

## Curriculum Vitae

### Biographical Information

Name: Prof. Carlos D. Bustamante, Ph.D.  
Current Positions: Venture Partner, F-Prime Capital Partners  
Investigator, Chang-Zuckerberg BioHub  
Prof. of Biomedical Data Science, Genetics, and (by courtesy) Biology, Stanford University (on leave through January, 2021)  
Updated: November, 2019

### Contact information

Department of Biomedical Data Science  
Stanford School of Medicine  
1265 Welch Road  
Medical School Office Building, x385  
Stanford, CA 94305-5464

Assistant: +1-650-721-5338  
Direct: +1-650-497-0771  
Fax: +1-650-497-1990  
Email: [cdbustam@stanford.edu](mailto:cdbustam@stanford.edu)

### Education

Harvard University, Cambridge, MA	B.A.	1997	Biology, <i>magna cum laude</i> with highest honors
Harvard University, Cambridge, MA	M.S.	2001	Statistics
Harvard University, Cambridge, MA	Ph.D.	2001	Biology
University of Oxford, Oxford, UK	Post-doc	2001-2	Mathematical Genetics

### Personal Statement

For the past 17 years, I have led a multidisciplinary team working on problems at the interface of computational and biological sciences. Much of our research has focused on genomics technology and its application in medicine, agriculture, and evolutionary biology. My first academic appointment was at Cornell University's College of Agriculture and Life Sciences. There, much of our work focused on population genetics and agricultural genomics motivated by a desire to improve the foods we eat and the lives of the animals upon which we depend. I moved to Stanford in 2010 to focus on enabling clinical and medical genomics on a global scale. I have been focused on reducing health disparities in genomics by: (1) calling attention to the problem raised by >95% of participants in large scale studies being of European descent; and (2) broadening representation of understudied groups in large NIH funded consortia, particularly minority groups from the U.S., the Americas, and Africa. My work has empowered decision-makers to utilize genomics and data science in the service of improving human health and wellbeing. In the next phase of my career, I will focus on opportunities for bringing these technologies to consumer and patients, directly, where this work can have the greatest impact.

I have a strong interest in building new academic units, non-profits, and companies. I was the Inaugural Chair of the Department of Biomedical Data Science—the first new department that Stanford has started in 14 years—and I was Founding Director (with Marc Feldman) of the Center for Computational, Evolutionary, and Human Genomics. I serve as an advisor to the US federal government, private companies, startups, and non-profits in the areas of computational genomics, population and medical genetics, veterinary and plant genomics, and business strategy. I am the founder of Arc-Bio, a pathogen detection company, a board member of the parent company (Eden Roc Sciences LLC) and current chair of its Scientific Advisory Board. I also consult for and invest in early stage biotechnology companies through my consulting firm, CDB Consulting, and as a Venture Partner with Fidelity's early stage biotechnology and healthcare group, F-Prime Capital Partners.

## **Positions and Employment**

2019-present	F-Prime Capital Partners, Venture Partner
2017-present	Investigator, Chan Zuckerberg Initiative (CZI) Biohub
2016-present	Principal and Chairman, CDB Consulting, LTD.
2015-present	Professor, Department of Biomedical Data Sciences, Stanford University
2010-present	Professor, Department of Genetics and (by courtesy) Biology, Stanford University
2015-2019	Inaugural Chair, Department of Biomedical Data Science, Stanford University
2012-2017	Founding Director, Stanford Center for Computational, Evolutionary, and Human Genomics
2012-2015	Director of Informatics, Stanford Center for Genomics and Personalized Medicine
2008-2010	Professor, Department of Biological Statistics and Computational Biology and Department of Statistical Sciences, Cornell University
2002-2008	Assistant Professor, Department of Biological Statistics and Computational Biology and Department of Statistical Sciences, Cornell University
2007-8	Associate Director, Cornell Center for Comparative and Population Genomics
2004	Visiting Scholar, NSF Institute for Pure and Applied Mathematics, Los Angeles

## **Fellowships and Honors**

2017	Chang-Zuckerberg Biohub Investigator
2016	Stanford Prize in Population Genetics and Society
2011-2016	2010 John D. & Catherine T. MacArthur Fellow
2008	Cornell Provost Award for Distinguished Research
2007-2009	Sloan Research Fellowship in Molecular Biology
2001-2002	Marshall-Sherfield Fellowship
1999-2001	Howard Hughes Medical Institute Graduate Research Fellowship
1997-1998	National Science Foundation Graduate Research Fellowship

## **Editorial Boards**

2013-2016	Evolution Section Editor, <i>PLoS-Genetics</i>
2013-2016	Editorial Board, <i>American Journal of Human Genetics</i>
2010-present	Editorial Board, <i>Human Biology</i>
2008-2017	Editorial Board, <i>Genome Research</i>
2008-2012	Editorial Board, <i>Molecular Biology and Evolution</i>
2005-2009	Editorial Board, <i>Bioinformatics</i>

## **Other Experience and Professional Memberships**

2015-present	Co-Chair (2015) and Chair (2017), Gordon Conference on Human Genetics and Genomics
2012-present	Co-Chair, Conference on Genomics and Health Disparities
2012-2015	Organizer, Cold Spring Harbor Biology of Genomes Conference
2012-2014	Organizer, Meeting of the Society for Molecular Biology and Evolution
2009-present	Member, The Center for Academic Research and Training in Anthropogeny (UCSD)
2004-2005	Fellow of NSF Institute for Pure and Applied Mathematics

## **Non-profit and Government Service**

2019-present	Chair, ClinGen Complex Disease Working Group
2015-present	Co-Chair, PAGE Analysis Working Group
2015-present	Co-Chair, ClinGen Ancestry and Diversity Working Group
	Chair, ClinGen Informatics Working Group

Member, ClinGen Sequence Variant Interpretation Committee  
 Member, ClinGen Steering Committee  
 Member, NHGRI Neuropsychiatric and Developmental Disorders Working Group  
 2011-2015 Member, National Human Genome Research Institute (NHGRI) Council  
 2011-2015 National Institutes of Health Council of Councils  
 2011-2014 External Evaluation committee NIDDK T2D GENES project  
 2010-2014 Scientific Advisory Board, SIGMA (Slim Initiative for Genomic Medicine in the Americas)  
 2009-present Scientific Advisory Board, NHGRI Center for Research on Genomics and Global Health

**Corporate Leadership**

2015-present Founder, Director, and Chair of Scientific Advisory Board, Arc-Bio, LLC: formerly Big Data Bio LLC (Boston, MA) and Identify Genomics (Menlo Park, CA)  
 2014-2016 Founder, Director, and Chair of Scientific Advisory Board, Identify Genomics LLC (Menlo Park, CA)  
 2015-present Director, Liberty Biosecurity (Washington, DC)  
 2016-present Director, Eden Roc Sciences LLC (Birmingham, MI)  
 2018-present Director, Etalon Diagnostics (Palo Alto, CA)  
 2019-present Board observer, Genomics PLC (Oxford, UK)  
 2019-present Board observer, Embark Veterinary, Inc. (Boston, MA)  
 2016-present Principal and Chairman, CDB Consulting LTD (Redwood City, CA)

**Scientific and Business Advisory Boards (current)**

2018-present Scientific and Business Advisory Boards, Awakens, Inc. (Berkeley, CA)  
 2018-present Scientific and Business Advisory Boards, Imprimed, Inc. (Palo Alto, CA)  
 2018-present Scientific Advisory Board, Columbia Care LLC (New York, NY)  
 2018-present Scientific Advisor, Nicklaus Children’s Hospital (Miami, FL)  
 2018-present Scientific and Business Advisory Boards, Fauna Bio, Inc. (San Francisco, CA)  
 2017-present Scientific Advisory Board, LunaDNA, Inc. (San Diego, CA)  
 2017-present Fellow, WebShield, Inc. (San Francisco, CA)  
 2016-present Scientific Advisory Board, Embark Veterinary (Boston, MA)  
 2016-present Scientific Advisory Board, Digitalis Ventures (New York, NY)  
 2015-present Scientific Advisory Board, Liberty Biosecurity (Washington, DC)  
 2013-present Scientific Advisory Board, Etalon Diagnostics (Menlo Park, CA)

**Consulting and Industry Service (past)**

2011-2019 Scientific Advisory Board, Ancestry.com (Provo, UT)  
 2016-2018 Mars/IBM Food Safety Board (McClean, VA)  
 2013-2018 Chair of Scientific Advisory Board, Etalon Diagnostics (Menlo Park, CA)  
 2011-2017 Scientific Advisory Board, 23andme “Roots into the Future” (Mountain View, CA)  
 2013-2015 Scientific Advisory Board, National Geographic Genographic Project (Washington, D.C.)  
 2012-2016 Scientific Advisory Board, Personalis, Inc. (Menlo Park, CA)  
 2012-2015 Scientific Advisory Board, InVita: formerly Locus Development, Inc. (San Francisco, CA)  
 2012-2014 Medical Advisory Board, Med-Tek (Miami, FL)  
 2011-2012 Scientific Advisory Board, Mubadala Healthcare (Abu Dhabi, UAE)  
 2009-2012 Scientific Advisory Board, Nature Source Genetics (Ithaca, NY)

## Patents and Patent Applications

1. COMPOSITIONS AND METHODS FOR TARGETED DEPLETION, ENRICHMENT, AND PARTITIONING OF NUCLEIC ACIDS USING CRISPR/CAS SYSTEM PROTEINS. Filing Date: 19.12.2015. Inventors: Meredith Carpenter, Carlos D. Bustamante, Stephane B. Gourguechon. Pub. No.: Wo/2016/100955.
2. ENRICHMENT OF DNA SEQUENCING LIBRARIES FROM SAMPLES CONTAINING SMALL AMOUNTS OF TARGET DNA. Pub. date: 17.12.2015. Inventors: Carlos D. Bustamante, Meredith Carpenter, Jason Buenrostro, and William J. Greenleaf. Pub. no: US 2015/0360194 A1.
3. PARTITIONING OF DNA SEQUENCING LIBRARIES INTO HOST AND MICROBIAL COMPONENTS. Pub. date: 20.08.2015. Inventors: Carlos D. Bustamante and Meredith Carpenter. Pub. no: US 2015/0232834 A1.
4. TECHNIQUES FOR DETERMINING HAPLOTYPE BY POPULATION GENOTYPE AND SEQUENCE DATA. Pub. date: 13.02.2014. Inventors: Carlos D. Bustamante and Fouad Zakharia. Pub. no: US 2014/0045705 A1.
5. PHASED WHOLE GENOME GENETIC RISK IN A FAMILY QUARTET. Pub. Date: March 28, 2013. Inventors: Frederick Dewey, Euan A. Ashley, Jake Byrnes, Carlos Daniel Bustamante, Atul J. Butte, Rong Chen. Pub no: 2015/0370959. Grant Number: 09443056 . Grant Date: 13.09.2016

## Publications

### 2019

225. Nieves-Colón MA, Pestle WJ, Reynolds AW, Llamas B, de la Fuente C, Fowler K, Skerry KM, Crespo-Torres E, **Bustamante CD**, Stone AC. [Ancient DNA reconstructs the genetic legacies of pre-contact Puerto Rico communities](#). *Mol Biol Evol*. pii: msz267. doi: 10.1093/molbev/msz267. Epub 2019 Nov 9. PMID: 31710665
224. Musharoff S, Shringarpure S, **Bustamante CD**, Ramachandran S. [The inference of sex-biased human demography from whole-genome data](#). *PLoS Genet*. 2019 Sep 20;15(9):e1008293. doi: 10.1371/journal.pgen.1008293. eCollection 2019 Sep. PMID: 31539367
223. Lopez M, Choin J, Sikora M, Siddle K, Harmant C, Costa HA, Silvert M, Mouguiama-Daouda P, Hombert JM, Froment A, Le Bomin S, Perry GH, Barreiro LB, **Bustamante CD**, Verdu P, Patin E, Quintana-Murci L. [Genomic Evidence for Local Adaptation of Hunter-Gatherers to the African Rainforest](#). *Curr Biol*. 2019 Sep 9;29(17):2926-2935.e4. doi: 10.1016/j.cub.2019.07.013. Epub 2019 Aug 8. PMID: 31402299
222. [Genetic analyses of diverse populations improves discovery for complex traits](#). Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, ... , **Bustamante CD**, Caberto C, Canizales-Quinteros S,..., Kooperberg C, Le Marchand L, Loos RJF, Matisse TC, North KE, Peters U, Kenny EE, Carlson CS. *Nature*. 2019 Jun;570(7762):514-518. doi: 10.1038/s41586-019-1310-4. Epub 2019 Jun 19. PMID: 31217584

221. Bucio D, Ormond KE, Hernandez D, **Bustamante CD**, Lopez Pineda A. [A genetic counseling needs assessment of Mexico](#). *Mol Genet Genomic Med*. 2019 May;7(5):e668. doi: 10.1002/mgg3.668. Epub 2019 Apr 1. PMID: 30938092
220. Fregel R, Ordóñez AC, Santana-Cabrera J, Cabrera VM, Velasco-Vázquez J, Alberto V, Moreno-Benítez MA, Delgado-Darias T, Rodríguez-Rodríguez A, Hernández JC, Pais J, González-Montelongo R, Lorenzo-Salazar JM, Flores C, Cruz-de-Mercadal MC, Álvarez-Rodríguez N, Shapiro B, Arnay M, **Bustamante CD**. [Mitogenomes illuminate the origin and migration patterns of the indigenous people of the Canary Islands](#). *PLoS One*. 2019 Mar 20;14(3):e0209125. doi: 10.1371/journal.pone.0209125. eCollection 2019. PMID:30893316
219. Gignoux CR, Torgerson DG, Pino-Yanes M, Uricchio LH, Galanter J, Roth LA, Eng C, Hu D, Nguyen EA, Huntsman S, Mathias RA, Kumar R, Rodriguez-Santana J, Thakur N, Oh SS, McGarry M, Moreno-Estrada A, Sandoval K, Winkler CA, Seibold MA, Padhukasahasram B, Conti DV, Farber HJ, Avila P, Brigino-Buenaventura E, Lenoir M, Meade K, Serebrisky D, Borrell LN, Rodriguez-Cintrón W, Thyne S, Joubert BR, Romieu I, Levin AM, Sienra-Monge JJ, Del Rio-Navarro BE, Gan W, Raby BA, Weiss ST, Bleecker E, Meyers DA, Martinez FJ, Gauderman WJ, Gilliland F, London SJ, **Bustamante CD**, Nicolae DL, Ober C, Sen S, Barnes K, Williams LK, Hernandez RD, Burchard EG. [An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos](#). *J Allergy Clin Immunol*. 2019 Mar;143(3):957-969. doi: 10.1016/j.jaci.2016.08.057. Epub 2018 Sep 7. PMID: 30201514
218. Li Q, Zhao K, **Bustamante CD**, Ma X, Wong WH. [Xrare: a machine learning method jointly modeling phenotypes and genetic evidence for rare disease diagnosis](#). *Genet Med*. 2019 Sep;21(9):2126-2134. doi: 10.1038/s41436-019-0439-8. Epub 2019 Jan 24. PMID: 3067503
- 2018**
217. Troll CJ, Putnam NH, Hartley PD, Rice B, Blanchette M, Siddiqui S, Ganbat JO, Powers MP, Ramakrishnan R, Kunder CA, **Bustamante CD**, Zehnder JL, Green RE, Costa HA. [Structural Variation Detection by Proximity Ligation from Formalin-Fixed, Paraffin-Embedded Tumor Tissue](#). *J Mol Diagn*. 2018 Dec 31. pii: S1525-1578(18)30172-7. doi: 10.1016/j.jmoldx.2018.11.003. [Epub ahead of print] PMID: 30605765
216. De La Vega FM, **Bustamante CD**. [Polygenic risk scores: a biased prediction?](#) *Genome Med*. 2018 Dec 27;10(1):100. doi: 10.1186/s13073-018-0610-x. PMID: 30591078
215. Lin M, Siford RL, Martin AR, Nakagome S, Möller M, Hoal EG, **Bustamante CD**, Gignoux CR, Henn BM. [Rapid evolution of a skin-lightening allele in southern African KhoeSan](#). *Proc Natl Acad Sci U S A*. 2018 Dec 26;115(52):13324-13329. doi: 10.1073/pnas.1801948115. Epub 2018 Dec 10. PMID: 30530665
214. Huddart R, Fohner AE, Whirl-Carrillo M, Wojcik GL, Gignoux CR, Popejoy AB, **Bustamante CD**, Altman RB, Klein TE. [Standardized Biogeographic Grouping System for Annotating Populations in Pharmacogenetic Research](#). *Clin Pharmacol Ther*. 2018 Dec 1. doi: 10.1002/cpt.1322. PMID: 30506572
213. Jha AR, Davenport ER, Gautam Y, Bhandari D, Tandukar S, Ng KM, Fragiadakis GK,

