|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| BIOGRAPHICAL SKETCH Provide the following information for the key personnel and other significant contributors in the order listed on Form Page 2. Follow this format for each person.  **DO NOT EXCEED FOUR PAGES.** | | | | |
|  | | | | |
| NAME  **Joshua T. Vogelstein** | | POSITION TITLE  **Assistant Professor, Dept of Biomedical Engineering, Institute for Computational Medicine,**  **Center for Imaging Science,**  **Institute for Data Intensive Engineering and Sciences, JHU** | | |
| eRA COMMONS USER NAME  jvogels3 | |
| EDUCATION/TRAINING *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)* | | | | |
| INSTITUTION AND LOCATION | DEGREE  *(if applicable)* | | YEAR(s) | FIELD OF STUDY |
| **Washington University, St. Louis, MO** | **BS** | | **2002** | **Biomedical Eng.** |
| **Johns Hopkins University, Baltimore MD** | **MS** | | **2009** | **Applied Math. & Stats.** |
| **Johns Hopkins University, Baltimore MD** | **PhD** | | **2009** | **Neuroscience** |
| **Johns Hopkins University, Baltimore MD** |  | | **2009-2012** | **Applied Math. & Stats.** |
| **Duke University, Durham, NC** |  | | **2012-2014** | **Math. & Stats.** |

1. **Personal Statement**

I am an expert in both neuroscience and statistical analysis, with graduate degrees in both. My contributions typically arise from a deep understanding of the neurobiological question of interest, as well as theoretical and practical knowledge in terms of developing and implementing methodologies to address these questions. I started my career focusing on single and populations of neurons, with contributions in deconvolution of calcium imaging, spike sorting, and neural coding. Since then, I have scaled both up and down in terms of spatial and temporal resolution. I have developed high-throughput pipelines for analysis of functional and diffusion magnetic resonance imaging, as well as similarly functional pipelines for electron microscopy. My code for these pipelines, as well as all of my groups work, is always developed open source. Several of my packages have already become the industry standards. My data analytics have included not just neural activity, but also brain connectivity (connectomes) and behavior. I have experience working across the evolutionary hierarchy, ranging from C. elegans to Drosophila to mice to humans. The four publications listed below are all examples of me developing and/or applying novel computational statistical methods to solve open problems, in neuroscience, and more generally.

1. **Vogelstein JT,** Park Y, Ohyama T, Kerr R, Truman JW, et al. (2014). Discovery of Brainwide Neural-Behavioral Maps via Multiscale Unsupervised Structure Learning. Science. 344(6182), 386-92.
2. Roberts NJ\*, **Vogelstein JT**\*, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE .(2012). The Predictive Capacity of Personal Genome Sequencing. Scientific Translational Medicine, 4(133):133ra58. PMCID: PMC3741669.
3. **Vogelstein JT**, Gray WR, Vogelstein RJ, Priebe CE. (2012). Graph Classification using Signal Subgraphs: Applications in Statistical Connectomics. IEEE Transactions on Pattern Recognition and Machine Intelligence, 5(7), 1539-51.
4. **Vogelstein JT**, Packer AM, Machado TA, Sippy T, Babadi B, et al. (2010). Fast non-negative deconvolution for spike train inference from population calcium imaging. Journal of Neurophysiology 104(6), 3691-704. PMCID: PMC3007657.
5. **Positions and Honors**

**Positions and Employment:**

2014 to present: Assistant Professor, Department of Biomedical Engineering, Johns Hopkins University

2014 to present: Faculty, Institute for Computational Medicine, Johns Hopkins University

2014 to present: Core Member, Center for Imaging Science, Johns Hopkins University

2010 to present: Adjunct Research Scientist, Human Language Technology Center of Excellence, Johns Hopkins University.

2012 to present: Member of the Institute for Data Intensive Engineering and Sciences, Johns Hopkins University.

2013 to 2014: Senior Research Scientist, Departments of Statistical Science and Mathematics, Duke University.

2013 to 2014: Assistant Consulting Professor, Department of Neurobiology, Duke University.

2013 to 2014: Senior Fellow at the Kenan Institute for Ethics, Duke University.

2012 to 2013: Visiting Assistant Research Professor, Department of Mathematics, Duke University.

2011 to 2012: Assistant Research Scientist, Department of Applied Mathematics and Statistics, Johns Hopkins University.

2004 to 2012: Chief Scientist, Global Domain Partners, LLC.

2010 to 2011: Postdoctoral Fellow under Professor Carey E Priebe, Department of Applied Mathematics and Statistics, Johns Hopkins University.

Summer of 2001: Research Assistant under Professor Randy O'Reilly, Department of Psychology, University of Colorado.

Summer of 2000: Clinical Engineer, Johns Hopkins Hospital.

Summer of 1999: Research Assistant under Dr. Jeffrey Williams, Department of Neurosurgery, Johns Hopkins Hospital.

Summer of 1998: Research Assistant under Professor Kathy Cho, Department of Pathology, Johns Hopkins School of Medicine.

