Julia Fukuyama

Email: jfukuyam@iu.edu
url: jfukuyama.github.io

Current position

Assistant Professor   2018-present
Department of Statistics
Indiana University, Bloomington

Appointments held

Postdoctoral Research Fellow   2017-2018
Department of Computational Biology
Fred Hutchinson Cancer Research Center

Education

PhD Statistics, Stanford University   2017
Advisor: Susan Holmes

MS Statistics, Stanford University   2012

BS Biology, *magna cum laude*, Yale University   2010

Preprints/Submitted

Julia Fukuyama. A decomposition of a phylogenetically-informed distance into basal and terminal components. Revisions sent back to *Theoretical Population Biology*. Preprint at https://www.biorxiv.org/content/10.1101/2025.04.17.649379v1.

Frederick A Matsen, Kevin Sung, Mackenzie M. Johnson, Will Dumm, David Rich, Tyler N Starr, Yun S. Song, Philip Bradley, *Julia Fukuyama*, Hugh Haddox. A sitewise model of natural selection on individual antibodies via a transformer-encoder. Submitted to *Molecular Biology and Evolution*.

Gerardo Flores, Abigail Schiff, Alyssa Smith, *Julia Fukuyama*, Ashia Wilson. A Consequentialist Critique of Binary Classification Evaluation Practices. https://doi.org/10.48550/arXiv.2504.04528.

Gerardo Flores, Alyssa Smith, *Julia Fukuyama*, Ashia Wilson. Aligning Evaluation with Clinical Priorities: Calibration, Label Shift, and Error Costs. Preprint at https://doi.org/10.48550/arXiv.2506.14540.

Mackenzie M. Johnson, Kevin Sung, Hugh Haddox, Ashni Vora, Tatsuya Araki, Gabriel Victoria, Yun Song, *Julia Fukuyama*, Frederick Matsen. Nucleotide context models outperform protein language models for predicting antibody affinity maturation. Preprint at https://doi.org/10.1101/2025.06.16.659977.

Kevin Sung, Mackenzie M. Johnson, Will Dumm, Noah Simon, Hugh Haddox, *Julia Fukuyama*, Frederick A Matsen IV. Thrifty wide-context models of B cell receptor somatic hypermutation. *eLife*, 2025. https://doi.org/10.7554/eLife.105471.1.sa3 (A reviewed preprint.)

Publications

Thayer Fisher, Kevin Sung, Noah Simon, *Julia Fukuyama**, Erick Matsen*. Inferring Mechanistic Parameters of Somatic Hypermutation Using Neural Networks and Approximate Bayesian Computation. *Annals of Applied Statistics*, 2025.

Luis Miguel Mestre, Roger S. Zoh, Cydne Perry, *Julia Fukuyama*, Maria A. Parker. Determining whether weight status mediates the association between number of cigarettes smoked per day and all-cause mortality among US adults who smoke cigarettes. *PLoS ONE*, 2025.

Justin Lin and *Julia Fukuyama*. Calibrating dimension reduction hyperparameters in the presence of noise. *PLoS Computational Biology*, 2024.

Hongxuan Zhai and *Julia Fukuyama*. A convenient correspondence between k-mer-based metagenomic distances and phylogenetically-informed β-diversity measures. *PLoS Computational Biology*, 2023.

*Julia Fukuyama**, Kris Sankaran*, Laura Symul*. Multiscale Analysis of Count Data through Topic Alignment. *Biostatistics*, 2023.

Julia Fukuyama. Local biplots for multidimensional scaling, with application to the microbiome. *Journal of Computational and Graphical Statistics*, 2022.

Brielle Stark and *Julia Fukuyama*. Leveraging big data to understand the interaction of task and language during monologic spoken discourse in speakers with and without aphasia. *Language, Cognition, and Neuroscience*, 2020.

Julia Fukuyama, Branden Olson, Frederick Matsen. Lack of evidence for a substantial rate of templated mutagenesis in B cell diversification. *Journal of Immunology*, 2020.

Julia Fukuyama. Emphasis on the deep or shallow parts of the tree provides a new characterization of phylogenetic distances. *Genome Biology*, 2019.

Julia Fukuyama. Adaptive gPCA: A method for structured dimensionality reduction. *Annals of Applied Statistics*, 2019.

Diana Proctor, *Julia Fukuyama*, Peter Loomer, Gary Armitage, Stacey Lee, Nicole Davis, Mark Ryder, Susan Holmes, David Relman. A spatial gradient of bacterial diversity in the human oral cavity shaped by salivary flow. *Nature Communications*, 2018.

Julia Fukuyama*, Laurie Rumker*, Kris Sankaran*, Pratheepa Jeganathan, Les Dethlefsen, David A. Relman, Susan Holmes. Multidomian analyses of a longitudinal human microbiome intestinal cleanout perturbation experiment. *PLoS Computational Biology*, 2017.

