

Alejandro Ochoa

Duke Center for Statistical Genetics and Genomics
Department of Biostatistics and Bioinformatics
Duke University Medical Center
Medical Science Research Building (MSRB) III
3 Genome Ct, Room 1241
Durham, NC 27710
alejandro.ochoa@duke.edu
<https://ochoalab.github.io/>

Education

- 2013 PhD in Molecular Biology, Princeton University
Dissertation: *Protein domain prediction using context statistics, the false discovery rate, and comparative genomics, with application to Plasmodium falciparum*
Advisers: Mona Singh (Computer Science) and Manuel Llinás (Molecular Biology)
- 2006 BS in Biology and Mathematics, Massachusetts Institute of Technology

Awards & Honors

- 2020 Whitehead Scholar, Whitehead Charitable Foundation. For Duke junior faculty with exceptional potential for research and research training in the biomedical sciences.
- 2008 NSF Graduate Research Fellowship. Computational Biology.
- 2008 Ford Foundation Diversity Fellowship Predoc. Comp. (Declined, listed as Honorable Mention.)
- 2006 MIT Department of Biology Merck Prize. Awarded to a graduating senior for outstanding research and academic performance in biophysics or bioinformatics.
- 2001 Mexican Math Olympiad Gold Medal. Awarded to top 15 competitors nationally.

Employment

- 2018-now *Assistant Professor*, Duke Center for Statistical Genetics and Genomics, Department of Biostatistics and Bioinformatics, Duke University
- 2013-2018 *Postdoctoral Research Associate* at John D Storey's group, Lewis-Sigler Institute for Integrative Genomics, and Center for Statistics and Machine Learning, Princeton University

Publications

JOURNAL ARTICLES

- 2021 **Alejandro Ochoa**, John D Storey. Estimating F_{ST} and kinship for arbitrary population structures. *PLoS Genet* 17(1): e1009241. PMID [33465078](https://pubmed.ncbi.nlm.nih.gov/33465078/).
- 2017 **Alejandro Ochoa**, Mona Singh. Domain prediction with probabilistic directional context. *Bioinf*

33(16) 2471-2478. PMID [28407137](#).

2016 Simon A Cobbold, Joana M Santos, **Alejandro Ochoa**, David H Perlman, Manuel Llinás. Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. *Sci Rep* 2016;6:19722. PMID [26813983](#).

2015 **Alejandro Ochoa**, John D Storey, Manuel Llinás, Mona Singh. Beyond the *E*-value: stratified statistics for protein domain prediction. *PLoS Comput Biol* 11 e1004509. PMID [26575353](#).

2013 Moriah L Szpara, Derek Gatherer, **Alejandro Ochoa**, Benjamin Greenbaum, Aidan Dolan, Rory J Bowden, Lynn W Enquist, Matthieu Legendre, Andrew J Davison. Evolution and diversity in human herpes simplex virus genomes. *J Virol* 88:1209-1227. PMID [24227835](#).

2011 **Alejandro Ochoa**, Manuel Llinás, Mona Singh. Using context to improve protein domain identification. *BMC Bioinformatics* 12:90. PMID [21453511](#).

2007 Gevorg Grigoryan, **Alejandro Ochoa**, Amy E Keating. Computing van der Waals energies in the context of the rotamer approximation. *Proteins* 68 (4), 863-78. PMID [17554777](#).

MANUSCRIPTS IN SUBMISSION

2020 Young-Sook Kim, Graham D Johnson, Jungkyun Seo, Alejandro Barrera, William H Majoros, **Alejandro Ochoa**, Andrew S Allen, Timothy E Reddy. Detecting regulatory elements in high-throughput reporter assays. Preprint: <https://doi.org/10.1101/2020.08.07.241901>

2019 Yiqi Yao, **Alejandro Ochoa**. Testing the effectiveness of principal components in adjusting for relatedness in genetic association studies. Preprint: <https://doi.org/10.1101/858399>

2019 **Alejandro Ochoa**, John D Storey. New kinship and F_{ST} estimates reveal higher levels of differentiation in the global human population. Preprint: <https://doi.org/10.1101/653279>

2016 **Alejandro Ochoa**, John D Storey. F_{ST} and kinship for arbitrary population structures I: Generalized definitions. Preprint: <https://doi.org/10.1101/083915>

Distributed software

2020 ligera: LIght GEnetic Robust Association. R, C++.
<https://github.com/OchoaLab/ligera>.

2019 human-differentiation-manuscript: Human differentiation analysis. R, bash, markdown.
<https://github.com/StoreyLab/human-differentiation-manuscript>.

2019-2020 genio: Genetics Input/Output Functions. R, C++.
Available on CRAN and <https://github.com/OchoaLab/genio>.