**Honors:**

2013: Spotlight, Neural Information Processing Systems (NIPS).

2011: Trainee Abstract Award, Organization for Human Brain Mapping.

2008: Spotlight, Computational and Systems Neuroscience (CoSyNe).

2008: Successfully completed Molecular Biology Summer Workshop. Smith College, Mass, USA.

2008: Successfully completed Advanced Techniques in Molecular Neuroscience. Cold Spring Harbor, New York, USA.

2005: Successfully audited Imaging Structure and Function of the Nervous System. Cold Spring Harbor, New York, USA.

2004: Successfully completed Advanced Course in Computational Neuroscience. Obidos, Portugal.

1. **Contribution to Science**
2. **Spike sorting from Calcium optophysiology experiments**. In my PhD, I became aware of a novel experimental method that seemed to have huge potential: calcium optophysiology experiments. At the time, it was not know how to automatically detect action potentials given a calcium trace. We developed two different algorithms to address this problem. Since then, the manuscripts describing those algorithms have been collectively cited approximately 200 times. I developed the algorithms with Liam Paninski, and then wrote all of the code, and continue to support its use. By making my code publicly available, it has become the de facto code for analysis used by tens of laboratories around the world.
3. **Vogelstein JT**, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R, Paninski L. (2010). Fast nonnegative deconvolution for spike train inference from population calcium imaging. Journal of neurophysiology. 104(6), 3691-704. PMCID: PMC3007657.
4. Paninski L, Ahmadian Y, Ferreira DG, Koyama S, Rad KR, Vidne M, **Vogelstein JT,** & Wu W. (2010). A new look at state-space models for neural data. Journal of computational neuroscience, 29(1-2), 107-26. PMCID: PMC3712521.
5. **Vogelstein JT**, Watson BO, Packer AM, Yuste R, Jedynak B, Paninski L. (2009). Spike inference from calcium imaging using sequential Monte Carlo methods. Biophysical Journal, 97(2), 636-55. PMCID: PMC2711341.
6. Hofer SB, Ko H., Pichler B, **Vogelstein JT**, Ros H, Zeng H, Lein E, Lesica NA, Mrsic-Flogel TD. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nature neuroscience, 14(8), 1045-52.
7. **Tools to estimate brain networks from different modalities.** Brains, to a first approximation, can quite reasonably be modeled as a network. However, until a few years ago, tools for efficient network inference were essentially absent from the community. Since then, we have developed tools to estimate brain networks from several different modalities of data, starting with calcium imaging (for which I co-developed the methods, and wrote much of the code), then diffusion and functional magnetic resonance imaging (for which I supervised students who wrote all the code), and more recently electron microscopy.
8. Mishchenko Y, **Vogelstein JT**, Paninski L. (2011). A Bayesian approach for inferring neuronal connectivity from calcium fluorescent imaging data. The Annals of Applied Statistics, 5(2B), 1229-1261. PMCID: PMC3704386.
9. Gray WR, Bogovic JA, **Vogelstein JT,** Landman BA, Prince JL, Vogelstein RJ. (2011). Magnetic Resonance Connectome Automated Pipeline. IEEE Pulse, 3(2) 42-48.
10. Sikka S, Cheung B, Khanuja R, Ghosh S, Yan C, Li Q, **Vogelstein JT**, Burns R, Colcombe S, Craddock C, Mennes M, Kelly C, Dimartino A, Castellanos F, Milham M. (2013). [Towards Automated Analysis of Connectomes: The Configurable Pipeline for the Analysis of Connectomes (C-PAC)](http://www.frontiersin.org/10.3389/conf.fninf.2014.08.00117/event_abstract). Conference Abstract: Neuroinformatics.
11. Gray Roncal W, Kleissas DM, **Vogelstein JT,** Manavalan P, Burns R, Vogelstein RJ, Priebe CE, Chevillet MA, Hager GD. (2014). An Automated Images-to-Graphs Pipeline for High Resolution Connectomics. arXiv:1411.6880.
12. **Foundational statistical theory for the analysis of populations of graphs.** As neuroscientists build better and better tools for estimating both physiological and anatomical networks, there is a growing need for statistical tools for the analysis of these data. Unfortunately, most published works on estimated graphs to not take advantage of the graph structure of the data, and therefore can be both under powered and overly confident. We have been building foundational statistical tools for the analysis of graphs. My role in this work has ranged from developing the method, to writing the code, to analyzing the data.
13. **Vogelstein JT,** Roncal Gray, W, Vogelstein RJ, Priebe CE. (2011). Graph Classification using Signal Subgraphs: Applications in Statistical Connectomics. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(7), 1539-1551.
14. **Vogelstein JT,** Vogelstein RJ, Priebe CE. (2011). Are mental properties supervenient on brain properties? Scientific reports, 1, 100. PMCID: PMC3216585
15. **Vogelstein JT,** Conroy JM, Podrazik LJ, Kratzer SG, Fishkind DE, Vogelstein RJ, Priebe CE. (2015). Fast Approximate Quadratic Programming for Graph Matching. PLoS One, 10(4):e0121002. PMCID: PMC4401723
16. Fiori M, Sprechmann P, **Vogelstein JT,** Musé P, Sapiro G. (2013). Robust multimodal graph matching: Sparse coding meets graph matching. Advances in Neural Information Processing Systems, 127-135.
17. **Tools for big data neuroscience.** As both the raw data (typically images) and the data derivatives (spike train and graphs) get increasingly large, the standard data analytic workflow starts breaking down, and scientists need new computational models. Therefore, I co-founded the “Open Connectome Project” (<http://openconnecto.me>) in 2011, to lower the barrier to entry into big data neuroscience. Since then, we have supported, hosted, and analyzed data for several of the most highly impactful papers. Below, I list several of our manuscripts describing our infrastructure and development.
18. Burns R, Gray Roncal W, Kleissas D, Lillaney K, Manavalan P, Perlman E, Berger DR, Bock DD, Chung K, Grosenick L, Kasthuri N, Weiler NC, Deisseroth K, Kazhdan M, Lichtman J, Reid RC, Smith SJ, Szalay AS, **Vogelstein JT**, Vogelstein RJ. (2013). The Open Connectome Project Data Cluster: Scalable Analysis and Vision for High-Throughput Neuroscience. Scientific and Statistical Database Management Conference. PMCID: PMC3881956
19. Kazhdan M, Burns R, Kasthuri B, Lichtman J, Vogelstein RJ, **Vogelstein, JT** (2013). Gradient-Domain Processing for Large EM Image Stacks. arXiv preprint arXiv:1310.0041.
20. Burns R, **Vogelstein JT,** Szalay AS. (2014). From Cosmos to Connectomes: The Evolution of Data-Intensive Science. Neuron, 83(6), 1249-1252.
21. Zheng D, Mhembere D, Burns R, **Vogelstein JT**, Priebe CE, and Szalay AS. (2015). FlashGraph: Processing Billion-Node Graphs on an Array of Commodity SSDs. Proceedings of the 13th USENIX Conference on File and Storage Technologies (FAST ’15), 45-58.