- Holmes S, Gautam GP, Leach J, Sherchand JB, **Bustamante CD**, Sonnenburg JL. [Gut microbiome transition across a lifestyle gradient in Himalaya](#). **PLoS Biol.** 2018 Nov 15;16(11):e2005396. doi: 10.1371/journal.pbio.2005396. eCollection 2018 Nov. PMID: 30439937
212. Popejoy AB, Ritter DI, Crooks K, Currey E, Fullerton SM, Hindorff LA, Koenig B, Ramos EM, Sorokin EP, Wand H, Wright MW, Zou J, Gignoux CR, Bonham VL, Plon SE, **Bustamante CD**; Clinical Genome Resource (ClinGen) Ancestry and Diversity Working Group (ADWG). [The clinical imperative for inclusivity: Race, ethnicity, and ancestry \(REA\) in genomics](#). **Hum Mutat.** 2018 Nov;39(11):1713-1720. doi: 10.1002/humu.23644. PMID:30311373
211. Cornejo OE, Yee MC, Dominguez V, Andrews M, Sockell A, Strandberg E, Livingstone D 3rd, Stack C, Romero A, Umaharan P, Royaert S, Tawari NR, Ng P, Gutierrez O, Phillips W, Mockaitis K, **Bustamante CD**, Motamayor JC. [Population genomic analyses of the chocolate tree, \*Theobroma cacao\* L., provide insights into its domestication process](#). **Commun Biol.** 2018 Oct 16;1:167. doi: 10.1038/s42003-018-0168-6. eCollection 2018. PMID:30345393
210. Ioannidis NM, Wang W, Furlotte NA, Hinds DA; 23andMe Research Team, **Bustamante CD**, Jorgenson E, Asgari MM, Whittemore AS. [Gene expression imputation identifies candidate genes and susceptibility loci associated with cutaneous squamous cell carcinoma](#). **Nat Commun.** 2018 Oct 15;9(1):4264. doi: 10.1038/s41467-018-06149-6. PMID:30323283
209. Gignoux CR, Torgerson DG, Pino-Yanes M, Uricchio LH, Galanter J, Roth LA, Eng C, Hu D, Nguyen EA, Huntsman S, Mathias RA, Kumar R, Rodriguez-Santana J, Thakur N, Oh SS, McGarry M, Moreno-Estrada A, Sandoval K, Winkler CA, Seibold MA, Padhukasahasram B, Conti DV, Farber HJ, Avila P, Brigino-Buenaventura E, Lenoir M, Meade K, Serebrisky D, Borrell LN, Rodriguez-Cintron W, Thyne S, Joubert BR, Romieu I, Levin AM, Sienna-Monge JJ, Del Rio-Navarro BE, Gan W, Raby BA, Weiss ST, Bleecker E, Meyers DA, Martinez FJ, Gauderman WJ, Gilliland F, London SJ, **Bustamante CD**, Nicolae DL, Ober C, Sen S, Barnes K, Williams LK, Hernandez RD, Burchard EG. [An admixture mapping meta-analysis implicates genetic variation at 18q21 with asthma susceptibility in Latinos](#). **J Allergy Clin Immunol.** 2018 Sep 7. pii: S0091-6749(18)31274-0. doi: 10.1016/j.jaci.2016.08.057. PMID: 30201514
208. López-Pineda A, Rodríguez-Moran MF, Álvarez-Aguilar C, Fuentes Valle SM, Acosta-Rosales R, Bhatt AS, Sheth SN, **Bustamante CD**. [Data mining of digitized health records in a resource-constrained setting reveals that timely immunophenotyping is associated with improved breast cancer outcomes](#). **BMC Cancer.** 2018 Sep 27;18(1):933. doi: 10.1186/s12885-018-4833-4. PMID: 30261931
207. Wojcik GL, Fuchsberger C, Taliun D, Welch R, Martin AR, Shringarpure S, Carlson CS, Abecasis G, Kang HM, Boehnke M, **Bustamante CD**, Gignoux CR, Kenny EE. [Imputation-Aware Tag SNP Selection To Improve Power for Large-Scale, Multi-ethnic Association Studies](#). **G3** (Bethesda). 2018 Aug 21. pii: g3.200502.2018. doi: 10.1534/g3.118.200502. PMID: 30131328
206. Cruz-Dávalos DI, Nieves-Colón MA, Sockell A, Poznik GD, Schroeder H, Stone AC, **Bustamante CD**, Malaspinas AS, Ávila-Arcos MC. [In-solution Y-chromosome capture-enrichment on ancient DNA libraries](#). **BMC Genomics.** 2018 Aug 14;19(1):608. doi:

205. Wang B, Pourshafeie A, Zitnik M, Zhu J, **Bustamante CD**, Batzoglou S, Leskovec J. [Network enhancement as a general method to denoise weighted biological networks.](#) **Nat Commun.** 2018 Aug 6;9(1):3108. doi: 10.1038/s41467-018-05469-x. PMID: 30082777
204. Fregel R, Méndez FL, Bokbot Y, Martín-Socas D, Camalich-Massieu MD, Santana J, Morales J, Ávila-Arcos MC, Underhill PA, Shapiro B, Wojcik G, Rasmussen M, Soares AER, Kapp J, Sockell A, Rodríguez-Santos FJ, Mikdad A, Trujillo-Mederos A, **Bustamante CD**. [Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe..](#) **Proc Natl Acad Sci U S A.** 2018 Jun 26;115(26):6774-6779. doi: 10.1073/pnas.1800851115. Epub 2018 Jun 12. PMID: 29895688
203. Waszak SM, Northcott PA, Buchhalter I, Robinson GW, Sutter C, Groebner S, Grund KB, Brugières L, Jones DTW, Pajtler KW, Morrissy AS, Kool M, Sturm D, Chavez L, Ernst A, Brabetz S, Hain M, Zichner T, Segura-Wang M, Weischenfeldt J, Rausch T, Mardin BR, Zhou X, Baciuc C, Lawerenz C, Chan JA, Varlet P, Guerrini-Rousseau L, Fults DW, Grajkowska W, Hauser P, Jabado N, Ra YS, Zitterbart K, Shringarpure SS, De La Vega FM, **Bustamante CD**, Ng HK, Perry A, MacDonald TJ, Hernáiz Driever P, Bendel AE, Bowers DC, McCowage G, Chintagumpala MM, Cohn R, Hassall T, Fleischhack G, Eggen T, Wesenberg F, Feychting M, Lannering B, Schüz J, Johansen C, Andersen TV, Rösli M, Kuehni CE, Grotzer M, Kjaerheim K, Monoranu CM, Archer TC, Duke E, Pomeroy SL, Shelagh R, Frank S, Sumerauer D, Scheurlen W, Ryzhova MV, Milde T, Kratz CP, Samuel D, Zhang J, Solomon DA, Marra M, Eils R, Bartram CR, von Hoff K, Rutkowski S, Ramaswamy V, Gilbertson RJ, Korshunov A, Taylor MD, Lichter P, Malkin D, Gajjar A, Korbel JO, Pfister SM. [Spectrum and prevalence of genetic predisposition in medulloblastoma: a retrospective genetic study and prospective validation in a clinical trial cohort.](#) **Lancet Oncol.** 2018 Jun;19(6):785-798. doi: 10.1016/S1470-2045(18)30242-0. Epub 2018 May 9. PMID: 29753700
202. Muzzio M, Motti JMB, Paz Sepulveda PB, Yee MC, Cooke T, Santos MR, Ramallo V, Alfaro EL, Dipierri JE, Bailliet G, Bravi CM, **Bustamante CD**, Kenny EE. [Population structure in Argentina.](#) **PLoS One.** 2018 May 1;13(5):e0196325. doi: 10.1371/journal.pone.0196325. eCollection 2018. PMID: 29715266
201. DeBoever C, Tanigawa Y, Lindholm ME, McInnes G, Lavertu A, Ingelsson E, Chang C, Ashley EA, **Bustamante CD**, Daly MJ, Rivas MA. [Medical relevance of protein-truncating variants across 337,205 individuals in the UK Biobank study.](#) **Nat Commun.** 2018 Apr 24;9(1):1612. doi: 10.1038/s41467-018-03910-9. PMID: 29691392
200. de Freitas PD, Mendez FL, Chávez-Congrains K, Galetti PM Jr, Coutinho LL, Pissinatti A, **Bustamante CD**. [Next-Generation Sequencing of the Complete Mitochondrial Genome of the Endangered Species Black Lion Tamarin \*Leontopithecus chrysopygus\* \(Primates\) and Mitogenomic Phylogeny Focusing on the Callitrichidae Family.](#) **G3** (Bethesda). 2018 May 31;8(6):1985-1991. doi: 10.1534/g3.118.200153. PMID: 29650540
199. de la Fuente C, Ávila-Arcos MC, Galimany J, Carpenter ML, Homburger JR, Blanco A, Contreras P, Cruz Dávalos D, Reyes O, San Roman M, Moreno-Estrada A, Campos PF, Eng C, Huntsman S, Burchard EG, Malaspinas AS, **Bustamante CD**, Willerslev E, Lloj E,

- Verdugo RA, Moraga M. [Genomic insights into the origin and diversification of late maritime hunter-gatherers from the Chilean Patagonia](#). *Proc Natl Acad Sci U S A*. 2018 Apr 24;115(17):E4006-E4012. doi: 10.1073/pnas.1715688115. Epub 2018 Apr 9. PMID: 29632188
198. Bhattacharya S, Li J, Sockell A, Kan MJ, Bava FA, Chen SC, Ávila-Arcos MC, Ji X, Smith E, Asadi NB, Lachman RS, Lam HYK, **Bustamante CD**, Butte AJ, Nolan GP. [Whole-genome sequencing of Atacama skeleton shows novel mutations linked with dysplasia](#). *Genome Res*. 2018 Apr;28(4):423-431. doi: 10.1101/gr.223693.117. Epub 2018 Mar 22. PMID: 29567674
197. Schroeder H, Sikora M, Gopalakrishnan S, Cassidy LM, Maisano Delsler P, Sandoval Velasco M, Schraiber JG, Rasmussen S, Homburger JR, Ávila-Arcos MC, Allentoft ME, Moreno-Mayar JV, Renaud G, Gómez-Carballa A, Laffoon JE, Hopkins RJA, Higham TFG, Carr RS, Schaffer WC, Day JS, Hoogland M, Salas A, **Bustamante CD**, Nielsen R, Bradley DG, Hofman CL, Willerslev E. [Origins and genetic legacies of the Caribbean Taino](#). *Proc Natl Acad Sci U S A*. 2018 Mar 6;115(10):2341-2346. doi: 10.1073/pnas.1716839115. Epub 2018 Feb 20. PMID: 29463742
196. Rappoport N, Toung J, Hadley D, Wong RJ, Fujioka K, Reuter J, Abbott CW, Oh S, Hu D, Eng C, Huntsman S, Bodian DL, Niederhuber JE, Hong X, Zhang G, Sikora-Wohfeld W, Gignoux CR, Wang H, Oehlert J, Jelliffe-Pawlowski LL, Gould JB, Darmstadt GL, Wang X, **Bustamante CD**, Snyder MP, Ziv E, Patsopoulos NA, Muglia LJ, Burchard E, Shaw GM, O'Brodovich HM, Stevenson DK, Butte AJ, Sirota M. [A genome-wide association study identifies only two ancestry specific variants associated with spontaneous preterm birth](#). *Sci Rep*. 2018 Jan 9;8(1):226. doi: 10.1038/s41598-017-18246-5. PMID: 29317701
195. Lins LSF, Trojahn S, Sockell A, Yee MC, Tatarenkov A, **Bustamante CD**, Earley RL, Kelley JL. [Whole-genome sequencing reveals the extent of heterozygosity in a preferentially self-fertilizing hermaphroditic vertebrate](#). *Genome*. 2018 Apr;61(4):241-247. doi: 10.1139/gen-2017-0188. Epub 2017 Dec 21. PMID: 29268023
194. Martin AR, Lin M, Granka JM, Myrick JW, Liu X, Sockell A, Atkinson EG, Werely CJ, Möller M, Sandhu MS, Kingsley DM, Hoal EG, Liu X, Daly MJ, Feldman MW, Gignoux CR, **Bustamante CD**, Henn BM. [An Unexpectedly Complex Architecture for Skin Pigmentation in Africans](#). *Cell*. 2017 Nov 30;171(6):1340-1353.e14. doi: 10.1016/j.cell.2017.11.015. PMID: 29195075
- 2017**
193. GTEx Consortium. [Genetic effects on gene expression across human tissues](#). *Nature*. 2017. 550(7675):204-213. doi: 10.1038/nature24277. PMID: 29022597
192. Cooke TF, Fischer CR, Wu P, Jiang TX, Xie KT, Kuo J, Doctorov E, Zehnder A, Khosla C, Chuong CM, **Bustamante CD**. [Genetic Mapping and Biochemical Basis of Yellow Feather Pigmentation in Budgerigars](#). *Cell*. 2017. 171(2):427-439.e21. doi:10.1016/j.cell.2017.08.016. PMID: 28985565
191. Ioannidis NM, Davis JR, DeGorter MK, Larson NB, McDonnell SK, French AJ, Battle AJ, Hastie TJ, Thibodeau SN, Montgomery SB, **Bustamante CD**, Sieh W, Whittemore AS. [FIRE: functional inference of genetic variants that regulate gene expression](#). *Bioinformatics*. 2017.



doi: 10.1093/bioinformatics/btx534. PMID: 28961785

190. Migicovsky Z, Sawler J, Gardner KM, Aradhya MK, Prins BH, Schwaninger HR, **Bustamante CD**, Buckler ES, Zhong GY, Brown PJ, Myles S. [Patterns of genomic and phenomic diversity in wine and table grapes](#). *Hortic Res.* 2017. 4:17035. doi: 10.1038/hortres.2017.35. eCollection 2017. PMID: 28791127
189. Patin E, Lopez M, Grollemund R, Verdu P, Harmant C, Quach H, Laval G, Perry GH, Barreiro LB, Froment A, Heyer E, Massougbodji A, Fortes-Lima C, Migot-Nabias F, Bellis G, Dugoujon JM, Pereira JB, Fernandes V, Pereira L, Van der Veen L, Mouguiama-Daouda P, **Bustamante CD**, Hombert JM, Quintana-Murci L. [Dispersals and genetic adaption of Bantu-speaking populations in Africa and North America](#). *Science.* 2017. 356(6337):543-546. doi: 10.1126/science.aal1988. PMID: 28473590
188. Holl HM, Brooks SA, Carpenter ML, Bustamante CD, Lafayette C. [A novel splice mutation within equine KIT and the W15 allele in the homozygous state lead to all white coat color phenotypes](#). *Anim Genet.* 2017. 48(4):497-498. doi: 10.1111/age.12554. Epub 2017 Apr 5. PMID: 28378922
187. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, **Bustamante CD**, Kenny EE. [Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations](#). *Am J Hum Genet.* 2017. 100(4):635-649. Doi: 10.1016/j.ajhg.2017.03.004. Epub 2017 Mar 30. PMID: 28366442
186. Costa HA, Neal JW, Bustamante CD, Zehnder JL. [Identification of a Novel Somatic Mutation Leading to Allele Dropout for EGFR L858R Genotyping in Non-Small Cell Lung Cancer](#). *Mol Diagn Ther.* 2017. (4):431-436. doi: 10.1007/s40291-017-0275-y. PMID: 28357677
185. McManus KF, Taravella AM, Henn BM, Bustamante CD, Sikora M, Cornejo OE. [Population genetic analysis of the DARC locus \(Duffy\) reveals adaptation from standing variation associated with malaria resistance in humans](#). *PLoS Genet.* 2017 13(3):e1006560. doi: 10.1371/journal.pgen.1006560. eCollection 2017 Mar. PMID:28282382
184. Shringarpure SS, Mathias RA, Hernandez RD, O'Connor TD, Szpiech ZA, Torres R, De La Vega FM, **Bustamante CD**, Barnes KC, Taub MA; CAAPA Consortium. 2016. [Using genotype array data to compare multi- and single-sample variant calls and improve variant call sets from deep coverage whole-genome sequencing data](#). *Bioinformatics.* 2017. 33(8):1147-1153. doi:10.1093/bioinformatics/btw786. PMID: 28035032.

