

### 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, Computer Science**, *University of Toronto*, Toronto, ON, Canada  
Specialist in Computer Science and Statistics

### Technical skills

- Languages** (fluent) Python, R, SQL, Java, (familiar) C/C++      **Software** (fluent) Git, VScode, Docker, tools pytest, PyTorch, Numpy/Scipy, (familiar) Numba, Tensorflow (learning) JAX

### Industry Experience

- Mar 2022 - **Senior Scientist**, *Recursion*, Toronto, ON
- Sep 2023
- Initially worked as the computational biologist on a program team, later moved to a data science role developing Recursion's software infrastructure for industrialized drug discovery
  - Lead all computational analyses for MYC program, integrating phenomic data, NGS data, proteomics data
  - Proposed and implemented a framework for evaluating improvements to a key software package used in Recursion's platform
  - Lead a hack-week project for proposing new programs directly from patient RNA samples (**picked up as a supported product for Q4 2023**)
  - Co-organized a Machine Learning journal club
- Feb 2019 - **Applied Research Scientist**, *ServiceNow*, Toronto, ON
- Mar 2022
- Contributing member of two research teams, focusing on NLP and explainable AI respectively.
  - Shipped two model explainability projects
  - Published two papers (one lead author)
  - Co-developed a tool for counterfactual explanation (**patent submitted**)
  - Co-developed a tool for improving quality and efficiency of crowd-sourcing data (**patent awarded**)
- July 2013 - **Researcher**, *Chematria*, Toronto, ON, Canada, (*internship*)
- Jan 2014
- Developed machine learning models for protein-ligand docking.
  - Supported by Mitacs Enterprise Award

Recursion – Toronto, ON

☎ 416-878-8297 • ✉ zamparo at gmail.com • 🌐 lzaparo.github.io/  
👤 lzaparo • 🐦 @lzaparo

---

## Research Experience

- Apr 2015 - **Post Doctoral Research Scholar**, *Memorial Sloan Kettering Cancer Center*, New York, NY
- Worked on DNA sequence embedding for TF-DNA binding prediction (**published**)
    - Contributed to development of sequence embedding model
    - Contributed supplemental figures
  - 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 - **PhD student**, *University of Toronto*, Toronto, ON
- Apr 2015
- Developed classification pipeline to estimate phenotype penetrance in yeast cell image data (**published**)
    - Contributed multiple figures, and wrote computation methods
  - Developed deep auto-encoder for dimensionality reduction of yeast cell image data (**submitted**)
    - Presented at NIPS computational biology workshop 2014
    - Preprint available at [arxiv.org/abs/1501.01348](http://arxiv.org/abs/1501.01348)

---

## Selected Publications

- Model Uncertainty **The Dynamics of Functional Diversity throughout Neural Network Training**, *Lee Zamparo, Marc-Etienne Brunet, Thomas George, Sepideh Kharaghani, Gintare Karolina Dziugaite*, Bayesian Deep Learning Workshop, NeurIPS (2021) <http://bayesiandeeplearning.org/2021/papers/63.pdf>
- Explainable AI **Beyond Trivial Counterfactual Explanations with Diverse Valuable Explanations**, *Pau Rodriguez, Massimo Caccia, Alexandre Lacoste, Lee Zamparo, Issam Laradji, Laurent Charlin, David Vazquez*, arXiv, arXiv:2103.10226 (2021) <https://arxiv.org/abs/2103.10226>
- Computational Biology **Sequential CRISPR gene editing in human iPSCs charts the clonal evolution of myeloid leukemia and identifies early disease targets**, *Tiansu Wang, Allison R Pine, Andriana G Kotini, Han Yuan, Lee Zamparo, Daniel T Starczynowski, Christina Leslie, Eirini P Papapetrou*, Cell Stem Cell, 28, 6, 1074-1089 (2021) <https://doi.org/10.1016/j.stem.2021.01.011>
- Computational Biology **BindSpace decodes transcription factor binding signals by large-scale sequence embedding**, *Han Yuan, Meghana Kshirsagar, Lee Zamparo, Yuheng Lu, Christina S Leslie*, Nature Methods, 16, 858-861 (2019) <https://doi.org/10.1038/s41592-019-0511-y>
- Bioinformatics **GuideScan: a comprehensive and customizable guide RNA design tool**, *Alexander Reinaldo-Perez, Yuri Pritykin, Joanna Alves Dos Campos Vidigal, Sagar Chhangawalla, Lee Zamparo, Christina S. Leslie, Andrea Ventura*, Nature Biotechnology, 35, 347-349 (2017) doi:10.1038/nbt.3804
- Computational Biology **An integrated model for detecting significant chromatin interactions from high-resolution Hi-C data**, *Mark Carty, Lee Zamparo, Alvaro González, Raphael Pelosof, Olivier Elemento, Christina S. Leslie*, Nature Communications, Volume 8, Article number: 15454 (2017), doi:10.1038/ncomms15454

Recursion – Toronto, ON

☎ 416-878-8297 • ✉ [zamparo@gmail.com](mailto:zamparo@gmail.com) • 🌐 [lzamparo.github.io/](https://lzamparo.github.io/)  
👤 [lzamparo](https://github.com/lzamparo) • 🐦 [@lzamparo](https://twitter.com/lzamparo)

2/2