2019-2020 simtrait: Simulate Complex Traits from Genotypes. R.
<https://github.com/OchoaLab/simtrait>.

2019-2020 popkinsuppl: Supplement to "popkin" package. R.
<https://github.com/OchoaLab/popkinsuppl>.

2017-2020 popkin: Estimate Kinship and FST under Arbitrary Population Structure. R, C++.
Available on CRAN and <https://github.com/StoreyLab/popkin>.

2017-2020 bnpsd: Model and Simulate Admixed Populations. R.
Available on CRAN and <https://github.com/StoreyLab/bnpsd>.

2014-2020 dPUC2: Domain Prediction Using Context, Version 2. Perl, C.
<https://github.com/alexviia/dpuc2>.

- 2014-2020 DomStratStats: Domain Stratified Statistics (q-values and local FDRs). Perl.
<https://github.com/alexviia/DomStratStats>.
- 2014-2019 RandProt: High-order Markov random models for protein sequences. Perl.
<https://github.com/alexviia/RandProt>.

Invited Talks

- 2020 *University Program in Genetics and Genomics (UPGG) seminar*. Room 147, Nanaline Duke building, Duke University, Durham, NC. 2020-02-18.
- 2019 *Computational Biology and Bioinformatics (CBB) retreat*. Holiday Inn Resort, Wrightsville Beach, NC. 2019-09-28.
- 2019 *Computational Biology and Bioinformatics (CBB) seminar*. French Family Science Center, Duke University, Durham, NC. 2019-02-04.
- 2019 *4th Mexico Population Genomics Meeting*. Amoxcalli Complex, Faculty of Sciences, University City, National Autonomous University of Mexico, Mexico City, DF, Mexico. 2019-01-11.
- 2018 Population Biology seminar, Department of Biology, Duke University, Durham, NC. 2018-09-27.
- 2018 Department of Biostatistics and Bioinformatics, Duke University, Durham, NC. 2018-02-21.
- 2018 Department of Biostatistics, Johns Hopkins University, Baltimore, MD. 2018-01-22.
- 2017 Department of Biology, University of Richmond, Richmond, VA. 2017-11-10.
- 2017 *Princeton Research Day*. Princeton University, Princeton, NJ. 2017-05-11.
- 2017 Department of Genetics, University of North Carolina, Chapel Hill, NC. 2017-02-13.
- 2016 *New York Area Population Genomics Workshop 2016*. Princeton University, Princeton, NJ. 2016-01-21.
- 2015 *Probabilistic Modeling in Genomics conference*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 2015-10-15.
- 2013 *Telepresentation for Yun Song's group*. UC Berkeley, Berkeley, CA. 2013-04-17.
- 2013 *Biological sequence analysis and probabilistic models conference*. HHMI Janelia Farm, Ashburn, VA. 2013-03-25.
- 2013 NCBI, NIH, Bethesda, MD. 2013-02-25.
- 2012 *Recruiting conference*. Department of Computer Science, Princeton University, Princeton, NJ. 2012-03-01.
- 2004 Rotary Club Paso del Norte, Ciudad Juarez, CHIH, Mexico. 2004-08-26.

Conference posters

- 2020 **Alejandro Ochoa**, John D Storey. New kinship and F_{ST} estimates applied to the global human population. *American Society of Human Genetics (ASHG) Annual Meeting*. Virtual.
- 2019 **Alejandro Ochoa**, John D Storey. Relatedness and Differentiation in Arbitrary Population Structures. *4th Mexico Population Genomics Meeting*. Amoxcalli Complex, Faculty of Sciences, University City, National Autonomous University of Mexico, Mexico City, DF, Mexico.
- 2018 **Alejandro Ochoa**, John D Storey. Relatedness and Differentiation in Arbitrary Population Structures. *Probabilistic Modeling in Genomics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

