Dr. John D. O'Brien

Updated: Oct. 2026

CONTACT INFORMATION

Department of Mathematics
Bowdoin College
8600 College Station
Brunswick, Maine 04011 USA

Voice: (+001) 207-798-4247
E-mail: jobrien@bowdoin.edu
Web: cascobayesian.github.io

CITIZENSHIP

USA

RESEARCH INTERESTS

Phylogenetics, Bayesian statistics, mixture models, statistical genetics, Markov chain Monte Carlo methods, statistical methods for social science

EDUCATION

University of California, Los Angeles, Los Angeles, CA 90095 USA

Ph.D., Biomathematics, 2008

- Thesis: Robust statistical methods for estimating the substitution distances and rates of change for genes from rapidly-evolving organisms
- Advisors: Professor Marc A. Suchard, Professor Zhen-Su She
- Area of Study: Statistics, genetics

M.Sc., Biomathematics, 2005

Pomona College, Claremont, CA 91711 USA

B.A., Mathematics (Honors), 2002

- Thesis: The topological symmetry groups of complete graphs
- Advisor: Professor Erica Flapan
- Area of Study: Knot theory

PROFESSIONAL HISTORY

Bowdoin College
Department of Mathematics
Searles Science Building
Brunswick, Maine 04011 USA

Associate Professor August 2018–present

- Instructor for statistics, data science, and Bayesian statistics
- Research work in genetics, sociology, and political science

Assistant Professor August 2012–July 2018

- Instructor for probability and biostatistics
- Research work in statistical genetics and ecology

University of Oxford
Centre for Genomics and Global Health
Wellcome Trust Centre for Human Genetics
Oxford, OX3 7BN UK

Senior Statistician June 2010–July 2012

- Implemented and tested next-generation sequencing pipelines for field-isolated *Plasmodium falciparum* genomes
- Using phylogenetic models to infer recombination, migration and gene flow for *Plasmodium falciparum*

University of Bristol
Department of Mathematics
Bristol, BS8 1TW UK

Research Associate in Statistical Genetics August 2008–June 2010

- Designed and implemented Bayesian methods for community metagenomics
- Lectured for *Introduction to Statistics*
- Co-organized `Statworks` conference on network inference

University of California, Los Angeles
Department of Biomathematics
Los Angeles, CA 90095 USA

Graduate Student September 2003–June 2008

- Extensive coursework in stochastic processes, statistics, genetics and ecology
- Teaching praeceptor for *Statistical Phylogenetics*
- Lecturer for *Calculus II* and *Introduction to Applied Mathematics*

Peking University
Center for Theoretical Biology
Beijing, PRC

Visiting Scholar Summers 2004, 2005 and 2006

- Developed the TREBLE algorithm for estimating the rate of substitution and phylogenies for time-sampled sequence data

West Land Competence Center
Bergen, Norway

Fulbright Scholar June 2002–August 2003

- Developed a dictionary of mathematics terms for Norwegian Sign Language

In draft

John O'Brien. "The topological analysis of mixture models."

John O'Brien and Aaron Kauffman. "The shape of local political space."

Peer-reviewed Publications

John D. O'Brien. 2026. "A Bayesian mixture model captures temporal and spatial structure of voting blocs within longitudinal referendum data." *Political Science Research and Methods*. 1–22. doi:10.1017/psrm.2025.10050. [link]

McBride, Felicia, Nur Schettino, John D. O'Brien, Wilder Harwood, Liz Perot, Grant Temple, Hugo Ayalo Solares et al. 2025. "A radio-quiet AGN as a candidate counterpart to neutrino event IceCube–200615A." *Monthly Notices of the Royal Astronomical Society*. 541(2):1613–1627. [link]

Dhiraj Murthy, John D. O'Brien, Alex Gross, Nathan Meyers. 2021. "Variations in the Temporal Structure of Sociability across American Cities." *Sociology.* 55 (1), 30-55 (21 pages) : ***Nominated for paper of the year by the British Sociological Association.***

Nicholas Record, John D. O'Brien, Karen Stamieszkin, and Jeffrey A. Runge. 2017. "Omic-style statistical clustering reveals old and new patterns in the Gulf of Maine ecosystem." *Canadian Journal of Fisheries and Aquatic Sciences*. 74(7):973–979. (7 pages)

CM Ndila, S Uyoga, AW Macharia, G Nyutu, N Peshu, J Ojal, M Shebe, *et al.* 2018. "Human candidate gene polymorphisms and risk of severe malaria in children in Kilifi, Kenya: a case-control association study." *The Lancet Haematology*. Volume 5, Issue 8, August 2018, Pages e333-e345. (13 pages)