## 2016

183. Bien SA, Wojcik GL, Zubair N, Gignoux CR, Martin AR, Kocarnik JM, Martin LW, Buyske S, Haessler J, Walker RW, Cheng I, Graff M, Xia L, Franceschini N, Matise T, James R, Hindorff L, Le Marchand L, North KE, Haiman CA, Peters U, Loos RJ, Kooperberg CL, **Bustamante CD**, Kenny EE, Carlson CS; PAGE Study. 2016. [Strategies for Enriching Variant Coverage in Candidate Disease Loci on a Multiethnic Genotyping Array](#). *PLoS One.* 2016. 11(12):e0167758. doi: 10.1371/journal.pone.0167758. PMID: 27973554.
182. Mathias RA, Taub MA, Gignoux CR, Fu W, Musharoff S, O'Connor TD, Vergara C, Torgerson DG, Pino-Yanes M, Shringarpure SS, Huang L, Rafaels N, Boorgula MP, Johnston HR, Ortega VE, Levin AM, Song W, Torres R, Padhukasahasram B, Eng C, Mejia-Mejia DA,

Ferguson T, Qin ZS, Scott AF, Yazdanbakhsh M, Wilson JG, Marrugo J, Lange LA, Kumar R, Avila PC, Williams LK, Watson H, Ware LB, Olopade C, Olopade O, Oliveira R, Ober C, Nicolae DL, Meyers D, Mayorga A, Knight-Madden J, Hartert T, Hansel NN, Foreman MG, Ford JG, Faruque MU, Dunston GM, Caraballo L, Burchard EG, Bleecker E, Araujo MI, Herrera-Paz EF, Gietzen K, Grus WE, Bamshad M, **Bustamante CD**, Kenny EE, Hernandez RD, Beaty TH, Ruczinski I, Akey J; CAAPA., Barnes KC. 2016. [A continuum of admixture in the Western Hemisphere revealed by the African Diaspora genome](#). **Nat Comm** 7:12522. doi: 10.1038/ncomms12522. PMID: 27725671.

181. Ioannidis NM, Rothstein JH, Pejaver V, Middha S, McDonnell SK, Baheti S, Musolf A, Li Q, Holzinger E, Karyadi D, Cannon-Albright LA, Teerlink CC, Stanford JL, Isaacs WB, Xu J, Cooney KA, Lange EM, Schleutker J, Carpten JD, Powell IJ, Cussenot O, Cancel-Tassin G, Giles GG, MaInnis RJ, Maier C, Hsieh CL, Wiklund F, Catalona WJ, Foulkes WD, Mandal D, Eeles RA, Kote-Jarai Z, **Bustamante CD**, Schaid DJ, Hastie T, Ostrander EA, Bailey-Wilson JE, Radivojac P, Thibodeau SN, Whittemore AS, Sieh W. 2016. [REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants](#). **Am J Hum Genet**. 2016 Oct 6;99(4):877-885. doi: 10.1016/j.ajhg.2016.08.016. PMID: 27666373.
180. Daneshjou R, Cavallari LH, Weeke PE, Karczewski KJ, Drozda K, Perera MA, Johnson JA, Klein TE, **Bustamante CD**, Roden DM, Shaffer C, Denny JC, Zehnder JL, Altman RB. 2016. [Population-specific single-nucleotide polymorphism confers increased risk of venous thromboembolism in African Americans](#). **Mol Genet Genomic Med**. 2016 Jun 21;4(5):513-20. doi: 10.1002/mgg3.226. PMID: 27652279.
179. Kelley JL, Yee MC, Brown AP, Richardson RR, Tatarenkov A, Lee CC, Harkins TT, **Bustamante CD**, Earley RL. 2016. [The Genome of the Self-Fertilizing Mangrove Rivulus Fish, \*Kryptolebias marmoratus\*: A Model for Studying Phenotypic Plasticity and Adaptations to Extreme Environments](#). **Genome Biol Evol**. 2016 Aug 16;8(7):2145-54. doi: 10.1093/gbe/evw145. PMID: 27324916.
178. Homburger JR, Green EM, Caleshu C, Sunitha MS, Taylor RE, Ruppel KM, Metpally RP, Colan SD, Michels M, Day SM, Olivotto I, **Bustamante CD**, Dewey FE, Ho CY, Spudich JA, Ashley EA. 2016. [Multidimensional structure-function relationships in human  \$\beta\$ -cardiac myosin from population-scale genetic variation](#). **Proc Natl Acad Sci U S A**. 2016 Jun 14;113(24):6701-6. doi: 10.1073/pnas.1606950113. PMID: 27247418.
177. Mendez FL, Poznik GD, Castellano S, **Bustamante CD**. 2016. [The Divergence of Neandertal and Modern Human Y Chromosomes](#). **Am J Hum Genet** 98(4):728-34. doi: 10.1016/j.ajhg.2016.02.023. PMID: 27058445.
176. Altman RB, Prabhu S, Sidow A, Zook JM, Goldfeder R, Litwack D, Ashley E, Asimenos G, **Bustamante CD**, Donigan K, Giacomini KM, Johansen E, Khuri N, Lee E, Liang XS, Salit M, Serang O, Tezak Z, Wall DP, Mansfield E, Kass-Hout T. 2016. [A research roadmap for next-generation sequencing informatics](#). **Sci Transl Med** 8(335):335ps10. doi: 10.1126/scitranslmed.aaf7314. Review. PMID: 27099173.
175. Poznik GD, Xue Y, Mendez FL, Willems TF, Massaia A, Wilson Sayres MA, Ayub Q, McCarthy SA, Narechania A, Kashin S, Chen Y, Banerjee R, Rodriguez-Flores JL, Cerezo M, Shao H, Gymrek M, Malhotra A, Louzada S, Desalle R, Ritchie GR, Cerveira E, Fitzgerald

- TW, Garrison E, Marcketta A, Mittelman D, Romanovitch M, Zhang C, Zheng-Bradley X, Abecasis GR, McCarroll SA, Flicek P, Underhill PA, Coin L, Zerbino DR, Yang F, Lee C, Clarke L, Auton A, Erlich Y, Handsaker RE; 1000 Genomes Project Consortium, **Bustamante CD**, Tyler-Smith C. 2016. [Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences](#). **Nat Genet** 48(6):593-9. doi: 10.1038/ng.3559. Epub 2016 Apr 25. PMID: 27111036.
174. Shringarpure SS, **Bustamante CD**, Lange K, Alexander DH. 2016. [Efficient analysis of large datasets and sex bias with ADMIXTURE](#). **BMC Bioinformatics** 17:218. doi: 10.1186/s12859-016-1082-x. PMID: 27216439.
173. Baharian S, Barakatt M, Gignoux CR, Shringarpure S, Errington J, Blot WJ, **Bustamante CD**, Kenny EE, Williams SM, Aldrich MC, Gravel S. 2016. [The Great Migration and African-American Genomic Diversity](#). **PLoS Genet** 12(5):e1006059. doi: 10.1371/journal.pgen.1006059. eCollection 2016 May. PMID: 27232753.
172. Homburger JR, Green EM, Caleshu C, Sunitha MS, Taylor RE, Ruppel KM, Metpally RP, Colan SD, Michels M, Day SM, Olivotto I, **Bustamante CD**, Dewey FE, Ho CY, Spudich JA, Ashley EA. 2016. [Multidimensional structure-function relationships in human  \$\beta\$ -cardiac myosin from population-scale genetic variation](#). **Proc Natl Acad Sci U S A** 113(24):6701-6. doi: 10.1073/pnas.1606950113. Epub 2016 May 31. PMID: 27247418.
171. Kelley JL, Arias-Rodriguez L, Martin DP, Yee MC, **Bustamante CD**, Tobler M. 2016. [Mechanisms underlying adaptation to life in hydrogen sulfide rich environments](#). **Mol Biol Evol**. 2016 Jun 19. pii: eww145. PMID: 26861137.
170. Davis JR, Fresard L, Knowles DA, Pala M, **Bustamante CD**, Battle A, Montgomery SB. 2016. [An Efficient Multiple-Testing Adjustment for eQTL Studies that Accounts for Linkage Disequilibrium between Variants](#). **Am J Hum Genet**. 98(1):216-24. PMID: 26749306.
169. Cooke TF, Yee MC, Muzzio M, Sockell A, Bell R, Cornejo OE, Kelley JL, Bailliet G, Bravi CM, **Bustamante CD**, Kenny EE. 2016. [GBStools: A Statistical Method for Estimating Allelic Dropout in Reduced Representation Sequencing Data](#). **PLoS Genet**. 12(2):e1005631. PMID: 26828719.
168. Henn BM, Botigué LR, Peischl S, Dupanloup I, Lipatov M, Maples BK, Martin AR, Musharoff S, Cann H, Snyder MP, Excoffier L, Kidd JM, **Bustamante CD**. 2016. [Distance from sub-Saharan Africa predicts mutational load in diverse human genomes](#). **Proc Natl Acad Sci U S A**. 113(4):E440-9. PMID: 26712023.
167. Freedman AH, Schweizer RM, Ortega-Del Vecchyo D, Han E, Davis BW, Gronau I, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B, Ramirez O, Hormozdiari F, Alkan C, Vilà C, Squire K, Geffen E, Kusak J, Boyko AR, Parker HG, Lee C, Tadiogtla V, Siepel A, **Bustamante CD**, Harkins TT, Nelson SF, Marques-Bonet T, Ostrander EA, Wayne RK, Novembre J. 2016. [Demographically-Based Evaluation of Genomic Regions under Selection in Domestic Dogs](#). **PLoS Genet**. 2016 Mar 4;12(3):e1005851. doi: 10.1371/journal.pgen.1005851.

## 2015

166. Stevison LS, Woerner AE, Kidd JM, Kelley JL, Veeramah KR, McManus KF; Great Ape Genome Project, **Bustamante CD**, Hammer MF, Wall JD. 2015. [The Time Scale of Recombination Rate Evolution in Great Apes](#). *Mol Biol Evol*. 2015 Dec 15. pii: msv331. PMID: 26671457.
165. Valenzano DR, Benayoun BA, Singh PP, Zhang E, Etter PD, Hu CK, Clément-Ziza M, Willemssen D, Cui R, Harel I, Machado BE, Yee MC, Sharp SC, **Bustamante CD**, Beyer A, Johnson EA, Brunet A. 2015. [The African Turquoise Killifish Genome Provides Insights into Evolution and Genetic Architecture of Lifespan](#). *Cell*. 163(6):1539-54. PMID: 26638078.
164. Homburger JR, Moreno-Estrada A, Gignoux CR, Nelson D, Sanchez E, Ortiz-Tello P, Pons-Estel BA, Acevedo-Vasquez E, Miranda P, Langefeld CD, Gravel S, Alarcón-Riquelme ME, **Bustamante CD** 2015. [Genomic Insights into the Ancestry and Demographic History of South America](#). *PLoS Genet*. 11(12):e1005602. PMID: 26636962.
163. Costa HA, Leitner MG, Sos ML, Mavrantoni A, Rychkova A, Johnson JR, Newton BW, Yee MC, De La Vega FM, Ford JM, Krogan NJ, Shokat KM, Oliver D, Halaszovich CR, **Bustamante CD**. 2015. [Discovery and functional characterization of a neomorphic PTEN mutation](#). *Proc Natl Acad Sci U S A*. 112(45):13976-81. PMID: 26504226.
162. Shringarpure SS, **Bustamante CD**. 2015. [Privacy Risks from Genomic Data-Sharing Beacons](#). *Am J Hum Genet*. 97(5):631-46. PMID: 26522470.
161. Shannon LM, Boyko RH, Castelhana M, Corey E, Hayward JJ, McLean C, White ME, Abi Said M, Anita BA, Bondjengo NI, Calero J, Galov A, Hedimbi M, Imam B, Khalap R, Lally D, Masta A, Oliveira KC, Pérez L, Randall J, Tam NM, Trujillo-Cornejo FJ, Valeriano C, Sutter NB, Todhunter RJ, **Bustamante CD**, Boyko AR. 2015. [Genetic structure in village dogs reveals a Central Asian domestication origin](#). *Proc Natl Acad Sci U S A*. 112(44):13639-44. PMID: 26483491.
160. Grubert F, Zaugg JB, Kasowski M, Ursu O, Spacek DV, Martin AR, Greenside P, Srivas R, Phanstiel DH, Pekowska A, Heidari N, Euskirchen G, Huber W, Pritchard JK, **Bustamante CD**, Steinmetz LM, Kundaje A, Snyder M. 2015. [Genetic Control of Chromatin States in Humans Involves Local and Distal Chromosomal Interactions](#). *Cell* 162(5):1051-65. PMID: 26300125.
159. Patwardhan A, Harris J, Leng N, Bartha G, Church DM, Luo S, Haudenschild C, Pratt M, Zook J, Salit M, Tirch J, Morra M, Chervitz S, Li M, Clark M, Garcia S, Chandratillake G, Kirk S, Ashley E, Snyder M, Altman R, **Bustamante C**, Butte AJ, West J, Chen R. 2015. [Achieving high-sensitivity for clinical applications using augmented exome sequencing](#). *Genome Med* 7(1):71. PMID: 26269718.
158. Raghavan M, Steinrücken M, Harris K, Schiffels S, Rasmussen S, DeGiorgio M, Albrechtsen A, Valdiosera C, Ávila-Arcos MC, Malaspinas AS, Eriksson A, Moltke I, Metspalu M, Homburger JR, Wall J, Cornejo OE, Moreno-Mayar JV, Korneliussen TS, Pierre T,

Rasmussen M, Campos PF, Damgaard PB, Allentoft ME, Lindo J, Metspalu E, Rodríguez-Varela R, Mansilla J, Henrickson C, Seguin-Orlando A, Malmström H, Stafford T Jr, Shringarpure SS, Moreno-Estrada A, Karmin M, Tambets K, Bergström A, Xue Y, Warmuth V, Friend AD, Singarayer J, Valdes P, Balloux F, LeBoreiro I, Vera JL, Rangel-Villalobos H, Pettener D, Luiselli D, Davis LG, Heyer E, Zollikofer CP, Ponce de León MS, Smith CI, Grimes V, Pike KA, Deal M, Fuller BT, Arriaza B, Standen V, Luz MF, Ricaut F, Guidon N, Osipova L, Voevoda MI, Posukh OL, Balanovsky O, Lavryashina M, Bogunov Y, Khusnutdinova E, Gubina M, Balanovska E, Fedorova S, Litvinov S, Malyarchuk B, Derenko M, Mosher MJ, Archer D, Cybulski J, Petzelt B, Mitchell J, Worl R, Norman PJ, Parham P, Kemp BM, Kivisild T, Tyler-Smith C, Sandhu MS, Crawford M, Villems R, Smith DG, Waters MR, Goebel T, Johnson JR, Malhi RS, Jakobsson M, Meltzer DJ, Manica A, Durbin R, **Bustamante CD**, Song YS, Nielsen R, Willerslev E. 2015. [Genomic evidence for the Pleistocene and recent population history of Native Americans](#). **Science**. 2015 Jul 21. pii: aab3884. PMID: 26198033.