- 2018 **Alejandro Ochoa**, John D Storey. Relatedness and Differentiation in Arbitrary Population Structures. *Population, Evolutionary and Quantitative Genetics Conference*. Madison Concourse Hotel, Madison, WI.
- 2016 **Alejandro Ochoa**, John D Storey, Srikanth Gottipati, Shashank Rohatagi, Andrew Forbes, Deborah Profit, Raymond Sanchez, Margaretta Nyilas, William Carson. Mixed Effects Modeling of Placebo Response Across 8 Placebo-controlled Aripiprazole Trials Spanning 20 Years in Acutely Relapsed Schizophrenia Patients. *Neuroscience Education Institute Psychopharmacology Congress*. Broadmoor Convention Center, Colorado Springs, CO.
- 2016 Srikanth Gottipati, **Alejandro Ochoa**, Shashank Rohatagi, Andrew Forbes, Deborah Profit, Raymond Sanchez, Margaretta Nyilas, William Carson. Canonical Loadings of PANSS Subscales Show Differential Placebo and Aripiprazole Drug Responses in Schizophrenia Patients. *Neuroscience Education Institute Psychopharmacology Congress*. Broadmoor Convention Center, Colorado Springs, CO.
- 2015 **Alejandro Ochoa**, John D Storey. F_{ST} generalized for arbitrary population structures. *John W. Tukey 100th Birthday Celebration conference*. Center for Statistics and Machine Learning, Princeton University, Princeton, NJ.
- 2013 **Alejandro Ochoa**, John D Storey, Manuel Llinás, Mona Singh. Forget the E -value: family-based q -values for protein domain prediction, and empirical error detection. *Biological sequence analysis and probabilistic models conference*. HHMI Janelia Farm, Ashburn, VA.
- 2010 **Alejandro Ochoa**, Manuel Llinás, Mona Singh. Using context to predict protein domains across diverse organisms. *Recomb Systems Biology conference*. Columbia University, New York, NY.

Teaching

PRINCETON UNIVERSITY

- 2017 Intro to Genomics and Comp Bio. Quant Comp Bio, Comp Sci. (Co-Lecturer Fall 2017)
- 2016-2017 Discussion Leader. Summer Undergraduate Research Program (Summer 2016, 2017)
- 2014-2017 Statistical Treatment of Data. Mol Bio, Quant Comp Bio. (Workshop Fall 2014, 2015, 2016, 2017)
- 2011 Intro to Genomics and Comp Mol Bio. Mol Bio, Comp Sci. (TA Fall 2011)
- 2008 Core Laboratory. Mol Bio. (TA Spring 2008)

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

- 2005 Calculus. OME Project Interphase (TA Summer 2005)
- 2003-2004 Multivar Calculus and Diff Eqs. OME Seminar XL (Tutor Fall 2003, Spring 2004)
- 2003-2004 Calculus. MITE2S Program (TA Summer 2003, 2004)

Mentoring

- 2019 Yiqi Yao (Duke), Master's student, Ochoa lab. Theoretical and empirical evaluation of genetic association study approaches.
- 2011 Sebastian Nasamu (MIT), summer student, Llinás lab. Cloned DNA of malaria protein domains, purified proteins, used protein-binding microarrays to test for DNA binding.

- 2010 Jeremy Bigness (Princeton), graduate rotation, Singh lab. Built HMMs of ortholog proteins.
- 2009 Neo Chung (Princeton), graduate rotation, Llinás lab. Wrote scripts to analyze sequencing data.

Research experience

- 2013-2018 Postdoctoral Research, Storey Lab, Princeton. Developed new definitions and tools to study arbitrary population structures. Also applied GWAS and mixed effects modeling to longitudinal psychiatric drug clinical trials.
- 2007-2013 Graduate Research Thesis, Singh Lab and Llinás Lab, Princeton. Developed probabilistic models and statistical methods that improve protein domain prediction. Studied the *Plasmodium falciparum* proteome, focusing on the AP2 transcription factors. Used experimental techniques (cloning, protein purification and protein-binding microarrays) to test predicted AP2 domains for DNA binding.
- 2007 Graduate Rotation, Tavazoie Lab, Princeton. Cloned an RNA aptamer for a phage display assay.
- 2006-2007 Graduate Rotation, Singh Lab, Princeton. Analyzed protein interactions predicted from the presence of interacting domain pairs.
- 2006 Graduate Rotation, Troyanskaya Lab, Princeton. Built and analyzed a potential "gold standard" of Gene Ontology terms integrated from predictions of multiple sources, in *Saccharomyces cerevisiae* and *Homo sapiens*.
- 2005-2006 Undergraduate Research Assistant, Keating Lab, MIT. Analyzed full-atom computational protein designs using modified van der Waals potentials.

Extracurricular University Activities

- 2006-2010 Latino Graduate Student Association, Princeton University. Board member: Technology Specialist 2007-2010 (handled website, email list, photos, calendar).
- 2002-2006 Association of Puerto Rican students, MIT. Regular member.

Outreach

- 2017 Panelist at HISPA Latinos in College Conference. Princeton University.
- 2010 Ivy-plus recruiting fair at University of Puerto Rico (4 campuses) for Princeton.
- 2008-2011 Science and Engineering Expo. HHMI and Princeton University.
- 2007 Helped local high school students with college personal statements. PUPP, Princeton University.

Languages

English. Native proficiency.
 Spanish. Native proficiency.
 French. Fluent reading and writing, conversational speaking.

Last updated: January 21, 2021