## Complete List of Published Work in Google Scholar: <https://scholar.google.com/citations?user=DWPfdT4AAAAJ&hl=en&oi=ao&cstart=-80&pagesize=100>

## 

1. **Research Support**

**Ongoing Research Support:**

R01 OD019123 (NIH – TRA) Smith (PI) 9/30/2014 – 6/30/2019

Synaptomes of Mouse and Man

The major goals of this project are to discover the synaptic diversity and complexity in mammalian brains, specifically comparing and contrasting humans with mice, the leading experimental animal.

GRANT11551224 Burns (PI) 2/1/2014 – 11/13/2015

Title: Scalable Brain Graph Analyses Using Big-Memory, High-IOPS Compute Architectures

The major goals of this project are to design, implement, and deploy graph analytics on heterogeneous hardware for the analysis of human (and other) brain graphs.

1R01EB016411 (NIH/NSF – CRNCS) Burns (PI) 9/10/2012 – 8/31/2015

Title: The EM Open Connectome Project

The major goals of this project are to develop open-source software and publicly accessible infrastructures for the neuroscience community to collect, curate, and analyze electron microscopy (EM) connectomes.

1R01DA036400 (NIH/NSF – BIGDATA) Mitra (PI) 3/15/2013 – 1/31/2016

Title: BigData: Small DCM: ESCA DA Computational Infrastructure for Massive Neuroscience

The major goals of this project are to develop a data-intensive architecture for the management and analysis of massive mouse brain imaging datasets.

FA8750-12-2-0303 (DARPA – XDATA) Priebe (PI) 10/1/2012 – 3/31/2017

Title: Fusion and Inference from Multiple and Massive Disparate Distributed Dynamic Data Sets

The major goals of this project are to develop and extend rigorous statistical theory and practice to perform various inferential tasks when the data are massive, distributed, and dynamic.

1R01MH099647-01 (NIH – TR01) Deisseroth (PI) 7/01/2012 – 6/30/2017

Title: CLARITY: Fully-Assembled Biology

The major goals of this project are to develop technology to enable peering into the structural and activity connectomes of nearly intact neural circuits.

N/A Vogelstein (PI) 2/11/2013 – 9/30/2015

Title: Endeavor Scientists Training Fellowship

The major goals of this project are to analyze MR derived brain imaging data using rigorous statistical methods in collaboration with the Child Mind Institute.

**Completed Research Support:**

964165 RJ Vogelstein (PI) 12/16/09 – 1/14/13

Title: National Center for Applied Neuroscience Project (NCAN)

The major goals of this project are to develop statistical methodologies with translational impact in neuroscience.

FA8750-12-C-0239 Andrews (PI) 11/8/2012 – 10/31/2013

Title: Graph-Based Scalable Analytics for Big Data

The major goals of this project are to derive scalable analytics for graph data.

Data Readiness Level Harer (PI) 1/1/2014 – 12/31/2014

Title: Data Readiness Level

The major goals of this project are to methods that can assess the value of a piece of data at any particular point in the analytic process.