157. Shringarpure SS, Carroll A, De La Vega FM, **Bustamante CD**. 2015. [Inexpensive and Highly Reproducible Cloud-Based Variant Calling of 2,535 Human Genomes](#). 2015. **PLoS One**. 2015 Jun 25;10(6):e0129277. doi: 10.1371/journal.pone.0129277. eCollection 2015. PMID: 26110529.
156. Rasmussen M, Sikora M, Albrechtsen A, Korneliusen TS, Moreno-Mayar JV, Poznik GD, Zollikofer CP, Ponce de León MS, Allentoft ME, Moltke I, Jónsson H, Valdiosera C, Malhi RS, Orlando L, **Bustamante CD**, Stafford TW Jr, Meltzer DJ, Nielsen R, Willerslev E. 2015. [The ancestry and affiliations of Kennewick Man](#). **Nature**. 2015 Jul 23;523(7561):455-8. doi: 10.1038/nature14625. PMID: 26087396.
155. **Bustamante CD**, Rasmussen M. 2015. [Beyond the reference genome](#). **Nat Biotechnol**. 2015 Jun;33(6):605-6. doi: 10.1038/nbt.3249. No abstract available. PMID: 26057977.
154. Rehm HL, Berg JS, Brooks LD, **Bustamante CD**, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS; ClinGen. 2015. [ClinGen--the Clinical Genome Resource](#). **N Engl J Med**. 2015 Jun 4;372(23):2235-42. doi: 10.1056/NEJMs1406261. Epub 2015 May 27. No abstract available. PMID: 26014595.
153. Tilgner H, Jahanbani F, Blauwkamp T, Moshrefi A, Jaeger E, Chen F, Harel I, **Bustamante CD**, Rasmussen M, Snyder MP. 2015. [Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events](#). **Nat Biotechnol**. 2015 Jul;33(7):736-42. doi: 10.1038/nbt.3242. Epub 2015 May 18. PMID: 25985263.
152. Henn BM, Botigué LR, **Bustamante CD**, Clark AG, Gravel S. 2015. [Estimating the mutation load in human genomes](#). **Nat Rev Genet**. 2015 Jun;16(6):333-43. doi: 10.1038/nrg3931. Epub 2015 May 12. Review. PMID: 25963372.
151. Rivas MA, Pirinen M, Conrad DF, Lek M, Tsang EK, Karczewski KJ, Maller JB, Kukurba KR, DeLuca DS, Fromer M, Ferreira PG, Smith KS, Zhang R, Zhao F, Banks E, Poplin R, Ruderfer DM, Purcell SM, Tukiainen T, Minikel EV, Stenson PD, Cooper DN, Huang KH, Sullivan TJ, Nedzel J; GTEx Consortium; Geuvadis Consortium, **Bustamante CD**, Li JB, Daly MJ, Guigo R, Donnelly P, Ardlie K, Sammeth M, Dermitzakis ET, McCarthy MI,

Montgomery SB, Lappalainen T, MacArthur DG. 2015. [Human genomics. Effect of predicted protein-truncating genetic variants on the human transcriptome.](#) *Science*. 2015 May 8;348(6235):666-9. PMID: 25954003.

150. Schroeder H, Ávila-Arcos MC, Malaspinas AS, Poznik GD, Sandoval-Velasco M, Carpenter ML, Moreno-Mayar JV, Sikora M, Johnson PL, Allentoft ME, Samaniego JA, Haviser JB, Dee MW, Stafford TW Jr, Salas A, Orlando L, Willerslev E, **Bustamante CD**, Gilbert MT. [Genome-wide ancestry of 17th-century enslaved Africans from the Caribbean.](#) *Proc Natl Acad Sci U S A*. 2015 Mar 24;112(12):3669-73. doi: 10.1073/pnas.1421784112. Epub 2015 Mar 9. PMID: 25755263.
149. McManus KF, Kelley JL, Song S, Veeramah KR, Woerner AE, Stevison LS, Ryder OA, Ape Genome Project G, Kidd JM, Wall JD, **Bustamante CD**, Hammer MF. [Inference of gorilla demographic and selective history from whole-genome sequence data.](#) *Mol Biol Evol*. 2015 Mar;32(3):600-12. doi: 10.1093/molbev/msu394. Epub 2014 Dec 21. PMID: 25534031.

## 2014

148. Pino-Yanes M, Gignoux CR, Galanter JM, Levin AM, Campbell CD, Eng C, Huntsman S, Nishimura KK, Gourraud PA, Mohajeri K, O'Roak BJ, Hu D, Mathias RA, Nguyen EA, Roth LA, Padhukasahasram B, Moreno-Estrada A, Sandoval K, Winkler CA, Lurmann F, Davis A, Farber HJ, Meade K, Avila PC, Serebrisky D, Chapela R, Ford JG, Lenoir MA, Thyne SM, Brigino-Buenaventura E, Borrell LN, Rodriguez-Cintrón W, Sen S, Kumar R, Rodriguez-Santana JR, **Bustamante CD**, Martinez FD, Raby BA, Weiss ST, Nicolae DL, Ober C, Meyers DA, Bleecker ER, Mack SJ, Hernandez RD, Eichler EE, Barnes KC, Williams LK, Torgerson DG, Burchard EG. [Genome-wide association study and admixture mapping reveal new loci associated with total IgE levels in Latinos.](#) 2015. *J Allergy Clin Immunol pii: S0091-6749(14)01575-9*. PMID: 25488688.
147. Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliussen TS, Grønnow B, Appelt M, Gulløv HC, Friesen TM, Fitzhugh W, Malmström H, Rasmussen S, Olsen J, Melchior L, Fuller BT, Fahrni SM, Stafford T Jr, Grimes V, Renouf MA, Cybulski J, Lynnerup N, Lahr MM, Britton K, Knecht R, Arneborg J, Metspalu M, Cornejo OE, Malaspinas AS, Wang Y, Rasmussen M, Raghavan V, Hansen TV, Khusnutdinova E, Pierre T, Dneprovsky K, Andreasen C, Lange H, Hayes MG, Coltrain J, Spitsyn VA, Götherström A, Orlando L, Kivisild T, Vilems R, Crawford MH, Nielsen FC, Dissing J, Heinemeier J, Meldgaard M, **Bustamante CD**, O'Rourke DH, Jakobsson M, Gilbert MT, Nielsen R, Willerslev E. [The genetic prehistory of the New World Arctic.](#) 2014. *Science* 345(6200):1255832. PMID: 25170159.
146. Daneshjou R, Gamazon ER, Burkley B, Cavallari LH, Johnson JA, Klein TE, Limdi N, Hillenmeyer S, Percha B, Karczewski KJ, Langae T, Patel SR, **Bustamante CD**, Altman RB, Perera MA. [Genetic variant in folate homeostasis is associated with lower warfarin dose in African Americans.](#) *Blood* 124(14):2298-305. PMID: 25079360.
145. Pino-Yanes M, Thakur N, Gignoux CR, Galanter JM, Roth LA, Eng C, Nishimura KK, Oh SS, Vora H, Huntsman S, Nguyen EA, Hu D, Drake KA, Conti DV, Moreno-Estrada A, Sandoval K, Winkler CA, Borrell LN, Lurmann F, Islam TS, Davis A, Farber HJ, Meade K, Avila PC, Serebrisky D, Bibbins-Domingo K, Lenoir MA, Ford JG, Brigino-Buenaventura E, Rodriguez-Cintrón W, Thyne SM, Sen S, Rodriguez-Santana JR, **Bustamante CD**, Williams LK, Gilliland FD, Gauderman WJ, Kumar R, Torgerson DG, Burchard EG. [Genetic ancestry influences asthma susceptibility and lung function among Latinos.](#) 2014. *J Allergy Clin*

**Immunol.** S0091-6749(14)01107-5. PMID: 25301036.

144. Kelley JL, Peyton JT, Fiston-Lavier AS, Teets NM, Yee MC, Johnston JS, **Bustamante CD**, Lee RE, Denlinger DL. [Compact genome of the Antarctic midge is likely an adaptation to an extreme environment.](#) 2014. **Nat Commun** 5:4611. PMID: 25118180.
143. Martin AR, Costa HA, Lappalainen T, Henn BM, Kidd JM, Yee MC, Grubert F, Cann HM, Snyder M, Montgomery SB, **Bustamante CD**. [Transcriptome sequencing from diverse human populations reveals differentiated regulatory architecture.](#) 2014. **PLoS Genet** 10(8):e1004549. PMID: 25121757.
142. Gardner KM, Brown PJ, Cooke TF, Cann S, Costa F, **Bustamante CD**, Velasco R, Troggio M, Myles S. [Fast and Cost-Effective Genetic Mapping in Apple Using Next-Generation Sequencing.](#) **G3**. 2014 Jul 16. pii: g3.114.011023. doi: 10.1534/g3.114.011023. PMID: 25031181.
141. Moreno-Estrada A, Gignoux CR, Fernández-López JC, Zakharia F, Sikora M, Contreras AV, Acuña-Alonzo V, Sandoval K, Eng C, Romero-Hidalgo S, Ortiz-Tello P, Robles V, Kenny EE, Nuño-Arana I, Barquera-Lozano R, Macín-Pérez G, Granados-Arriola J, Huntsman S, Galanter JM, Via M, Ford JG, Chapela R, Rodríguez-Cintrón W, Rodríguez-Santana JR, Romieu I, Sienna-Monge JJ, del Rio Navarro B, London SJ, Ruiz-Linares A, Garcia-Herrera R, Estrada K, Hidalgo-Miranda A, Jimenez-Sanchez G, Carnevale A, Soberón X, Canizales-Quinteros S, Rangel-Villalobos H, Silva-Zolezzi I, Burchard EG, **Bustamante CD**. [The genetics of Mexico recapitulates Native American substructure and affects biomedical traits.](#) **Science**. 2014. 344:1280-5. doi: 10.1126/science.1251688. PMID: 24926019.
140. Sikora M, Carpenter ML, Moreno-Estrada A, Henn BM, Underhill PA, Sánchez-Quinto F, Zara I, Pitzalis M, Sidore C, Busonero F, Maschio A, Angius A, Jones C, Mendoza-Revilla J, Nekhrizov G, Dimitrova D, Theodossiev N, Harkins TT, Keller A, Maixner F, Zink A, Abecasis G, Sanna S, Cucca F, **Bustamante CD**. [Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe.](#) 2014. **PLoS Genet** 10(5):e1004353. PMID: 24809476.
139. Kidd JM, Sharpton TJ, Bobo D, Norman PJ, Martin AR, Carpenter ML, Sikora M, Gignoux CR, Nemat-Gorgani N, Adams A, Guadalupe M, Guo X, Feng Q, Li Y, Liu X, Parham P, Hoal EG, Feldman MW, Pollard KS, Wall JD, **Bustamante CD**, Henn BM. [Exome capture from saliva produces high quality genomic and metagenomic data.](#) 2013. **BMC Genomics** 15(1):262. PMID: 2470809.
138. Underhill PA, Poznik GD, Rootsi S, Järve M, Lin AA, Wang J, Passarelli B, Kanbar J, Myres NM, King RJ, Di Cristofaro J, Sahakyan H, Behar DM, Kushniarevich A, Sarac J, Saric T, Rudan P, Pathak AK, Chaubey G, Grugni V, Semino O, Yepiskoposyan L, Bahmanimehr A, Farjadian S, Balanovsky O, Khusnutdinova EK, Herrera RJ, Chiaroni J, **Bustamante CD**, Quake SR, Kivisild T, Vilems R. [The phylogenetic and geographic structure of Y-chromosome haplogroup R1a.](#) 2014. **Eur J Hum Genet.** (in press). doi: 10.1038/ejhg.2014.50. PMID: 24667786.
137. Rasmussen M, Anzick SL, Waters MR, Skoglund P, DeGiorgio M, Stafford TW Jr, Rasmussen S, Moltke I, Albrechtsen A, Doyle SM, Poznik GD, ..., **Bustamante CD**, Jakobsson M, Nielsen R, Willerslev E. [The genome of a Late Pleistocene human from a Clovis burial site in western Montana.](#) **Nature**. 2014 Feb 13;506(7487):225-9. doi: 10.1038/nature13025. PMID: 24522598.
136. Drake KA, Torgerson DG, Gignoux CR, Galanter JM, Roth LA, Huntsman S, Eng C, Oh SS, Yee SW, Lin L, **Bustamante CD**, Moreno-Estrada A, Sandoval K, Davis A, Borrell LN, Farber HJ, Kumar R, Avila PC, Brigino-Buenaventura E, Chapela R, Ford JG, Lenoir MA, Lurmann F, Meade K, Serebrisky

D, Thyne S, Rodríguez-Cintrón W, Sen S, Rodríguez-Santana JR, Hernandez RD, Giacomini KM, Burchard EG. [A genome-wide association study of bronchodilator response in Latinos implicates rare variants](#). 2014. *J Allergy Clin Immunol*. 133(2):370-378.e15. doi: 10.1016/j.jaci.2013.06.043. PMID: 23992748.

135. Martin AR, Tse G, **Bustamante CD**, Kenny EE. [Imputation-based assessment of next generation rare exome variant arrays](#). *Pac Symp Biocomput*. 2014:241-52. PMID: 24297551.
134. Daneshjou R, Zappala Z, Kukurba K, Boyle SM, Ormond KE, Klein TE, Snyder M, **Bustamante CD**, Altman RB, Montgomery SB. [Path-scan: a reporting tool for identifying clinically actionable variants](#). 2014. *Pac Symp Biocomput* 2014:229-40. PMID: 24297550.
133. Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, Silva PM, Galaverni M, Fan Z, Marx P, Lorente-Galdos B, Beale H, Ramirez O, Hormozdiari F, Alkan C, Vilà C, Squire K, Geffen E, Kusak J, Boyko AR, Parker HG, Lee C, Tadiogola V, Siepel A, **Bustamante CD**, Harkins TT, Nelson SF, Ostrander EA, Marques-Bonet T, Wayne RK, Novembre J. [Genome sequencing highlights the dynamic early history of dogs](#). 2014. *PLoS Genet*. 10(1):e1004016. doi: 10.1371/journal.pgen.1004016.

## 2013

132. Gravel S, Zakharia F, Moreno-Estrada A, Byrnes JK, Muzzio M, Rodriguez-Flores JL, Kenny EE, Gignoux CR, Maples BK, Guiblet W, Dutil J, Via M, Sandoval K, Bedoya G; 1000 Genomes Project, Oleksyk TK, Ruiz-Linares A, Burchard EG, Martinez-Cruzado JC, **Bustamante CD**. 2013. [Reconstructing native american migrations from whole-genome and whole-exome data](#). *PLoS Genet*. 9(12):e1004023. doi: 10.1371/journal.pgen.1004023.
131. Auton A, Rui Li Y, Kidd J, Oliveira K, Nadel J, Holloway JK, Hayward JJ, Cohen PE, Grealley JM, Wang J, **Bustamante CD**, Boyko AR. [Genetic recombination is targeted towards gene promoter regions in dogs](#). 2013. *PLoS Genet*. 9(12):e1003984. doi: 10.1371/journal.pgen.1003984. PMID: 24348265.
130. Carpenter ML, Buenrostro JD, Valdiosera C, Schroeder H, Allentoft ME, Sikora M, Rasmussen M, Gravel S, Guillén S, Nekhrizov G, Leshtakov K, Dimitrova D, Theodossiev N, Pettener D, Luiselli D, Sandoval K, Moreno-Estrada A, Li Y, Wang J, Gilbert MT, Willerslev E, Greenleaf WJ, **Bustamante CD**. [Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries](#). 2013. *Am J Hum Genet*. 7;93(5):852-64. doi: 10.1016/j.ajhg.2013.10.002. PMID: 24568772.
129. Moreno-Estrada A, Gravel S, Zakharia F, McCauley JL, Byrnes JK, Gignoux CR, Ortiz-Tello PA, Martínez RJ, Hedges DJ, Morris RW, Eng C, Sandoval K, Acevedo-Acevedo S, Norman PJ, Layrissa Z, Parham P, Martínez-Cruzado JC, Burchard EG, Cuccaro ML, Martin ER, **Bustamante CD**. [Reconstructing the population genetic history of the Caribbean](#). 2013. *PLoS Genet*. 9(11):e1003925. doi: 10.1371/journal.pgen.1003925. Epub 2013 Nov 14. PMID: 24244192.
128. Goldstein O, Mezey JG, Schweitzer PA, Boyko AR, Gao C, **Bustamante CD**, Jordan JA, Aguirre GD, Acland GM. [QCB1 and PDE6B mutations cause similar early onset retinal degenerations in two closely related terrier dog breeds](#). 2013. *Invest Ophthalmol Vis Sci*. 54(10):7005-19. doi: 10.1167/iovs.13-12915. PMID: 24045995.
127. Kumar R, Nguyen EA, Roth LA, Oh SS, Gignoux CR, Huntsman S, Eng C, Moreno-Estrada



