interactions in HiC data **(published)**
  - Contributed half of the manuscript figures
  - Presented work at ICML Computational Biology Workshop 2016
- Worked as part of multiple post-doc and PhD students teams from different labs
- Co-developed and taught quantitative biology course (https://github.com/qbio-gsk/qbio-gsk)

Sep 2008 - Apr 2015	**PhD student**, *University of Toronto*, Toronto, ON.

- Developed classification pipeline to estimate phenotype penetrance in yeast cell image data **(published)**
  - Contributed multiple figures, and wrote computation methods
- Developed deep autoencoder for dimensionality reduction of yeast cell image data **(submitted)**
  - Presented at NIPS computational biology workshop 2014
  - Preprint available at arxiv.org/abs/1501.01348

▬▬ Industry Experience

July 2013 - **Researcher**, *Chematria*, Toronto, ON, Canada, *(internship).*
Jan 2014 ○ Developed machine leaning models for protein-ligand docking.
 ○ Supported by Mitacs Enterprise Award

2004-2005 **Technical Developer**, *The Jonah Group*, Toronto, ON, Canada, *(contract).*
 ○ Developed JSP and Apache Struts web applications for a variety of clients.

▬▬ Publications

Bioinformatics **GuideScan: a comprehensive and customizable guide RNA design tool**, *Alexander Reinaldo-Perez, Yuri Pritykin, Joanna Alves Dos Campos Vidigal, Sagar Chhangawalla, <u>Lee Zamparo</u>, Christina S. Leslie, Andrea Ventura*, Nature Biotechnology, 35, 347–349 (2017) doi:10.1038/nbt.3804.

Bioinformatics **An integrated model for detecting significant chromatin interactions from high-resolution Hi-C data**, *Mark Carty, <u>Lee Zamparo</u>, Alvaro González, Raphael Pelossof, Olivier Elemento, Christina S. Leslie*, Nature Communications, Volume 8, Article number: 15454 (2017), doi:10.1038/ncomms15454.

Bioinformatics **Exploring quantitative yeast phenomics with single cell analysis of DNA damage foci**, *Erin B. Styles, Karen J. Founk, <u>Lee Zamparo</u>, et al.*, Cell Systems, Volume 3, Issue 3, p264–277.e10, September 2016.

Bioinformatics **PhenoM: a database of morphological phenotypes caused by mutation of essential genes in Saccharomyces cerevisiae**, *K. Jin*, J. Li*, F.S. Vizeacoumar, Z. Li, R. Min, <u>L. Zamparo</u>, F. J. Vizeacoumar, A. Datti, B. Andrews, C. Boone, Z. Zhang*, NAR, vol. 40, No. D1, 2012.

Bioinformatics **Statistical lower bounds on protein copy number from fluorescence expression images**, *<u>L. Zamparo*</u>, T. Perkins*, Bioinformatics, vol. 25, no. 20, 2009.

▬▬ Talks and Posters

Lightning **HiC-DC: a model for detecting significant interactions via Hi-C**, *M. Carty,*
talk, poster *<u>L. Zamparo*</u>, A. González, R. Pelossof, O. Elemento, C. Leslie*, ICML MLCB workshop, 2016, New York, NY.

Lightning **Dimensionality Reduction for High Content Screening Data**, *<u>L. Zamparo*</u>,*
talk, poster *Z. Zhang*, NIPS MLCB workshop, 2014, Montreal, QC.

Conference **Dimensionality Reduction for High Content Screening Data**, *<u>L. Zamparo*</u>,*
talk *Z. Zhang*, MASAMB 2014, Sheffield, UK.

Tech talk **Dimensionality Reduction for High Content Screening Data**, *<u>L. Zamparo*</u>,*
Z. Zhang, TorBUG 2014, Toronto, ON.

Poster **Genome-wide methods to identify phenotypes**, *<u>L. Zamparo*</u>, E. Styles*, B. Andrews, Z. Zhang*, CSHL: Automated Imaging and High Throughput Phenotyping, 2012.

Poster **Estimating protein copy number based on fluctuations in intensity**, *<u>L. Zamparo*</u>, T. Perkins*, ISMB, 2008.

Memorial Sloan Kettering Cancer Center – New York, NY
📞 *646–824–7786* • ✉ *zamparol at cbio.mskcc.org*
🌐 *lzamparo.github.io/* • ⌨ *lzamparo* • 🐦 *@lzamparo* *2/2*