John D. O'Brien, Lucas Amenga-Etego, and Ruiqi Li. 2016. "Approaches to estimating inbreeding coefficients in clinical isolates of *Plasmodium falciparum* from genomic sequence data." *Malaria Journal*. 15:473. (10 pages)

Alistair Miles, Zamin Iqbal, Paul Vauterin, Richard Pearson, Susanna Campino, Marie Theron, Keith Gould, Danial Mead, Elizabeth Drury, John D. O'Brien, Valentin Rubio, *et al.* 2016. "Indels, structural variation, and recombination drive genomic diversity in *Plasmodium falciparum*." *Genome Research.* 26(9):1288-99. (13 pages)

John D. O'Brien, Kathryn Lin, and Scott MacEachern. 2016. "Mixture model of pottery decorations from Lake Chad Basin archaeological sites reveals ancient segregation patterns." *Proceedings of the Royal Society B*. 283(1827):20152824. (8 pages)

John D. O'Brien, Zamin Iqbal, Jason Wendler and Lucas Amenga-Etego. 2016. "Inferring strain mixture within clinical *Plasmodium falciparum* isolates from genomic sequence data." *PLoS Computational Biology*. 12(6):e1004824. (20 pages)

Vladimir Douhovnikoff, Samuel Taylor, Eric Hazelton, Cindi Smith and John D. O'Brien. 2016. "Maximal stomatal conductance to water and plasticity in stomatal traits differ between native and invasive introduced lineages of *Phragmites australis* in North America." *AoB Plants*, 8:plw006. (11 pages)

John D. O'Brien, Xavier Didelot, Zamin Iqbal, Lucas Amenga-Etego, Bartu Ahiska, and Daniel Falush. 2014. "A Bayesian approach to inferring the phylogenetic structure of communities from metagenomic data." *Genetics*. 197(3):925–937. (13 pages)

Jason Wendler, John Okombo, Roberto Amato, Olivo Miotto, Steven Kiara, Leah Mwai, Lewa Pole, John D. O'Brien, Magnus Manske, Daniel Alcock, *et al.*, 2014. "A Genome-wide association study of *Plasmodium falciparum* susceptibility to 22 antimalarial drugs in Kenya." *PLoS ONE*. 9(5):e96486. (9 pages)

Olivo Miotto, Jacob Almagro-Garcia, Magnus Manske, Bronwyn MacInnis, Susana Campino, Kirk A. Rockett, Chanaki Amaratunga, *et al.* 2013. "Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia." *Nature Genetics*. 45(6) 648-655. (26 pages)

Samuel Oyola, Yong Gu, Magnus Manske, Thomas Otto, John D. O'Brien, Daniel Alcock, Bronwyn MacInnis, Matthew Berriman, Chris Newbold, Dominic Kwiatkowski, *et al.* 2013. "Efficient depletion of host DNA contamination in malaria clinical sequencing." *Journal of Clinical Microbiology*. 51(3):745-751. (7 pages)

Magnus Manske, Olivo Miotto, Susanna Campino, Sarah Auburn, Jacob Almagro-Garcia, Gareth Maslen, John D. O'Brien, Abdoulaye Djimde, *et al.* 2012. "Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing." *Nature*.

487(7407):375-9. (5 pages)

Vladimir Minin, John D. O'Brien and Arseni Seregin. 2011. "Imputation estimators partially correct for model misspecification." *Statistical Applications in Genetics and Molecular Biology*. 10(1). doi:10.2202/1544-6115.1650.

Dwayne Chambers, Erica Flapan and John D. O'Brien. 2011. "Topological symmetry groups of K_{4r+3}." *Discrete and Continuous Dynamical Systems – Series S*. 4(6):1401–1411. (11 pages)

John D. O'Brien, Vladimir N. Minin, and Marc A. Suchard. 2009. "Learning to count: robust estimates for labeled distances between molecular sequences." *Molecular Biology and Evolution*. 26(4):801–814. (14 pages)

John D. O'Brien, Zhen-Su She, and Marc A Suchard. 2008. "Dating the time of viral subtype divergence." *BMC Evolutionary Biology*. 8:172. (9 pages)

Zhu Yang, John D. O'Brien (joint first author), Xiao-Bin Zheng, Huai-Qu Zhu, and Zhen-Su She. 2007. "Tree and rate estimation by local evaluation of heterochronous nucleotide data." *Bioinformatics*. 23(2):169–176. (8 pages)