- A, Sandoval K, Peñaloza-Espinosa RI, López-López M, Avila PC, Farber HJ, Tcheurekdjian H, Rodriguez-Cintron W, Rodriguez-Santana JR, Serebrisky D, Thyne SM, Williams LK, Winkler C, **Bustamante CD**, Pérez-Stable EJ, Borrell LN, Burchard EG. *Factors associated with degree of atopy in Latino children in a nationwide pediatric sample: the Genes-environments and Admixture in Latino Asthmatics (GALA II) study.* 2013. **J Allergy Clin Immunol** 132(4):896-905.e1. doi: 10.1016/j.jaci.2013.02.046. PMID: 23684070.
126. Roots S, Behar DM, Järve M, Lin AA, Myres NM, Passarelli B, Poznik GD, Tzur S, Sahakyan H, Pathak AK, Rosset S, Metspalu M, Grugni V, Semino O, Metspalu E, **Bustamante CD**, Skorecki K, Villems R, Kivisild T, Underhill PA. *Phylogenetic applications of whole Y-chromosome sequences and the Near Eastern origin of Ashkenazi Levites.* 2013. **Nat Commun** 4:2928. doi: 10.1038/ncomms3928. PMID: 24346185.
125. Ma X, Kelley JL, Eilertson K, Musharoff S, Degenhardt JD, Martins AL, Vinar T, Kosiol C, Siepel A, Gutenkunst RN, **Bustamante CD**. 2013. *Population genomic analysis reveals a rich speciation and demographic history of orang-utans (Pongo pygmaeus and Pongo abelii).* **PLoS One** 23;8(10):e77175. doi: 10.1371/journal.pone.0077175. eCollection 2013. PMID: 24194868
124. Maples BK, Gravel S, Kenny EE, Bustamante CD. 2013. *RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference.* **Am J Hum Genet.** 93(2):278-88 PMID: 23910464
123. Poznik GD, Henn BM, Yee M-C, Sliwerska E, Euskirchen GM, Lin AA, Snyder MP, Quintana-Murci L, Kidd JM, Underhill PA, Bustamante CD. 2013. *Sequencing Y chromosomes resolves discrepancy in time to common ancestor of males versus females.* **Science** 341:562-5. PMID: 23908239
122. Prado-Martinez J, Sudmant PH, Kidd JM,..., **Bustamante CD**, Hammer MF, Eichler EE, Marques-Bonet T. 2013. *Great ape genetic diversity and population history.* **Nature** 499:471—5.
121. Botigué LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, Atzmon G, Burns E, Ostrer H, Flores C, Bertranpetit J, Comas D, and **Bustamante CD**. 2013. *Gene flow from North Africa contributes to differential human genetic diversity in southern Europe.* **PNAS** 110(29):11791—6. PMID: 23733930.
120. Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D 3rd, Cornejo O, Findley SD, Zheng P, Utró F, Royaert S, Saski C, Jenkins J, Podicheti R, Zhao M, Scheffler BE, Stack JC, Feltus FA, Mustiga GM, Amores F, Phillips W, Marelli JP, May GD, Shapiro H, Ma J, **Bustamante CD**, Schnell RJ, Main D, Gilbert D, Parida L, Kuhn DN. 2013. *The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color.* **Genome Biol** 3;14(6):R53. PMID: 23731509.
119. Corona E, Chen R, Sikora M, Morgan AA, Patel CJ, Ramesh A, **Bustamante CD**, Butte AJ. 2013. *Analysis of the genetic basis of disease in the context of worldwide human relationships and migration.* **PLoS Genet.** 9(5):e1003447. PMID: 23717210; PMCID: PMC3662561.
118. Kumar R, Nguyen EA, Roth LA, Oh SS, Gignoux CR, Huntsman S, Eng C, Moreno-Estrada A, Sandoval K, Peñaloza-Espinosa RI, López-López M, Avila PC, Farber HJ, Tcheurekdjian H, Rodriguez-Cintron W, Rodriguez-Santana JR, Serebrisky D, Thyne SM, Williams LK, Winkler C, **Bustamante CD**, Pérez-Stable EJ, Borrell LN, Burchard EG. 2013. *Factors associated with degree of atopy in Latino children in a nationwide pediatric sample: the Genes-environments and Admixture in Latino Asthmatics (GALA II) study.* **J Allergy Clin**

**Immuno** S0091-6749(13)00586-1. PMID: 23684070.

117. Cornejo OE, Lefébure T, Bitar PD, Lang P, Richards VP, Eilertson K, Do T, Beighton D, Zeng L, Ahn SJ, Burne RA, Siepel A, **Bustamante CD**, Stanhope MJ. 2013. [Evolutionary and population genomics of the cavity causing bacteria \*Streptococcus mutans\*](#). **Mol Biol Evol** 30(4):881-93. PMID: 23228887.

## 2012

116. Casto AM, Henn BM, Kidd JM, **Bustamante CD**, Feldman MW. 2012. [A tale of two haplotypes: the EDA2R/AR Intergenic region is the most divergent genomic segment between Africans and East Asians in the human genome](#). **Hum Biol.** 84(6):641-94. PMID: 23959643.
115. Brisbin A, Bryc K, Byrnes J, Zakharia F, Omberg L, Degenhardt J, Reynolds A, Ostrer H, Mezey JG, and **Bustamante CD**. 2012. [PCAdmix: principal components-based assignment of ancestry along each chromosome in individuals with admixed ancestry from two or more populations](#). **Hum Biol.** 84(4):343-64. PMID: 23249312.
114. Eilertson KE, Booth JG, and **Bustamante CD**. 2012. *SnIPRE: Selection Inference Using a Poisson Random Effects Model*. **PLoS Comput Biol** 8(12): e1002806. doi:10.1371/journal.pcbi.1002806.
113. 1000 Genomes Project Consortium. *An integrated map of genetic variation from 1,092 human genomes*. 2012. **Nature** 491:56-65. PMID: 23128226.
112. Pickrell JK, Patterson N, Barbieri C, Berthold F, Gerlach L, Güldemann T, Kure B, Mpoloka SW, Nakagawa H, Naumann C, Lipson M, Loh PR, Lachance J, Mountain J, **Bustamante CD**, Berger B, Tishkoff SA, Henn BM, Stoneking M, Reich D, and Pakendorf B. 2012. *The genetic prehistory of southern Africa*. **Nat Commun.** 3:1143. PMID: 23072811.
111. Sánchez-Quinto F, Botigué LR, Civit S, Arenas C, Avila-Arcos MC, **Bustamante CD**, Comas D, and Lalueza-Fox C. 2012. *North African populations carry the signature of admixture with Neandertals*. **PLoS One** 7(10):e47765. PMID: 23082212; PMCID: PMC3474783.
110. Kidd JM, Gravel G, Byrnes J, Moreno-Estrada A, Musharoff S, Bryc K, Degenhardt JD, Brisbin A, Sheth V, Chen R, McLaughlin SF, Peckham HE, Omberg L, Bormann Chung CA, Stanley S, Pearlstein K, Levandowsky E, Acevedo-Acevedo S, Auton A, Keinan A, Acuña-Alonzo V, Barquera-Lozano R, Canizales-Quinteros S, Eng S, Burchard EG, Russell A, Reynolds A, Clark AG, Reese MG, Lincoln SE, Butte AJ, De La Vega FM, and **Bustamante CD**. 2012. *Population Genetic Inference from Personal Genome Data: Impact of Ancestry and Admixture on Human Genomic Variation*. **Am J Hum Genet** 91(4):660-71. PMID: 23040495.
109. Granka JM, Henn BM, Gignoux CR, Kidd JM, **Bustamante CD**, and Feldman MW. 2012. *Limited Evidence for Classic Selective Sweeps in African Populations*. **Genetics**. PMID: 22960214 (cover article).
108. Kenny EE, Timpson, NJ, Sikora M., Yee, M-C, Moreno-Estrada A, Eng C, Huntsman S, Burchard EG, Stoneking M, **Bustamante CD\***, and Myles S.\* 2012. *Melanesian blond hair is caused by an amino acid change in TYRP1*. **Science** 336(6081):554. PMID: 22556244.
107. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, McGee S, Do R, Liu X, Jun G, Kang HM, Jordan D, Leal SM, Gabriel S, Rieder MJ, Abecasis G, Altshuler D, Nickerson DA, Boerwinkle E, Sunyaev S, **Bustamante CD**, Bamshad MJ, Akey JM; Broad

GO; Seattle GO. 2012 *Evolution and Functional Impact of Rare Coding Variation from Deep Sequencing of Human Exomes*. **Science** 337(6090):64-9. PMID: 22604720.

106. Campbell CL, Palamara PF, Dubrovsky M, Botigué LR, Fellous M, Atzmon G, Oddoux C, Pearlman A, Hao L, Henn BM, Burns E, **Bustamante CD**, Comas D, Friedman E, Pe'er I, and Ostrer H. 2012. *North African Jewish and non-Jewish populations form distinctive, orthogonal clusters*. **Proc Natl Acad Sci U S A**. 2012 Aug 21;109(34):13865-70. PMCID: PMC3427049.
105. Clark RT, Famoso AN, Zhao K, Shaff JE, Craft EJ, **Bustamante CD**, McCouch SR, Aneshansley DJ, and Kochian LV. 2012. *High-throughput two-dimensional root system phenotyping platform facilitates genetic analysis of root growth and development*. **Plant Cell Environ**. (in press) PMID: 22860896.
104. Ma X, Rogacheva MV, Nishant KT, Zanders S, **Bustamante CD**, and Alani E. 2012. *Mutation hot spots in yeast caused by long-range clustering of homopolymeric sequences*. **Cell Rep**. 1(1):36—42.
103. Kelley JL, Yee MC, Lee C, Levandowsky E, Shah M, Harkins T, Earley RL, and **Bustamante CD**. 2012. *The Possibility of De Novo Assembly of the Genome and Population Genomics of the Mangrove Rivulus *Kryptolebias marmoratus**. **Integr Comp Biol** 52(6):737-42 PMID: 22723055
102. Schoenebeck JJ, Hutchinson SA, Byers A, Beale HC, Carrington B, Faden DL, Rimbault M, Decker B, Kidd JM, Sood R, Boyko AR, Fondon JW 3rd, Wayne RK, **Bustamante CD**, Ciruna B, and Ostrander EA. 2012. *Variation of BMP3 Contributes to Dog Breed Skull Diversity*. **PLoS Genet**. 8(8):e1002849. PMCID: PMC3410846.
101. Lam HY, Pan C, Clark MJ, Lacroute P, Chen R, Haraksingh R, O'Huallachain M, Gerstein MB, Kidd JM, **Bustamante CD**, and Snyder M. 2012. *Detecting and annotating genetic variations using the HigeSeq pipeline*. **Nat Biotechnol** 30(3):226—9. PMID: 22398614.
100. Henn, B. M., Botigué, L. R., Gravel, S., Wang, W., Brisbin, A., Byrnes, J. K., Fadhlouzi-Zid, K., Zalloua, P. A., Moreno-Estrada, A., Bertranpetit, J., Bustamante, C. D., and Comas, D. 2012. *Genomic ancestry of North Africans supports back-to-Africa migrations*. **PLoS Genet** 8(1): e1002397. PMCID: PMC3257290.
99. Chen, R., Corona, E., Sikora, M., Dudley, J. T., Morgan, A. A., Moreno-Estrada, A., Nilsen, G. B., Ruau, D., Lincoln, S. E., **Bustamante, C. D.**, and Butte, A. J. 2012. *Type 2 diabetes risk alleles demonstrate extreme directional differentiation among human populations, compared to other diseases*. **PLoS Genet** 8(4):e1002621. PMCID: PMC3325177.
98. Keller A, Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Rootsi S, Myres NM, Cabrera VM, Underhill PA, **Bustamante CD**, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus H, O'Connor BD, Carlson MR, Meder B, Blin N, Meese E, Pusch CM, Zink A. 2012. *New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing*. **Nat Commun**. 3:698.

## 2011

97. Kenny EE, **Bustamante CD**. 2011. *SnapShot: Human biomedical genomics*. **Cell** 147(1):248- e1.

96. **Bustamante CD**, Burchard EG, De la Vega FM. 2011. *Genomics for the world*. **Nature** 475:163-5.
95. Dewey FE, Chen R, Cordero SP, Ormond KE, Caleshu C, Karczewski KJ, Whirl-Carrillo M, Wheeler MT, Dudley JT, Byrnes JK, Cornejo OE, Knowles JW, Woon M, Sangkuhl K, Gong L, Thorn CF, Hebert JM, Capriotti E, David SP, Pavlovic A, West A, Thakuria JV, Ball MP, Zaranek AW, Rehm HL, Church GM, West JS, **Bustamante CD**, Snyder M, Altman RB, Klein TE, Butte AJ, Ashley EA. 2011. *Phased whole-genome genetic risk in a family quartet using a major allele reference sequence*. **PLoS Genet** 7(9):e1002280. PMID: 3174201.
94. Gao H, Bryc K, **Bustamante CD**. *On identifying the optimal number of population clusters via the deviance information criterion*. **PLoS One**. 2011;6(6):e21014. PMID: 3125185.
93. Marth GT, Yu F, Indap AR, Garimella K, Gravel S, Leong WF, Tyler-Smith C, Bainbridge M, Blackwell T, Zheng-Bradley X, Chen Y, Challis D, Clarke L, Ball EV, Cibulskis K, Cooper DN, Fulton B, Hartl C, Koboldt D, Muzny D, Smith R, Sougnez C, Stewart C, Ward A, Yu J, Xue Y, Altshuler D, **Bustamante CD**, Clark AG, Daly M, Depristo M, Flicek P, Gabriel S, Mardis E, Palotie A, Gibbs RA, 1000 Genomes Project. *The functional spectrum of low-frequency coding variation*. **Genome Biol**. 2011;12(9):R84.
92. Rasmussen M, Guo X, Wang Y, Lohmueller KE, ..., **Bustamante CD**, ..., Nielsen R, Jun W, Willerslev E. *An Aboriginal Australian Genome Reveals Separate Human Dispersals into Asia*. **Science**. 2011.
91. Xie X, Molina J, Hernandez R, Reynolds A, Boyko AR, **Bustamante CD**, Purugganan MD. *Levels and patterns of nucleotide variation in domestication QTL regions on rice chromosome 3 suggest lineage-specific selection*. **PLoS One**. 2011;6(6):e20670. PMID: 3108957.
90. Zhao K, Tung CW, Eizenga GC, Wright MH, Ali ML, Price AH, Norton GJ, Islam MR, Reynolds A, Mezey J, McClung AM, **Bustamante CD\***, McCouch SR\*. *Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa**. **Nat Commun**. 2011;2:467.
89. Vonholdt BM, Pollinger JP, Earl DA, Knowles JC, Boyko AR, Parker H, Geffen E, Pilot M, Jedrzejewski W, Jedrzejewska B, Sidorovich V, Greco C, Randi E, Musiani M, Kays R, **Bustamante CD**, Ostrander EA, Novembre J, Wayne RK. *A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids*. **Genome Res**. 2011.
88. Brisbin A, Cruickshank J, Moise NS, Gunn T, **Bustamante CD**, Mezey JG. *Fast, exact linkage analysis for categorical traits on arbitrary pedigree designs*. **Genet Epidemiol**. 2011.
87. Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, Huang P, Jackson S, Schaal BA, **Bustamante CD**, Boyko AR, Purugganan MD. *Molecular evidence for a single evolutionary origin of domesticated rice*. **Proc Natl Acad Sci U S A**. 2011.108(20):8351-6.
86. Myers RA, Casals F, Gauthier J, Hamdan FF, Keebler J, Boyko AR, **Bustamante CD**, Piton AM, Spiegelman D, Henrion E, Zilversmit M, Hussin J, Quinlan J, Yang Y, Lafreniere RG, Griffing AR, Stone EA, Rouleau GA, Awadalla P. *A population genetic approach to mapping neurological disorder genes using deep resequencing*. **PLoS Genet**. 2011;7(2):e1001318. PMID: 3044677.
85. Gravel S, Henn BM, Gutenkunst RN, Indap AR, Marth GT, Clark AG, Yu F, Gibbs RA, **Bustamante CD**. *Demographic history and rare allele sharing among human populations*. **Proc Natl Acad Sci U S A**. 2011;108(29):11983-8. PMID: 3142009.