Elena Vendrame*, *Julia Fukuyama**, Dara Strauss-Albee, Susan Holmes, Catherine Blish. Mass cytometry analytical approaches reveal cytokine-induced changes in natural killer cells. *Cytometry Part B: Clinical Cytometry*, 2016.

Benjamin Callahan, Kris Sankaran, *Julia Fukuyama*, Paul McMurdie, and Susan Holmes. Bioconductor workflow for microbiome data analysis: from raw reads to community analyses. *F1000Research*, 2016.

Benjamin Callahan, Diana Proctor, *Julia Fukuyama*, David A. Relman, Susan Holmes. Reproducible research workflow in R for the analysis of personalized human microbiome data. *Pacific Symposium on Biocomputing*, 2016.

Dara Strauss-Albee, *Julia Fukuyama*, Emily Liang, Yi Yao, Justin Jarrell, Alison Drake, John Kinuthia, Ruth Montgomery, Grace John-Stewart, Susan Holmes, Catherine Blish. Human NK cell repertoire diversity reflects immune experience and correlates with viral susceptibility. *Science Translational Medicine*, 2015.

Alex Kay, Nick Bayless, *Julia Fukuyama*, Natali Aziz, Cornelia Dekker, Sally Mackey, Gary Swan, Mark Davis, Catherine Blish. Pregnancy does not attenuate the antibody or plasmablast response to inactivated influenza vaccine. *Journal of Infectious Diseases*, 2015.

Alex Kay, *Julia Fukuyama*, Natali Aziz, Cornelia Dekker, Sally Mackey, Gary Swan, Mark Davis, Susan Holmes, Catherine Blish. Enhanced natural killer-cell and T-cell responses to influenza A virus during pregnancy. *Proceedings of the National Academy of Sciences*, 2014.

Miling Yan, Sunje Pamp, *Julia Fukuyama*, Peter Hwang, Do-Yeon Cho, Susan Holmes, David A. Relman. Nasal microenvironments and interspecific interactions influence nasal microbiota complexity and *S. aureus* carriage. *Cell Host and Microbe*, 2013.

Julia Fukuyama, Paul McMurdie, Les Dethlefsen, David A. Relman, Susan Holmes. Comparisons of Distance Methods for Combining Covariates and Abundances in Microbiome Studies. *Pacific Symposium on Biocomputing*, 2012.

Jeffrey Isenberg, Yifeng Jia, *Julia Fukuyama*, Christopher Switzer, David A. Wink, David D. Roberts. Thrombospondin-1 inhibits nitric oxide signaling via CD36 by inhibiting myristic acid uptake. *Journal of Biological Chemistry*, 2007.

Invited Presentations

The convenient correspondence between k-mer-based metagenome distances and phylogenetically-informed β-diversity measures
Department of Mathematics and Statistics, Washington University in St. Louis, St. Louis, MO, October 2022.

Phylogenetically-informed distance methods for microbiome data analysis
Joint Statistical Meetings, Washington, DC, August 2022.

Multiscale alignment of count data through topic alignment
EcoSta, Kyoto, Japan, June 2022.

Phylogenetically-informed distance methods for microbiome data analysis
Joint Statistical Meetings, Philadelphia, PA, August 2020.

Local biplots and their application to microbiome data
CMStatistics, London, UK, December 2019.

Phylogenetically-informed distance methods for microbial ecology
Notre Dame Department of Statistics, South Bend, IN, October 2019.

Phylogenetically-informed distance methods for microbiome data analysis
University of Louisville Department of Statistics, Louisville, KY, September 2019.

Phylogenetically-informed distance methods: Their uses, properties, and potential
The Fourth Workshop on Higher-Order Aspmytotics and Post-Selection Inference
Washington University, St. Louis MO, August 2019.

Unboxing a Black Box: Properties of Phylogenetically-Informed Distances
Workshop on "The Role of Genomics and Metagenomics in Human Health: Recent Developments in Statistical and Computational Methods"
Banff International Research Station, Banff, Canada, February 2019.

Function-Based Clustering of Amplicon Sequence Variants
CMStatistics, University of Pisa, Pisa, Italy, December 2018.

Using Phylogenetic Information to Understand the Microbiome.
Biostatistics Colloquium, Fred Hutchinson Cancer Research Center, Seattle, WA, May 2018.

Using Phylogenetic Information to Understand the Microbiome
University of California Berkeley Center for Computational Biology, Berkeley, CA, March 2018.

Using Phylogenetic Information to Understand the Microbiome
University of Vermont Department of Statistics, Burlington, VT, March 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Wharton Department of Statistics, Philadelphia, PA February 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Varderbilt University Department of Biostatistics, Nashville, TN, February 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Yale University Department of Statistics, New Haven, CT, February 2018.