Greg Lakatos, John D. O'Brien, and Tom Chou. 2006. "Hydrodynamic mean field solutions of 1D exclusion processes with spatially varying hopping rates." *Journal of Physics A: Mathematical and General*. 39(10):2253-2264. (13 pages)

External Awards

Royal Society
- International Travel Grant, 2009

University of California
- Chancellor's Fellowship, 2003-2004 and 2007-2008

National Institutes of Health
- NIH Systems and Integrative Biology Fellowship, 2004-2007

United States Department of State
- Fulbright Scholar, 2002-2003

Pomona College, Department of Mathematics
- Honors Thesis, 2002

Courses taught

Bowdoin College

Biostatistics — Each semester : Fall 2012–Spring 2015, Fall 2017

- Accelerated introductory statistics with computation
- Intensive use of `R` computational environment
- Extensive outside readings in the social and natural sciences

Probability — Fall 2012, Spring 2013, Fall 2016, Fall 2018, Fall 2021

- Introductory probability course
- Proof-based development finishing with central limit theorem

- Introduction to `R` computational environment

Statistics Spring 2014 - Spring 2024

- A survey of techniques in statistics and probabilistic machine learning
- Covers theoretical statistical concepts and applied techniques for data analysis
- Intensive use of `R` computational environment

Advanced topics in statistics Fall 2013, Fall 2014, Fall 2017, Fall 2019, Fall 2023, Spring 2024

- Applied Bayesian statistics with computation
- Theoretical discussion of Bayesian statistical approach followed by development of Markov chain Monte Carlo techniques for practical inference
- Intensive use of `R` computational environment

University of Oxford

Instructor for Seminar in Statistics Spring 2011–Spring 2012

- Led weekly seminar for students and employees at the Wellcome Trust Center for Human Genetics
- Used data from existing research projects to develop course concepts

Instructor for Cape Town Workshop in Genetic Epidemiology March 2011

- Led short course for African researchers investigating the genetic basis of malaria
- Worked closely with several graduate students to develop research projects

University of Bristol

Instructor for Introductory Statistics Spring 2009, Spring 2010

- Taught eight week course in introductory statistics for biology students
- Used `R` computational environment for exploring course concepts

University of California, Los Angeles

Teaching Praeceptor for Statistical Phylogenetics Spring 2007

- Assisted with instruction for graduate course in phylogenetics methods
- Instructed lectures for unit on evolutionary models

Thesis Supervision

Nathan Kellerman Fall 2025–Spring 2026

- Honors thesis in mathematics: *"A Rapid Algorithm for Inferring Latent Mixture Structure in Replicate Social Science Data"*

Avery Cutler Fall 2025–Spring 2026

- Honors thesis in mathematics: *"Beyond the Docket: A Bayesian Exploration of Settlement Outcomes in Boston Evictions"*

Brian Liu Fall 2024-Spring 2025

- Honors thesis in mathematics: *"Chains for Change: Mathematical Modeling of Boston's Eviction Court Cases"*

Max Hukill Fall 2019-Spring 2021

- Honors thesis in mathematics: *"A Bayesian Hierarchical Mixture Model with Continuous-Time Markov Chains to Capture Bumblebee Foraging Behavior"*

Nan Ding Fall 2018–Spring 2019

- Honors thesis in mathematics: *"Rein Control Homeostasis and the Dynamics of Complex Ecosystem Stability"*

Jacob Spertus Fall 2014–Spring 2015

- Honors thesis in mathematics: *"A Bayesian Exploration of the Arabidopsis thaliana Chloroplast Genome"*
- Development of Markov chain Monte Carlo techniques to model haplotype structure in mixed genomes
- Broad readings from primary biological literature

Chunyi Zhao Fall 2014–Spring 2015

- Honors thesis in mathematics: *"The Card Swiper's Coloring Book: Bayesian Inference of Social Networks on Spatio-temporal Data"*
- Hidden Markov model for inferring social structures from dining hall swipe data
- Use of `R` for visualization and inference

Zoe Karp Fall 2013–Spring 2014

- Honors thesis in mathematics: *"Bayesian Parameter Estimation without Likelihoods: An Example in Lobster Heart Modeling"*
- Approximate Bayesian inference scheme for inferring parameters in differential system from experimental data

Nathan Meyers Fall 2012–Spring 2013

- Honors thesis in mathematics: *"The Twitterscape in Time and Space"*
- Cluster analysis for large-scale sociological spatio-temporal data
- Extensive use of `R` for visualization and hypothesis testing

Consulting projects

CLYNK January 2021 – June 2025

- Developed statistical techniques to move Maine's largest recycling program to an AI-assisted sampling program.
- Provided explanatory additional resources for legislators, businesses, and regulators.