84. Henn BM, Gignoux CR, Jobin M, Granka JM, Macpherson JM, Kidd JM, Rodriguez-Botigue L, Ramachandran S, Hon L, Brisbin A, Lin AA, Underhill PA, Comas D, Kidd KK, Norman PJ, Parham P, **Bustamante CD**, Mountain JL, Feldman MW. *Feature Article: Hunter-gatherer genomic diversity suggests a southern African origin for modern humans.* **Proc Natl Acad Sci U S A.** 2011. 108(13):5154-62. PMID: 3069156.
83. Locke DP, Hillier LW, Warren WC, ... , **Bustamante CD**, ... , Gibbs RA, Wilson RK. *Comparative and demographic analysis of orang-utan genomes.* **Nature.** 2011;469(7331):529-33. PMID: 3060778.
82. Myles S, Boyko AR, Owens CL, Brown PJ, Grassi F, Aradhya MK, Prins B, Reynolds A, Chia JM, Ware D, **Bustamante CD\***, Buckler ES\*. *From the Cover: Genetic structure and domestication history of the grape.* **Proc Natl Acad Sci U S A.** 2011;108(9):3530-5. PMID: 3048109.
81. Lohmueller KE, **Bustamante CD**, Clark AG. *Detecting directional selection in the presence of recent admixture in african-americans.* **Genetics.** 2011;187(3):823-35.

## 2010

79. **Bustamante CD**, Henn BM. *Human origins: Shadows of early migrations.* 2010. **Nature** 468:1044-5.
78. FM DLV, **Bustamante CD**, Leal SM. *Genome-Wide Association Mapping And Rare Alleles: From Population Genomics To Personalized Medicine - Session Introduction.* **Pac Symp Biocomput.** 2011:74-5.
77. Andres AM, Dennis MY, Kretschmar WW, Cannons JL, Lee-Lin SQ, Hurle B, Schwartzberg PL, Williamson SH, **Bustamante CD**, Nielsen R, Clark AG, Green ED. *Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation.* **PLoS Genet.** 2010;6(10):e1001157. PMID: 2954825.
76. Wright MH, Tung CW, Zhao K, Reynolds A, McCouch SR, **Bustamante CD.** *ALCHEMY: a reliable method for automated SNP genotype calling for small batch sizes and highly homozygous populations.* **Bioinformatics.** 2010;26(23):2952-60. PMID: 2982150.
75. Henn BM, Gravel S, Moreno-Estrada A, Acevedo-Acevedo S, **Bustamante CD.** *Fine-scale population structure and the era of next-generation sequencing.* **Hum Mol Genet.** 2010;19(R2):R221-6. PMID: 2953744.
74. Brisbin A, Weissman MM, Fyer AJ, Hamilton SP, Knowles JA, **Bustamante CD**, Mezey JG. *Bayesian linkage analysis of categorical traits for arbitrary pedigree designs.* **PLoS One.** 2010;5(8):e12307. PMID: 2928726.
73. Nishant KT, Wei W, Mancera E, Argueso JL, Schlattl A, Delhomme N, Ma X, **Bustamante CD**, Korbel JO, Gu Z, Steinmetz LM, Alani E. *The baker's yeast diploid genome is remarkably stable in vegetative growth and meiosis.* **PLoS Genet.** 2010;6(9). PMID: 2936533.
72. Goldstein O, Mezey JG, Boyko AR, Gao C, Wang W, **Bustamante CD**, Anguish LJ, Jordan JA, Pearce-Kelling SE, Aguirre GD, Acland GM. *An ADAM9 mutation in canine cone-rod dystrophy 3 establishes homology with human cone-rod dystrophy 9.* **Mol Vis.** 2010;16:1549-69. PMID: 2925905.
71. Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, Lohmueller KE, Zhao K, Brisbin A, Parker HG, vonHoldt BM, Cargill M, Auton A, Reynolds A, Elkhouloun AG,

- Castelhano M, Mosher DS, Sutter NB, Johnson GS, Novembre J, Hubisz MJ, Siepel A, Wayne RK, **Bustamante CD**, Ostrander EA. *A simple genetic architecture underlies morphological variation in dogs*. **PLoS Biol.** 2010;8(8):e1000451. PMID: 2919785.
70. Zanders S, Ma X, Roychoudhury A, Hernandez RD, Demogines A, Barker B, Gu Z, **Bustamante CD**, Alani E. *Detection of heterozygous mutations in the genome of mismatch repair defective diploid yeast using a Bayesian approach*. **Genetics.** 2010;186(2):493-503. PMID: 2954485.
69. Zhao K, Wright M, Kimball J, Eizenga G, McClung A, Kovach M, Tyagi W, Ali ML, Tung CW, Reynolds A, **Bustamante CD**, McCouch SR. *Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome*. **PLoS One.** 2010;5(5):e10780. PMID: 2875394.
68. Bryc K, Velez C, Karafet T, Moreno-Estrada A, Reynolds A, Auton A, Hammer M, **Bustamante CD**, Ostrer H. *Colloquium paper: genome-wide patterns of population structure and admixture among Hispanic/Latino populations*. **Proc Natl Acad Sci U S A.** 2010;107 Suppl 2:8954-61. PMID: 3024022.
67. Williams LM, Ma X, Boyko AR, **Bustamante CD**, Oleksiak MF. *SNP identification, verification, and utility for population genetics in a non-model genus*. **BMC Genet.** 2010;11:32. PMID: 2874759.
66. Brideau CM, Eilertson KE, Hagarman JA, **Bustamante CD**, Soloway PD. *Successful computational prediction of novel imprinted genes from epigenomic features*. **Mol Cell Biol.** 2010;30(13):3357-70. PMID: 2897571.
65. Lohmueller KE, **Bustamante CD**, Clark AG. *The effect of recent admixture on inference of ancient human population history*. **Genetics.** 2010;185(2):611-22. PMID: 2881141.
64. Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, Reynolds A, Bryc K, Brisbin A, Knowles JC, Mosher DS, Spady TC, Elkahouloun A, Geffen E, Pilot M, Jedrzejewski W, Greco C, Randi E, Bannasch D, Wilton A, Shearman J, Musiani M, Cargill M, Jones PG, Qian Z, Huang W, Ding ZL, Zhang YP, **Bustamante CD**, Ostrander EA, Novembre J, Wayne RK. *Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication*. **Nature.** 2010;464(7290):898-902.
63. Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S, Froment A, Bodo JM, Wambebe C, Tishkoff SA, **Bustamante CD**. *Genome-wide patterns of population structure and admixture in West Africans and African Americans*. **Proc Natl Acad Sci U S A.** 2010;107(2):786-91. PMID: 2818934.

## 2009

62. Gutenkunst RN, Hernandez RD, Williamson SH, **Bustamante CD**. *Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data*. **PLoS Genet.** 2009;5(10):e1000695. PMID: 2760211.
61. Cadieu E, Neff MW, Quignon P, Walsh K, Chase K, Parker HG, Vonholdt BM, Rhue A, Boyko A, Byers A, Wong A, Mosher DS, Elkahouloun AG, Spady TC, Andre C, Lark KG, Cargill M, **Bustamante CD**, Wayne RK, Ostrander EA. *Coat variation in the domestic dog is governed by variants in three genes*. **Science.** 2009;326(5949):150-3. PMID: 2897713.
60. Andres AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN,

- White TJ, Green ED, **Bustamante CD**, Clark AG, Nielsen R. *Targets of balancing selection in the human genome*. **Mol Biol Evol**. 2009;26(12):2755-64. PMID: 2782326.
59. Boyko AR, Boyko RH, Boyko CM, Parker HG, Castelhana M, Corey L, Degenhardt JD, Auton A, Hedimbi M, Kityo R, Ostrander EA, Schoenebeck J, Todhunter RJ, Jones P, **Bustamante CD**. *Complex population structure in African village dogs and its implications for inferring dog domestication history*. **Proc Natl Acad Sci U S A**. 2009;106(33):13903-8. PMID: 2728993.
58. Torgerson DG, Boyko AR, Hernandez RD, Indap A, Hu X, White TJ, Sninsky JJ, Cargill M, Adams MD, **Bustamante CD**, Clark AG. *Evolutionary processes acting on candidate cis-regulatory regions in humans inferred from patterns of polymorphism and divergence*. **PLoS Genet**. 2009;5(8):e1000592. PMID: 2714078.
57. Parker HG, VonHoldt BM, Quignon P, Margulies EH, Shao S, Mosher DS, Spady TC, Elkahoun A, Cargill M, Jones PG, Maslen CL, Acland GM, Sutter NB, Kuroki K, **Bustamante CD**, Wayne RK, Ostrander EA. *An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs*. **Science**. 2009;325(5943):995-8. PMID: 2748762.
56. McNally KL, Childs KL, Bohnert R, Davidson RM, Zhao K, Ulat VJ, Zeller G, Clark RM, Hoen DR, Bureau TE, Stokowski R, Ballinger DG, Frazer KA, Cox DR, Padhukasahasram B, **Bustamante CD**, Weigel D, Mackill DJ, Bruskiwich RM, Ratsch G, Buell CR, Leung H, Leach JE. *Genomewide SNP variation reveals relationships among landraces and modern varieties of rice*. **Proc Natl Acad Sci U S A**. 2009;106(30):12273-8. PMID: 2718348.
55. Gibbs RA, Taylor JF, Van Tassell CP, ... , **Bustamante CD**, ..., Dodds KG. *Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds*. **Science**. 2009;324(5926):528-32. PMID: 2735092.
54. Nielsen R, Hubisz MJ, Hellmann I, Torgerson D, Andres AM, Albrechtsen A, Gutenkunst R, Adams MD, Cargill M, Boyko A, Indap A, **Bustamante CD**, Clark AG. *Darwinian and demographic forces affecting human protein coding genes*. **Genome Res**. 2009;19(5):838-49. PMID: 2675972.
53. Santos VJ, **Bustamante CD**, Valero-Cuevas FJ. *Improving the fitness of high-dimensional biomechanical models via data-driven stochastic exploration*. **IEEE Trans Biomed Eng**. 2009;56(3):552-64. PMID: 2841988.
52. Lohmueller KE, **Bustamante CD**, Clark AG. *Methods for human demographic inference using haplotype patterns from genomewide single-nucleotide polymorphism data*. **Genetics**. 2009;182(1):217-31. PMID: 2674818.
51. Auton A, Bryc K, Boyko AR, Lohmueller KE, Novembre J, Reynolds A, Indap A, Wright MH, Degenhardt JD, Gutenkunst RN, King KS, Nelson MR, **Bustamante CD**. *Global distribution of genomic diversity underscores rich complex history of continental human populations*. **Genome Res**. 2009;19(5):795-803. PMID: 2675968.
50. Anderson TM, vonHoldt BM, Candille SI, Musiani M, Greco C, Stahler DR, Smith DW, Padhukasahasram B, Randi E, Leonard JA, **Bustamante CD**, Ostrander EA, Tang H, Wayne RK, Barsh GS. *Molecular and evolutionary history of melanism in North American gray wolves*. **Science**. 2009;323(5919):1339-43. PMID: 2903542.
49. Gray MM, Granka JM, **Bustamante CD**, Sutter NB, Boyko AR, Zhu L, Ostrander EA, Wayne RK. *Linkage disequilibrium and demographic history of wild and domestic canids*. **Genetics**. 2009;181(4):1493-505. PMID: 2666515.

48. Degenhardt JD, de Candia P, Chabot A, Schwartz S, Henderson L, Ling B, Hunter M, Jiang Z, Palermo RE, Katze M, Eichler EE, Ventura M, Rogers J, Marx P, Gilad Y, **Bustamante CD**. *Copy number variation of CCL3-like genes affects rate of progression to simian-AIDS in Rhesus Macaques (Macaca mulatta)*. **PLoS Genet**. 2009;5(1):e1000346. PMID: 2621346.
47. **Bustamante CD**, Ramachandran S. *Evaluating signatures of sex-specific processes in the human genome*. **Nat Genet**. 2009;41(1):8-10. PMID: 2766813.

## 2008

46. Brooks MB, Catalfamo JL, Etter K, Brisbin A, **Bustamante CD**. *Exclusion of ABCA-1 as a candidate gene for canine Scott syndrome*. **J Thromb Haemost**. 2008;6(9):1608-10. PMID: 2766810.
45. Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, **Bustamante CD**. *Genes mirror geography within Europe*. **Nature**. 2008;456(7218):98-101. PMID: 2735096.
44. Nelson MR, Bryc K, King KS, Indap A, Boyko AR, Novembre J, Briley LP, Maruyama Y, Waterworth DM, Waeber G, Vollenweider P, Oksenberg JR, Hauser SL, Stirnadel HA, Kooner JS, Chambers JC, Jones B, Mooser V, **Bustamante CD**, Roses AD, Burns DK, Ehm MG, Lai EH. *The Population Reference Sample, POPRES: a resource for population, disease, and pharmacological genetics research*. **Am J Hum Genet**. 2008;83(3):347-58. PMID: 2556436.
43. Kosiol C, Vinar T, da Fonseca RR, Hubisz MJ, **Bustamante CD**, Nielsen R, Siepel A. *Patterns of positive selection in six Mammalian genomes*. **PLoS Genet**. 2008;4(8):e1000144. PMID: 2483296.
42. Blekhman R, Man O, Herrmann L, Boyko AR, Indap A, Kosiol C, **Bustamante CD**, Teshima KM, Przeworski M. *Natural selection on genes that underlie human disease susceptibility*. **Curr Biol**. 2008;18(12):883-9. PMID: 2474766.
41. Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, Lohmueller KE, Adams MD, Schmidt S, Sninsky JJ, Sunyaev SR, White TJ, Nielsen R, Clark AG, **Bustamante CD**. *Assessing the evolutionary impact of amino acid mutations in the human genome*. **PLoS Genet**. 2008;4(5):e1000083. PMID: 2377339.
40. Padhukasahasram B, Marjoram P, Wall JD, **Bustamante CD**, Nordborg M. *Exploring population genetic models with recombination using efficient forward-time simulations*. **Genetics**. 2008;178(4):2417-27. PMID: 2323826.
39. Shimizu KK, Reininga JM, Caicedo AL, Mays CA, Moore RC, Olsen KM, Ruzsa S, Coop G, **Bustamante CD**, Purugganan MD. *Retraction*. **Science**. 2008;320(5873):176.
38. Lohmueller KE, Indap AR, Schmidt S, Boyko AR, Hernandez RD, Hubisz MJ, Sninsky JJ, White TJ, Sunyaev SR, Nielsen R, Clark AG, **Bustamante CD**. *Proportionally more deleterious genetic variation in European than in African populations*. **Nature**. 2008;451(7181):994-7. PMID: 2923434.
37. Huerta-Sanchez E, Durrett R, **Bustamante CD**. *Population genetics of polymorphism and divergence under fluctuating selection*. **Genetics**. 2008;178(1):325-37. PMID: 2206081.