Learning Interpretable Representations of the Microbiome
Yale University Department of Immunobiology, New Haven, CT, February 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
University of Michigan Department of Biostatistics, Ann Arbor, MI, February 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
University of Toronto Department of Statistics, Toronto, Canada, January 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Florida State University Department of Statistics, Tallahassee, FL, January 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
University of Virginia Department of Statistics, Charlottesville, VA, January 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
University of Oregon Department of Biology, OR, January 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Colorado State University Department of Statistics, Fort Collins, CO, January 2018.

Dimensionality Reduction with Structured Variables and Applications to the Microbiome
Columbia University Department of Statistics, New York, NY, January 2018.

Dimension Reduction with Structured Variables
Indiana University Department of Statistics, Bloomington, IN, January 2018.

Adaptive gPCA. Strategies and Techniques for Analyzing Microbial Population Structures
(STAMPS), Woods Hole, MA, August 2017.

Beyond UniFrac. Interdisciplinary Microbiome Perspectives, Stanford, CA, June 2017.

Contributed Presentations

Improved phylogenetic ordinations for microbiome data. Joint Statistical Meetings, Chicago,
IL, August 2016.

Phylogenetically informed analysis of microbiome data using adaptive gPCA in R, UseR, Stanford,
CA, June 2016.

Interpretable ordinations for microbiome data using sparse double principal coordinates analysis.
Biomedical Computation at Stanford, Stanford, CA, April 2015.

Comparisons of distance methods for combining covariates and abundances in microbiome studies.
Pacific Symposium on Biocomputing, Kona, HI, January 2012.

Teaching

Courses

Info I609: Advanced Seminar in Informatics (One of many co-instructors, responsible for
organizing one seminar)   Spring 2025

Stat S690: Statistical Consulting   Fall 2023

Stat S610: Introduction to Statistical Computing   Fall 2024, 2023, 2022, 2020, 2019

Stat S470/670: Exploratory Data Analysis	Spring 2025, 2023, 2022, 2021, Fall 2018
Stat S692: Internship in Statistical Consulting	Spring 2020
Stat S710: Introduction to Statistical Computing	Spring 2019
Strategies and Techniques for Analyzing Microbial Population Structures Marine Biological Lab, Woods Hole, MA Research Facilitator	August 2017

Mentoring

PhD Thesis primary advisor: Hongxuan Zhai (Statistics, graduated 2023), Justin Lin (Math, in progress), Paul Hunt (Statistics, in progress).

PhD Thesis committee: Hongxi Li, Madeleine Bartz, Sarthak Mishra (Computer Science, in progress), Genevieve Mortensen (Bioinformatics, in progress), Nicolas Castro-Perdomo (Geology, in progress), Yuanyuan Luan (Biostatistics, graduated 2024), Catherine Mortensen (Bioinformatics, in progress), Elizabeth Sherrill (Geology, graduated 2024), Anupama Chandroth (Geology, in progress), Miguel Trujillo (Statistics, graduated 2023), John Koo (Statistics, graduated 2023), Xixi Hu (Statistics, graduated 2019), William Shoemaker (Biology, graduated 2020), Fatma Parlak (Statistics, graduated 2023), Luis Miguel Mestre (Biostatistics, graduated 2023), Etienne Nzabarushimana (Bioinformatics, graduated 2022), Tony Lam

DA project primary advisor: Chelsie Parker (Statistics), Paul Hunt (Statistics)

Master's thesis: Sriram Sridhar (Statistics, graduated 2020)

Undergraduate research advisor: Benjamin Shores (Capstone Project in Data Science, graduated 2023), Morgan Fissel (Capstone Project in Data Science, graduated 2023), Carolina Coria (Summer Research Experience)

Mentor for the CAREER MODE program https://www.publichealth.columbia.edu/academics/non-degree-special-programs/professional-non-degree-programs/career-mode-program, 2023-2024 and 2024-2025.

Service to the profession

Associate Editor at *Biostatistics* since August 2024.

Referee for the *Annals of Applied Statistics*, *PLoS Computational Biology*, *Microbiome*, *Molecular Ecology Resources*, *PLoS Genetics*, *mSystems*, *Bioinformatics*, *Biometrika*, *Genome Biology*, *Journal*

of Applied Statistics, Nature Methods, Journal of Computational and Graphical Statistics, Journal of the American Statistical Association, BMC Bioinformatics, WIREs Computational Statistics, Bioinformatics Advances Journal of Computational Biology.

Awards

Provost's Travel Award for Women in Science	2022
Indiana University Trustees' Teaching Award	2022
Bio-X Stanford Interdisciplinary Graduate Fellowship	2014-2017
Yale College William R. Belknap Prize for Excellence in Biology	2010
Yale College Dean's Research Fellowship in the Sciences	2009
Yale College Fellowship for International Research in the Sciences and Health Studies	2008