Legal Consulting March 2017

- Investigated the structure of errors in administrative oversight
- Provided statistical summary of the probability of administrative errors

Expert Witness January 2014

- Created bibliography on measurement of skill in games of chance
- Provided testimony on the amount of skill involved in poker

Computing Environments

C++, C, Java, R, Python, Matlab, Unix, MacOS X, Windows

Presentations and workshops

Conference on Computational Social Science Philadelphia, PA, July 2024

- "The persistent structure of political ideology from referendum data" (*Reviewed*)

Symposium on Data Science and Statistics Richmond, VA, June 2024

- "The topology of mixture models" (*Reviewed*)

Colby Bates Bowdoin Mathematics Meeting Waterville, ME, October 2023

- "Inferring the shape of Maine's political ideology with referendum data" (*Invited*)

Political and Computational Social Science Cambridge, MA, June 2022

- "Inferring the structure of political ideology with a Bayesian model of referendum data" (*Contributed*)

Meeting of the Indian Statistical Association Gainesville, FL, August 2018

- "Bayesian mixtures of mixture for metagenomic data" (*Contributed*)

Joint Statistical Meeting Baltimore, MD, August 2017

- "All is everywhere: latent Dirichlet dispersion for ecological count data" (*Contributed*)

Joint Statistical Meeting Chicago, IL, August 2016

- "What's in that ecology? A latent factor Dirichlet-multinomial model for metagenomic count data" (*Contributed*)

Genetic Epidemiology of Malaria Workshop June 2015

- Working group: Merging population genetics with epidemiology (*Invited*)

Bigelow Laboratories for Ocean Science December 2014

- "Regression in complex microbial ecosystems" (*Invited*)

Joint Statistical Meeting August 2014

- "A Bayesian approach to inferring community structure and evolutionary history from whole-genome metagenomic data" (*Invited*)

Newton Institute Workshop on Metagenomics March 2014

- "The lineage model: A Bayesian approach to inferring ecology and evolution" (*Invited*)

Colby-Bates-Bowdoin Seminar November 2013

- "Measurably evolving ecologies: statistical methods for metagenomic sequence data" (*Invited*)

Joint Statistical Meeting August 2013

- "The lineage model: inferring ecoevolutionary dynamics from metagenomic data" (*Contributed*)

Genetic Epidemiology of Malaria Conference June 2012

- "Testing pipelines with crosses" (*Invited*)

International Biogeographic Society Conference September 2011

- "Towards a reconstruction of the historical migration patterns of *P. falciparum* through its organelles" (*Contributed*)

Bristol Complexity Seminar March 2010

- "Simultaneously estimating microbial ecology and evolution" (*Invited*)

Oxford Mathematical Biology Seminar November 2009

- "The lineage model for microbial community composition" (*Invited*)

Joint Statistical Meeting August 2009

- "The lineage model for microbial community composition" (*Contributed*)

Stochastic Processes and their Applications General Meeting August 2009

- "A more general model of the evolution of influenza subtypes" (*Contributed*)

Gordon Research Conference in Applied Microbiology July 2009

- "A Bayesian model for metagenomic population structure" (*Contributed*)

University of Bath June 2009

- "Phylogenetics and GMRFs to model conformational changes" (*Invited*)

European Bioinformatics Institute October 2008

- "Robust counting for unusual distances" (*Invited*)

Centers for Disease Control - Influenza Division April 2008

- "Modeling distance, rates and non-neutrality in viral systems" (*Invited*)

Joint Statistical Meeting July 2007

- "Counting codon changes carefully: Markov-chain-induced counting processes for robust distance estimation" (*Contributed*)

Peking University - Center for Theoretical Biology July 2006

- "Watching change in influenza A viruses" (*Invited*)

DIMACS Workshop on Rapidly Evolving Pathogens July 2006

- "Characterizing the rate of nucleotide substitution along branches in influenza phylogenies" (*Contributed*)

Society for Molecular Biology and Evolution General Meeting May 2006

- "Breaking the molecular clock (assumption)" (*Contributed*)

University of Bergen - Department of Linguistics March 2003

- "Cognitive aspects of the grammar of Norwegian Sign Language" (*Invited*)