BIOGRAPHICAL SKETCH

Provide the following information for the Senior/key personnel and other significant contributors.
Follow this format for each person. **DO NOT EXCEED FIVE PAGES.**

NAME: Bar-Joseph, Ziv

eRA COMMONS USER NAME: ZBARJO

POSITION TITLE: FORE Systems Professor of Computational Biology and Machine Learning

EDUCATION/TRAINING

INSTITUTION AND LOCATION	DEGREE *(if applicable)*	Completion Date MM/YYYY	FIELD OF STUDY
Hebrew University, Jerusalem, Israel	B. Sc.	07/1997	Computer Science, Math
Hebrew University, Jerusalem, Israel	M. Sc.	08/1999	Computer Science
Massachusetts Institute of Technology	Ph. D.	06/2003	Computer Science

A. Personal Statement

My work focuses on the analysis, processing and modeling of high throughput biological data and biological networks using machine learning techniques. We have specifically been focused on the analysis of time series and longitudinal lung multi-omic data, on modeling scRNA-Seq temporal data and on the use of deep learning methods for the analysis of high throughput biological data. We have developed methods for integrating static interaction data with time series data and methods to obtain and improve condition and species specific data so that it can be used to model dynamic regulation. We have also focused on the analysis and comparison of cross species data. Our methods have been used, by us and others, to model a large number of biological systems in several species. The software tools we developed are actively used by several researchers around the world.

a. J. Ding, B. Aronow, N. Kaminski, J. Kitzmiller, J. Whitsett and Z. Bar-Joseph. Reconstructing differentiation networks and their regulation from time series single cell expression data. *Genome Res.* 28: 383-395, 2018
b. J. Ding, F. Ahangari, C.R. Espinoza, D. Chhabra, T. Nicola, X. Yan, C.V. Lal, J.S. Hagood, N. Kaminski, Z. Bar-Joseph*, N. Ambalavanan*. Integrating multi-omics longitudinal data to reconstruct networks underlying lung development. *AJP Lung*, 317(5):L556-L568, 2019
c. C. Lin and Z. Bar-Joseph. Continuous State HMMs for Modeling Time Series Single Cell RNA-Seq Data. *Bioinformatics.* 35(22), 4707–4715, 2019.
d. K. Hurley, J. Ding, C. Villacorta-Martin, M.J. Herriges, A. Jacob, M. Vedaie, K.D. Alysandratos, Y.L. Sun, C. Lin, J. Huang, A.A. Wilson, A. Mithal, G. Mostoslavsky, I. Oglesby, I. Caballero, S.H. Guttentag, F. Ahangari, N. Kaminski, A. Rodriguez-Fraticelli, F. Camargo, Z. Bar-Joseph* and D.N. Kotton*. Reconstructed Single-Cell Fate Trajectories Define Lineage Plasticity Windows during Differentiation of Human PSC-Derived Distal Lung Progenitors. *Cell Stem Cell*. Epub ahead of print. 2020.
* corresponding authors

B. Positions and Honors

Positions and Employment

1999-2003 Research Assistant (Teaching Assistant, 2001), EECS, MIT, Cambridge, MA
2003 Postdoctoral Associate, Rick Young and David Gifford Lab, Whitehead Institute, MIT, Cambridge, MA
2003-2009 Assistant professor, Machine Learning Dept. School of Computer Science, Carnegie Mellon University, Pittsburgh, PA
2009-2015 Associate professor, Lane Center for Computational Biology, Machine Learning Dept., School of Computer Science, Carnegie Mellon University, Pittsburgh, PA

2015- Professor, Department of Computational Biology, Machine Learning Dept., School of Computer Science, Carnegie Mellon University, Pittsburgh, PA
2017- FORE Systems Professor of Computational Biology and Machine Learning, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA

Other Experience and Professional Memberships

2001- Member, International Society for Computational Biology
2009-2012 Editorial Board, Bioinformatics
2009, 2010 PC chair, Recomb Regulatory Genomics and Systems Biology
2013- 2017 Associate Editor, Bioinformatics
2013-2015 PC chair: Workshop on Biological Distributed Algorithms (BDA)
2014-2017 Stirring Committee Member: NIAID Systems Biology Program
201-2018 Standing member: NIH Modeling and Analysis of Biological Systems (MABS) study section
2015-2019 Co-Director, Big Data for Better Health (BD4BH) in Pennsylvania
2017- Director, joint CMU-Pitt PhD Program in Computational Biology (CPCB)
2018 - Scientific Advisory Board Member, Cancer Systems Biology Center, UC Irvine
2018- Director, HuBMAP HIVE Computational Tools Center
2019- Advisory Board Member, Salk / Allen Center for Aging and Alzheimer's Disease

Honors

1998- Best student paper award, ACM Symposium on Principles of Distributed Computing (PODC)
2001-2003 Program in Mathematics and Molecular Biology (PMMB), national fellow
2003 Dimacs-Celera Genomics graduate student award in computational molecular biology
2005 National Science Foundation CAREER award
2012 Overton Prize, International Society for Computational Biology (ISCB)
2016 Best paper award, 20th Conference on Research in Computational Molecular Biology (RECOMB)
2017 FORE Systems Chair, School of Computer Science, Carnegie Mellon University