## 2007



36. Nielsen R, Hellmann I, Hubisz M, **Bustamante C**, Clark AG. *Recent and ongoing selection in the human genome.* **Nat Rev Genet.** 2007;8(11):857-68. PMID: 2933187.
35. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR, **Bustamante CD**, Purugganan MD. *Genome-wide patterns of nucleotide polymorphism in domesticated rice.* **PLoS Genet.** 2007;3(9):1745-56. PMID: 1994709.
34. Sweeney MT, Thomson MJ, Cho YG, Park YJ, Williamson SH, **Bustamante CD**, McCouch SR. *Global dissemination of a single mutation conferring white pericarp in rice.* **PLoS Genet.** 2007;3(8):e133. PMID: 1941752.
33. Hernandez RD, Williamson SH, Zhu L, **Bustamante CD**. *Context-dependent mutation rates may cause spurious signatures of a fixation bias favoring higher GC-content in humans.* **Mol Biol Evol.** 2007;24(10):2196-202.
32. Jensen JD, Thornton KR, **Bustamante CD**, Aquadro CF. *On the utility of linkage disequilibrium as a statistic for identifying targets of positive selection in nonequilibrium populations.* **Genetics.** 2007;176(4):2371-9. PMID:1950638.
31. Williamson SH, Hubisz MJ, Clark AG, Payseur BA, **Bustamante CD**, Nielsen R. *Localizing recent adaptive evolution in the human genome.* **PLoS Genet.** 2007;3(6):e90. PMID: 1885279.
30. Hernandez RD, Williamson SH, **Bustamante CD**. *Context dependence, ancestral misidentification, and spurious signatures of natural selection.* **Mol Biol Evol.** 2007;24(8):1792-800.
29. Mosher DS, Quignon P, **Bustamante CD**, Sutter NB, Mellersh CS, Parker HG, Ostrander EA. *A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs.* **PLoS Genet.** 2007;3(5):e79. PMID: 1877876.
28. Gao H, Williamson S, **Bustamante CD**. *A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data.* **Genetics.** 2007;176(3):1635-51. PMID: 1931536.
27. Hernandez RD, Hubisz MJ, Wheeler DA, Smith DG, Ferguson B, Rogers J, Nazareth L, Indap A, Bourquin T, McPherson J, Muzny D, Gibbs R, Nielsen R, **Bustamante CD**. *Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques.* **Science.** 2007;316(5822):240-3.
26. Gibbs RA, Rogers J, Katze MG, ... , **Bustamante CD**, ..., Smith KE, Zwiag AS. *Evolutionary and biomedical insights from the rhesus macaque genome.* **Science.** 2007;316(5822):222-34.
25. Sutter NB, **Bustamante CD**, Chase K, Gray MM, Zhao K, Zhu L, Padhukasahasram B, Karlins E, Davis S, Jones PG, Quignon P, Johnson GS, Parker HG, Fretwell N, Mosher DS, Lawler DF, Satyaraj E, Nordborg M, Lark KG, Wayne RK, Ostrander EA. *A single IGF1 allele is a major determinant of small size in dogs.* **Science.** 2007;316(5821):112-5. PMID: 2789551.
24. Abecasis G, Tam PK, **Bustamante CD**, Ostrander EA, Scherer SW, Chanock SJ, Kwok PY, Brookes AJ. *Human Genome Variation 2006: emerging views on structural variation and large-scale SNP analysis.* **Nat Genet.** 2007;39(2):153-5.

23. Pollinger JP, **Bustamante CD**, Fledel-Alon A, Schmutz S, Gray MM, Wayne RK. *Selective sweep mapping of genes with large phenotypic effects*. **Genome Res**. 2005;15(12):1809-19. PMID: 1356119.
22. Nielsen R, Williamson S, Kim Y, Hubisz MJ, Clark AG, **Bustamante C**. *Genomic scans for selective sweeps using SNP data*. **Genome Res**. 2005;15(11):1566-75. PMID: 1310644.
21. Clark AG, Hubisz MJ, **Bustamante CD**, Williamson SH, Nielsen R. *Ascertainment bias in studies of human genome-wide polymorphism*. **Genome Res**. 2005;15(11):1496-502. PMID: 1310637.
20. **Bustamante CD**, Fledel-Alon A, Williamson S, Nielsen R, Hubisz MT, Glanowski S, Tanenbaum DM, White TJ, Sninsky JJ, Hernandez RD, Civello D, Adams MD, Cargill M, Clark AG. *Natural selection on protein-coding genes in the human genome*. **Nature**. 2005;437(7062):1153-7.
19. Dimmic MW, Hubisz MJ, **Bustamante CD**, Nielsen R. *Detecting coevolving amino acid sites using Bayesian mutational mapping*. **Bioinformatics**. 2005;21 Suppl 1:i126-35.
18. Jensen JD, Kim Y, DuMont VB, Aquadro CF, **Bustamante CD**. *Distinguishing between selective sweeps and demography using DNA polymorphism data*. **Genetics**. 2005;170(3):1401-10. PMID: 1451184.
17. Williamson SH, Hernandez R, Fledel-Alon A, Zhu L, Nielsen R, **Bustamante CD**. *Simultaneous inference of selection and population growth from patterns of variation in the human genome*. **Proc Natl Acad Sci U S A**. 2005;102(22):7882-7. PMID: 1142382.
16. Zhu L, **Bustamante CD**. *A composite-likelihood approach for detecting directional selection from DNA sequence data*. **Genetics**. 2005;170(3):1411-21. PMID: 1451173.
15. Nielsen R, **Bustamante C**, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, J JS, Adams MD, Cargill M. *A scan for positively selected genes in the genomes of humans and chimpanzees*. **PLoS Biol**. 2005;3(6):e170. PMID: 1088278.
14. Dimmic MW, Sunyaev S, **Bustamante CD**. *Inferring SNP function using evolutionary, structural, and computational methods*. **Pac Symp Biocomput**. 2005:382-4.
13. Shimizu KK, Cork JM, Caicedo AL, Mays CA, Moore RC, Olsen KM, Ruzsa S, Coop G, **Bustamante CD**, Awadalla P, Purugganan MD. *Darwinian selection on a selfing locus*. **Science**. 2004;306(5704):2081-4.
12. Williamson S, Perry SM, **Bustamante CD**, Orive ME, Stearns MN, Kelly JK. *A statistical characterization of consistent patterns of human immunodeficiency virus evolution within infected patients*. **Mol Biol Evol**. 2005;22(3):456-68.
11. Williamson S, Fledel-Alon A, **Bustamante CD**. *Population genetics of polymorphism and divergence for diploid selection models with arbitrary dominance*. **Genetics**. 2004;168(1):463-75. PMID: 1448126.
10. Sawyer SA, Kulathinal RJ, **Bustamante CD**, Hartl DL. *Bayesian analysis suggests that most amino acid replacements in Drosophila are driven by positive selection*. **J Mol Evol**. 2003;57

Suppl 1:S154-64.

9. Gilad Y, **Bustamante CD**, Lancet D, Paabo S. *Natural selection on the olfactory receptor gene family in humans and chimpanzees*. **Am J Hum Genet**. 2003;73(3):489-501. PMID: 1180675.
8. Barrier M, **Bustamante CD**, Yu J, Purugganan MD. *Selection on rapidly evolving proteins in the Arabidopsis genome*. **Genetics**. 2003;163(2):723-33. PMID: 1462452.
7. **Bustamante CD**, Nielsen R, Hartl DL. *Maximum likelihood and Bayesian methods for estimating the distribution of selective effects among classes of mutations using DNA polymorphism data*. **Theor Popul Biol**. 2003;63(2):91-103.
6. **Bustamante CD**, Nielsen R, Sawyer SA, Olsen KM, Purugganan MD, Hartl DL. *The cost of inbreeding in Arabidopsis*. **Nature**. 2002;416(6880):531-4.
5. **Bustamante CD**, Wakeley J, Sawyer S, Hartl DL. *Directional selection and the site-frequency spectrum*. **Genetics**. 2001;159(4):1779-88. PMID: 1461920.
4. **Bustamante CD**, Nielsen R, Hartl DL. *A maximum likelihood method for analyzing pseudogene evolution: implications for silent site evolution in humans and rodents*. **Mol Biol Evol**. 2002;19(1):110-7.
3. Nurminsky D, Aguiar DD, **Bustamante CD**, Hartl DL. *Chromosomal effects of rapid gene evolution in Drosophila melanogaster*. **Science**. 2001;291(5501):128-30.
2. **Bustamante CD**, Townsend JP, Hartl DL. *Solvent accessibility and purifying selection within proteins of Escherichia coli and Salmonella enterica*. **Mol Biol Evol**. 2000;17(2):301-8.
1. Jimenez JJ, Alvarez E, **Bustamante CD**, Yunis AA. *Pretreatment with 1,25(OH)<sub>2</sub>D<sub>3</sub> protects from Cytoxan-induced alopecia without protecting the leukemic cells from Cytoxan*. **Am J Med Sci**. 1995;310(2):43-7.

### Book Section (3)

1. Kulathinal R, Sawyer SA, **Bustamante CD**, Nurminsky DI, Ponce R, Ranz JM, Hartl DL. *Selective Sweep in the Evolution of a New Sperm-Specific Gene in Drosophila*. In: Nurminsky D, editor. *Selective Sweeps*; Springer; 2005.
2. Przeworski M, **Bustamante CD**. *Genetic signatures of Natural Selection*. *Encyclopedia of Genetics*. London, U.K.: Wiley & Sons; 2005.
3. Ramachandran S, Tang H, Gutenkunst R, **Bustamante C**. *Genetics and Genomics of Human Population Structure*. In: Speicher MR, Motulsky AG, Antonarakis SE, editors. *Vogel and Motulsky's Human Genetics: Problems and Approaches*. Hiedelberg: Springer; 2010. p. 589-615.

### Research Support

#### Ongoing Research Support (grants transferred while CDB on leave)

U41HG009649 (Bustamante [C])

08/01/2017 – 07/31/2021

SPO# 109590

NIH

Clinical Genome Resource (ClinGen)

Major goal is to create a unified, public, and freely available database of genetic alterations relevant to clinical care. Our ultimate goal is to empower clinicians, genetic counselors, and patients to make informed decisions based on DNA testing.

Role: PI

1U01HG009080 (Bustamante [C])

05/20/2016 – 03/31/2020

SPO# 121626

NIH

Center for Multi- and Trans-ethnic Mapping of Mendelian and Complex Disease

Major goal is to accelerate the discovery of DNA variation relevant to health and disease by analyzing data from over 225,000 ethnically and racially diverse patients that will undergo genome sequencing. Of particular importance is ensuring we have powerful statistical methods for analyzing data from underserved groups including U.S. minority populations. Achieving this goal requires expertise across many domains of knowledge including: medical and population genomics, algorithm development for disease mapping, and expertise in large-scale databases.

Role: PI

5R24 HL117756-02 (Wu)

04/15/2014 – 03/31/2019

NIH/NHLBI

SPO# 110465/ Cardiovascular Institute Operations

Biorepository of Human iPSCs for Studying Dilated and Hypertrophic Cardiomyopathy

The project seeks to create a bank of 600 iPSC lines created from cardiovascular patients and distribute to the NHLBI community.

Role: Co-PI

1U54MD01724-01 (Cullen[C], Bustamante)

04/01/2016 – 03/31/2021

SPO# 121927

NIH/NIA

Stanford Precision Health for Ethnic and Racial Equity (SPHERE)

Major goal is to provide services to the different projects in the center as well as any other group in the consortium. We already have a robust infrastructure including equipment, expertise, and personnel for performing state-of-the-art genotyping, genome and exome sequencing, transcriptomics, metabolomics, and microbiome characterization. These activities will be implemented into the individual projects.

Role: Lab Core PI

1U24EB02367401 (Ashley)

12/12/2016 – 11/30/2022

SPO# 124803

NIH

Stanford MoTrPAC Bioinformatics Center

Major goal is to manage the bioinformatics for the MoTrPAC Clinical Site

Role: Co-PI

**Completed Research Support (including grants with No-Cost Extensions expiring by 9/1/2019)**

1R01HG008150-01 (Montgomery)

08/01/2015 – 06/30/2018

SPO# 114738

NIH

SPO# 114738/ Pathology Research

Predicting casual non-coding variants in a founder population

The project seeks to develop methods for interpreting non-coding genetic variation: to provide a Bayesian method for predicting causal variants from full genomes, to identify and validate a large set of functional non-coding variants using CRISPR technology, and to predict disease-relevant traits likely to be affected by each variant.

Role: Co-PI

1U24HG008956-01 (Matis [C], Bustamante)

01/14/2016 – 11/30/2019

SPO# 120615

NIH

Rutgers University

NHGRI Genome Sequencing Program Coordinating Center

Major goal is to identify of relevant control sets for data sets coming through GSPCC, and on the use of diverse trans-ethnic populations in boosting power, identifying novel causal variants, and characterizing the genetic background of the common control resource.

Role: Subcontract PI

IOS-1642184 (Bustamante)

08/01/2016 – 07/31/2018

SPO# 124489

NSF

EAGER: Establishing the genetic basis of hibernation by building and utilizing a next-generation genomics resource for the model hibernator, the thirteen-lined ground squirrel

Major goal is to establish a genetic basis for the hibernating phenotype in the 13-lined ground squirrel by developing a high-quality genome resource for the species. Specifically, we will employ quantitative trait loci analysis to identify genetic variants associated with hibernation to determine whether hibernation-related traits are heritable.

Role: PI

U01FD004979 (Altman)

02/01/2015 – 03/31/2018

SPO# 119541

UCSF

UCSF-Stanford CERSI: Improving the Efficiency and Rigor of Pharmacovigilance (PV) at FDA

Role: Co-PI

4U01HG0074919-04 (Matisse [C], Bustamante)

09/01/2013 – 05/31/2018

SPO# 109405

Rutgers University (Prime NIH/NHGRI)

NHGRI PAGE Coordinating Center

The major goals of this study are to provide flexible support in data analysis, management, synthesis, and dissemination, and program logistics, for the PAGE II studies.

Role: Subcontract PI

4U01HG007417 (Loos [C], Bustamante)

09/01/2013 – 05/31/2018

SPO# 109391

MSSM (Prime NIH/NHGRI)

Diverse ancestry biobank to map biomedical traits and elucidate health disparities

The major goal of this project is to characterize the population architecture of differences in disease burden among populations in a multi-ethnic cohort using common and rare coding variants. We will also develop methods for disease mapping based on local ancestry information and will assess aggregate effects of rare variants in the context of the PAGE II network.

Role: Co-PI

66864T (Bustamante, Martin, PIs)  
NIH/NIGMS

09/01/2011 – 08/31/2017

Genomic Origins and Admixture in Latinos (GOAL)

The major goal of this project is to generate and analyze dense genotype data from 180 parent-offspring triads of diverse Hispanic/Latino ancestry sampled in Miami, as well as sequence the complete genomes of 30 H/L triads.

ED-9697-STS2 (Bustamante, Dermitzakis[C], Guigo, McCarthy) 08/01/2013 – 06/30/2016 NIH  
Methods for High-Resolution Analysis of Genetic Effects on Gene Expression

The major goal of this project is to identify genetic effects on gene expression using GTEx data. Specifically, we will provide statistical expression quantitative trait methodology, network methods for detecting gene regulatory interactions, and bioinformatics methods for improving identification of allele-specific expression.

0255-6424-4609 (Loos, Bustamante)  
NIH/NHGRI

09/01/2013 - 05/31/2017

Diverse Ancestry Biobank to Map Biomedical Traits and Elucidate Health Disparities

The major goal of this project is to characterize the population architecture of differences in disease burden among populations in a multi-ethnic cohort using common and rare coding variants. We will also develop methods for disease mapping based on local ancestry information and will assess aggregate effects of rare variants in the context of the PAGE II network.

5943 (Matisse)  
NIH/NHGRI

09/01/2013 - 05/31/2017

NHGRI PAGE Coordinating Center

The major goals of this study are to provide flexible support in data analysis, management, synthesis, and dissemination, and program logistics, for the PAGE II studies.

2R01HG003229 (Clark, Bustamante, Nielsen, PIs)  
NIH/NHGRI

01/01/2010 – 6/30/2013

Population Genetic Inferences from Dense Genotype Data

The major goals of this project are to develop methods for population genetic inference from next-generation sequence data, including ascertainment bias correction, inference of natural selection, and prediction of functionality for novel SNPs and CNVs in the human genome.

U01HG005715 (Bustamante, PI)  
NIH/NHGRI

09/09/2010 – 06/30/2014

Population Structure Admixture and Selection across the 1000 Genomes Data Set

The main goals of the project are to develop novel approaches for characterizing population structure, analyzing patterns of admixture, and localizing signatures of selection across the 2,500 samples sequenced by the 1000 Genomes Project.

58-6631-0-100 (Bustamante, PI)

07/01/2010 – 01/07/2014

USDA

Identification of Genes Regulating Major Agronomic Traits in Cacao Through Association and Linkage Mapping Using Large Sequence Data Sets

The major goal of this project is provide analytical and bioinformatic support to the MARS-USDA *Theobroma cacao* L. sequencing project, including SNP diversity and association mapping analysis of cacao populations as well as gene expression analysis of diverse *Theobroma* species.