C. Contribution to Science

1. My early work focused on developing one of the first methods for integrating the (then very new) protein-DNA interaction data (primarily from ChIP-CHIP experiments) with gene expression data. More recently we extended these methods and applied them to integrate several other types of biological data. Some of the papers related to that work are:

a. Afek, N. Alon, O, Barad, E. Hornstein, N. Barkai, and Ziv Bar-Joseph. A biological solution to a fundamental distributed computing problem. *Science*, 331(6014):183-5, 2011.
b. G.E. Zinman, S. Naiman, Y. Kanfi, H. Cohen, and Z. Bar-Joseph. ExpressionBlast: Mining large, unstructured, expression databases. Nature Methods,10:925–926, 2013.
c. L. Song, S.C. Huang, A. Wise, R. Castanon, J.R. Nery, H. Chen, M. Watanabe, J. Thomas1, Z. Bar-Joseph and J.R. Ecker. A transcription factor hierarchy defines an environmental stress response network. *Science*, 354(6312). pii: aag1550, 2016.
d. S. Rashid, Z. Long, S. Sing, M. Kohram, H. Vashistha, S. Navlakha, H. Salman, Z. N. Oltvai, and Ziv Bar-Joseph. Adjustment in tumbling rates improves bacterial chemotaxis on obstacle-laden terrains. *Proceedings of the National Academy of Science (PNAS)*, 16(24):11770-11775, 2019.

2. A major focus over the last couple of years have been on developing and applying methods for reconstructing dynamic regulatory networks. Our methods extend HMMs enabling the integration of time series and static data including mRNA and miRNA gene expression data, epigenetic data and static protein-DNA interaction data and single cell time series data. These methods were applied, by us and several other groups, to model various responses, to study development and to model several other dynamic processes.

a. Z. Bar-Joseph, A. Gitter, I. Simon. Studying and modelling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics*, 13, 552-564, 2012
b. E. Sefer, M. Kleyman and Z. Bar-Joseph. Tradeoffs between dense and replicate sampling strategies for high throughput time series experiments. *Cell Systems*, 3(1):35-42, 2016.
Prior version in *Proceedings of the 20th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, 2016

*Recipient of the 'Best Paper Award' in RECOMB 2016

c. J. Ding, J.S. Hagood, N. Ambalavanan, N. Kaminski, Z. Bar-Joseph. iDREM: Interactive visualization of dynamic regulatory networks. *PLoS Comput Biol.* 14(3):e1006019, 2018
d. Y. Yuan and Ziv Bar-Joseph. Deep learning for inferring gene relationships from single-cell expression data. *Proceedings of the National Academy of Science (PNAS)*, 116 (52) 27151-27158, 2019.

3. More recently we have focused on the use of single cell RNA-Seq data (scRNA-Seq) data for the analysis and modeling of developmental and differentiation processes. We have developed several methods for this data including for initial processing, for temporal reconstruction and for integrating scRNA-Seq data with bulk genomics data.

a. A. Alavi, M. Ruffalo, A. Parvangada, Z. Huang, and Z. Bar-Joseph. scQuery: a web server for comparative analysis of single-cell RNA-seq data. *Nature Communications*, 9(1):4768, 2018.
b. C. Lin and Z. Bar-Joseph. Continuous State HMMs for Modeling Time Series Single Cell RNA-Seq Data. *Bioinformatics. 35(22):4707-4715,* 2019.
c. J. Ding, C. Lin, and Z. Bar-Joseph, Cell lineage inference from SNP and scRNA-Seq data. *Nucleic Acids Research*, 47(10):e56, 2019.
d. M.P. Snyder, S. Lin, A. Posgai, M. Atkinson, R. Satija, N. Gehlenborg, J. Laskin, P. Harbury, N.A. Nystrom, J.C. Silverstein, Z. Bar-Joseph*, K. Zhang, K. Börner, L. Cai, S.A. Teichmann, B. Paten, P. Mabee R. Conroy. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature*, 574(7777):187-192, 2019.
* corresponding author

D. Ongoing Research Support

2T32EB009403 Bar-Joseph (MPI) 08/01/2019-07/30/2024
Integrated interdisciplinary, inter-university phd program computational biology
Goal: Train the next generation of computational biology PhD students

1U01HL145567-01 Kaminski (PI) 01/01/2019-11/30/2022
Normal Aging Lung Cell Atlas (NALCA)
Goals: Develop method for modeling time series studies of developing lungs

OT2OD026682 Bar-Joseph (PI) 09/01/2018 – 08/31/2022
Comprehensive Flexible and FAIR Tools for the HuBMAP HIVE
Goal: Develop computational tools to analyze imaging and expression data generated by HUBMAP.

1R01GM122096 Bar-Joseph (PI) 09/01/2017-08/31/2021
Reconstructing regulatory networks from time series single cell data
Goals: Develop methods for modeling iPSC differentiation to lung cells using scRNA-Seq data

1R01HL128172 Kotton (PI) 07/01/2015-06/30/2020
Epigenenomic and transcriptomic networks in normal and defective lung development
Goals: Study epigenetic modifications in population of cells for a lung development model
Role: Investigator

1U01 HL122626-01 Ambalavanan (PI) 04/01/14- 03/30/20 (NCE).
Alveolar DevMAP
Goals: Model lung development in mice using time series transcription data from population of cells
Role: Investigator

James S. McDonnell Foundation Bar-Joseph (PI) 10/01/13 - 09/31/21 (NCE)
Scholars Award in Studying Complex Systems
Goals: Study bi-directional computational and biological systems to improve our understanding of brain development and the robustness and efficiently of networks.

<u>Completed Research Support within the last 3 years</u>

1. CURE, PA Department of Health, "Big Data for Better Health (BD4BH) in Pennsylvania". Bar-Joseph (PI), 2015-2019
2. 1R01HL127349 "Genomic Analysis of Tissue and Cellular Heterogeneity in IPF". Kaminski (PI) 2015-2019
3. NSF DBI- 1356505 "ABI Innovation: BCSP: Understanding the design and usage of distributed biological networks". To Bar-Joseph (PI), 2014- 2019