1R01GM083606 (Bustamante, PI)  
NIH/NIGMS

08/01/2007–07/31/2012

Association Mapping in Structured Populations

The major goal of this project is to develop novel statistical methods for association mapping in structured populations and to use these methods to map several complex traits in ethnically diverse African populations.

1R01MH084695 (Clark, Bustamante, Nielsen, PIs)  
NIH/NIMH

09/28/2008–07/31/2012

Inferring multiple-SNP disease association with DNA resequence data

The major goal of this project is to develop novel statistical methods for association mapping of genetic variants underlying risk for complex disease, using resequenced DNA samples from cases and controls.

DBI 0701382 (Purugganan, PI, Bustamante, Co-PI)  
NSF/DBI

09/01/2007–08/31/2012

GEPR: The Evolutionary Genomics of Rice Domestication

The major goals of this project are to fine-map domestication genes and examine the role of evolutionary selection in the origin of cultivated rice.

DEB 0516310 (Bustamante)

09/01/2005 – 08/31/2010 NSF

Computational Methods for Detecting Natural Selection using Comparative Population Genomic Data

The major goal of this project is to develop novel population genetic methods for identifying genes and genomic regions that are involved in adaptive molecular evolution through the comparison both within and between species genomic variation data.

5U01HL084706 -03 (Clark, PI, Bustamante, Co-PI)  
NIH/NHLBI

06/15/2006 – 05/31/2010

Dimension Reduction Approaches for Genome-Wide Association Testing

The major goal of this project is to develop novel Bayesian shrinkage methods for genome-wide association mapping and to apply these methods to genome-wide expression level and genotype data from the 270 HapMap cell lines in order to map expression quantitative trait loci (eQTLs).

BIO 0948510 (Bustamante, PI)  
NSF/DBI/EAGER

08/15/2009–7/31/2011 (NCE)

Genome-wide diversity in village dogs: uncovering the history of dog domestication, the genetics of natural dog populations, and consequences of artificial selection in breed dogs.

The major goals of this project are to develop population genetic methods for inferring the evolutionary history of dog domestication from genotyping and sequencing data, and to determine the impact of dog domestication on the genetic basis of complex traits in dogs.

DBI 0606461 (McCouch, PI, Bustamante, Co-PI)

08/01/2006 – 07/31/2011 (NCE)

NSF/DBI

GEPR: Exploring the Genetic Basis of Transgressive Variation in Rice

The major goal of this project is to perform association mapping in rice (*Oryza sativa*) using 500 individuals genotyped at 10,000 SNPs in order to identify quantitative trait loci underlying 26 agronomic traits of interests.

1R01AI073368 (Stanhope, PI, Bustamante, Co-PI) 5/1/2010–6/15/2011 NIH/NIAID

#### Evolutionary Genomics and Population Genetics of Pathogenic Streptococci

The major goal of this project is to reach a thorough understanding of the molecular specifics correlated with adaptive differences within and between the pathogenic taxa of the genus *Streptococcus*.

673446Z (Nickerson, PI, Bustamante, Co-PI) 09/30/2009 - 06/30/2012 NIH

#### Northwest Genomic Center

The main goals of my subcontract on this project is to develop population genetic methods for quantifying population structure from whole-genome exome data and estimate selection coefficients on different categories of functional sites in the human genome.

### Educational and Outreach Activities

#### **Post-doctoral fellows trained**

1. Dr. Adam Auton, D. Phil. (10/2007—6/2009), Statistical Geneticist, 23andme
2. Dr. Maria Avila-Arcos, Ph.D. (05/2014—12/2017), Assistant Professor, National Autonomous University of Mexico
3. Dr. Linda Boettger (Goodman), Ph.D. (1/2018 – 8//2018), Co-founder and CTO at Fauna Bio
4. Dr. Adam Boyko, Ph. D. (9/2005—8/2011), Associate Professor, Cornell University
5. Dr. Jake Byrnes, Ph. D. (9/2010—10/2011), Population Genomics Manager, Ancestry.com
6. Dr. Meredith Carpenter, Ph. D. (7/2011—1/2014), Chief Scientific Officer, ArcBio LLC
7. Dr. Jessica Chen, Ph.D. (3/2017 – 9/2019), Research Scientist, Genentech
8. Dr. Omar Cornejo, Ph. D. (8/2009—8/2013), Assistant Professor, Washington State University
9. Dr. Christopher DeBoever, Ph.D. (10/2016—4/2018), Vice President of Platform at Data4Cure, Inc.
10. Dr. Matthew J. Dimmic, Ph. D. (3/2003—2005), Genomics and Data Science Lead, Monsanto
11. Dr. Anna Katharina Elmer, Ph.D. (4/2014—9/2014), Senior Scientist, ArcBio LLC
12. Dr. Rosa Lorenzo Fregel, Ph.D. (6/2014—8/2017), Assistant Professor Tenerife
13. Dr. Christopher Gignoux, Ph.D. (8/2013—2/2017), Associate Professor, University of Colorado
14. Dr. Ryan Gutenkunst, Ph. D. (1/2008—12/2008), Assistant Professor, University of Arizona
15. Dr. Katherine Grabek, Ph.D. (7/2014—8/2018), Co-founder and CSO at Fauna Bio
16. Dr. Simon Gravel, Ph.D. (10/2009—12/2012), Assistant Professor, McGill University
17. Dr. Ragan Hart, Ph.D. (9/2018—6/2019), Entrepreneur-In-Residence, F-Prime Capital
18. Dr. Brenna Henn, Ph. D. (12/2009—12/2012), Assistant Professor, SUNY Stony Brook
19. Dr. Patrick House, Ph. D. (9/2014—9/2015), Neuroscientist and writer
20. Dr. Aashish Jha, Ph.D. (9/2015—6/2019), Director of Bioinformatics, NomNomNow, Berkeley, CA
21. Dr. Alex Ioannidis, Ph.D. (2018—current), Postdoctoral scholar, Stanford
22. Dr. Joanna Kelley, Ph. D. (8/2010—8/2013), Assistant Professor, Washington State University
23. Dr. Eimear Kenny, Ph. D. (10/2010—10/2012), Assistant Professor, Mount Sinai School of Medicine
24. Dr. Jeffrey Kidd, Ph. D. (2/2010—12/2011), Assistant Professor, University of Michigan School of Medicine
25. Dr. Carolin Kosiol, Ph. D. (9/2006—1/2009), Bioinformatics Group Leader, University of Veterinary Medicine (vetmeduni) Vienna, Austria



26. Dr. Tuuli Lappalainen (1/2013—12/2013), Assistant Professor, Department of Systems Biology, Columbia University and Junior Investigator and Core Member, New York Genome Center
27. Dr. Arturo Lopez Pineda (1/2016—4/2019), Computational biologist, Genentech
28. Dr. Nilah Monnier Ioannidis, Ph.D. (9/2013—10/2019), Assistant Professor, U. California Berkeley
29. Dr. Andres Moreno Estrada, M.D./Ph. D. (3/2009—12/2014), Professor and Group Leader, Langebio (Mexico)
30. Dr. Fernando Mendez, Ph.D., (7/2013 – 7/2017), Senior Scientist - Population Geneticist at Helix
31. Dr. Marina Muzzio, Ph. D. (8/2011—8/2013), Researcher, Instituto Multidisciplinario de Biología Celular (IMBICE), National Scientific and Technical Research Council (CONICET), Buenos Aires, Argentina
32. Dr. Sean Myles, Ph.D. (1/2008—7/2010), Assistant Professor, Dalhousie University
33. Dr. Badri Padhukasahasram, Ph. D. (9/2006—1/2008), Bioinformatics Scientist, Illumina, Inc.
34. Dr. Alice Popejoy, Ph.D. (11/2017 – current)
35. Dr. Morten Rasmussen, Ph.D., (8/2013—7/2016), Director of Technology, ArcBio, LLC
36. Dr. Arindam RoyChoudhury, Ph. D. (7/2008—8/2009), Assistant Professor, Columbia University
37. Dr. Anna Rychkova, Ph.D. (9/2013—04/2016), Scientist, Amgen Comparative Biology and Safety Sciences Department
38. Dr. Suyash Shringarpure, Ph.D. (9/2012—06/2016), Statistical Geneticist at 23andme
39. Dr. Karla Sandoval Mendoza, Ph.D. (4/2010—12/2014), Postdoctoral Scholar, Langebio (Mexico)
40. Dr. Martin Sikora, Ph. D. (7/2010—6/2014), Associate Professor and Group Leader, Center for GeoGenetics, Copenhagen, Denmark
41. Dr. Elena Sorokin, Ph. D. (9/2015—2/28), Statistical geneticist at Ancestry DNA
42. Dr. Scott Williamson, Ph. D. (6/2003—6/2006), deceased (was Assistant Professor at Cornell University)
43. Dr. Genevieve Wojcik, Ph. D. (1/2014—9/2019), Assistant Professor, John Hopkins School of Public Health
44. Dr. Keyan (Kevin) Zhao, Ph.D. (9/2007—12/2010), Staff Scientist, IonTorrent
45. Dr. Ashley Zehnder, DVM, Ph.D. (10/2016—8/2018), Co-founder and CEO at Fauna Bio

### **Graduate Students Trained**

1. Abra Brisbin (Ph. D. Applied Math) (2006 – 2010, joint with J. Mezey), Assistant Professor of Mathematics, University of Wisconsin-Eau Claire
2. Kasia Bryc (Ph. D. Biometry) (2006 — 2010), Population Geneticist, 23andMe
3. Thomas Cooke (Ph. D. Genetics) (2011—2017), Postdoc at Whitehead Institute
4. Helio Costa (Ph. D. Genetics) (2011—2015), Fellow in Molecular Clinical Genetics at Stanford
5. Roxana Daneshjou (M.D./Ph.D. Genetics) (2012—2014, joint w/ Russ Altman), MSTP student at Stanford, completing medical training
6. Joe Davis (Ph. D. Genetics) (2012—2017, joint with S. Montgomery), Data Scientist at Cardinal Analytix Solutions
7. Jeremiah Degenhardt (Ph. D. Biometry) (2005 — 2010), Senior Scientist in Bioinformatics, Onyx Pharmaceuticals
8. Kirsten Eilertson (Ph. D. Statistics) (2007—2011), Assistant Professor, Penn State
9. Hong Gao (Ph. D. Genetics) (2004—2008), Senior Bioinformatics Scientist, Illumina, Inc.
10. Alex Han (Ph.D. Genetics) (2014—current)
11. Ryan Hernandez (Ph. D. Biometry) (2003—2008), Associate Professor, UCSF
12. Julian Homburger (Ph. D. Genetics) (2014—2017, joint with E. A. Ashley), Data Scientist at Color Genomics
13. Alex Ioannidis (Ph.D. Computational and Mathematical Engineering) (2013—2018), Postdoctoral scholar, Stanford

14. Sandra Indarraga (M.S./Ph. D. Genetics) (2009—2011), self-employed
15. Emilia Huerta-Sanchez (Ph. D. Applied Math, joint with R. Durrett) (2005—2009), Assistant Professor, UC Merced
16. Jeffrey Jensen (Ph. D. Genetics & Development, joint with C.F. Aquadro) (2002—2006), Head of Laboratory and Professor, Ecole Polytechnique Federale de Laussane
17. Lin Lin (Ph. D. Computational Biology) (2006—2010), Postdoctoral Scholar, Harvard University
18. Kirk Lohmueller (Ph. D. Genetics & Development) (2005 – 2009, joint with A. G. Clark), Assistant Professor, UCLA
19. Xin (Maria) Ma (Ph. D. Statistics) (2008—2011), Postdoctoral Scholar, Stanford University
20. Brian Maples (Ph.D. Biomedical Informatics) (2011–2014), Computational Data Scientist, Nuna, Inc.
21. Alicia Martin (Ph.D. Genetics) (2011—2015), Postdoctoral Scholar, Harvard University
22. Kimberly McManus (M.S. Genetics) (2012—2015), Data Scientist, LinkedIn
23. Shaila Musharoff (Ph.D. Genetics) (2009—2015), Postdoctoral Scholar, UC San Francisco
24. Patricia Ortiz-Tello (M.D./Ph.D. Genetics) (2010—2016), MSTP student at Stanford, completed medical training
25. Armin Pourshafeie (Ph.D. Physics) (2015—current)
26. David Poznik (Ph.D. Biomedical Informatics) (2011—2015), Computational Biologist and Population Geneticist, 23andme
27. Jessica Torres (Ph.D. Biomedical Informatics) (2014—current)
28. Mark Wright (Ph.D. Genetics & Development) (2007—2010), 23andMe
29. Fouad Zakharia (Ph.D. Genetics) (2010—2012), Vice President, Morgan Stanley
30. Lan Zhu (M.S. /Ph.D. Biometry) (2002—2006), Associate Professor (tenured), Oklahoma State University

### **Courses Taught**

- BIODS210: Configuration of the Health Care System and the Application of Big Data/Analytics (F2016)
- BTRY 407: Principles of Probability and Theoretical Statistics (F2007)
- BTRY 694: Statistical inference in Population Genomics (S2004,S2006, F2006, F2007, S2008, F2010)
- BTRY 484/684: Statistical Genomics (F2005, F2006)
- BTRY 694: Statistical Inference in Evolutionary Genomics (S2005)
- BTRY 384: Intro. Bioinformatics and Statistical Genomics (F2004, F2005)
- BTRY 652: Computationally Intensive Statistical Inference (S2003)
- CSRE 110: Comparative Studies of Race and Ethnicity (W2011, guest lecture)
- GENE 210: Personalized Genomics (S2010, S2011, guest lecture)
- GENE 203: Graduate Genetics (S2013, F2010, lectures on human genetics and genomics)
- GENE 203: Graduate Genetics (F2011, primary lecturer)
- GENE 202: Genetics for Medical Students (F2010, lectures on human population genetics)
- GENE 209: Population Genetics (S2011, S2012, S2013, S2014)

### **Invited Conferences, Keynote Lectures, and Departmental Seminars (>175 total)**

#### 2019

1. Etalon Diagnostic Leadership Retreat, Santa Rosa, CA
2. Helix Leadership Retreat, Menlo Park, CA
3. DRIFT Symposium, American Society of Human Genetics 2019, Houston, TX
4. NHGRI Workshop: Perspectives in Comparative Genomics & Evolution, Washington, D.C.

5. Human Genetics and Genomics Gordon Research Conference 2019, Closing Keynote, Waterville, NH
6. Stanford Executive Summer Program, Stanford University Graduate School of Business
7. STAT+ Inaugural Event on Venture Capital in Pharma and Biotechnology, San Francisco, CA
8. RECOMB 2019 Keynote Speaker, Washington, D.C.
9. California Institute of Regenerative Medicine (CIRM) Conference on Brainstorming Neurodegeneration” Keynote, San Francisco, CA
10. Nation Health Care Council (NHCC) Fellow’s Program 2019, Nashville, TN
11. Personalized Medicine World Conference 2019, Santa Clara, CA

## 2018

1. Advances in Genome Biology and Technology Personalized Medicine Keynote, San Diego, CA
2. Genomic Medicine XI, San Diego, CA
3. Genetic Health Information Network Summit, Nashville, TN
4. Siriraj Hospital Genomics and Medicine Keynote, Bangkok, Thailand
5. NHGRI Town Hall, Stanford, CA
6. 2018 Marshall Scholars’ Forum, Stanford, CA
7. Distributed18 (Blockchain in Healthcare) Conference, San Francisco, CA
8. Big Data in Biomedicine 18 (Blockchain organizer), Stanford, CA
9. Stanford Executive Summer Program, Stanford University Graduate School of Business
10. Broad Institute of Harvard/MIT Medical and Population Genomics, Cambridge, MA
11. Frontiers in Genomic Medicine Keynote, Sioux Falls, SD
12. American College of Medical Genetics Annual Meeting Keynote, Charlotte, NC
13. Vanderbilt University Medical Center Personalized Medicine Keynote, Nashville, TN
14. Colorado State University on Veterinary Electronic Health Records, Ft. Collins, CO
15. Pacific Symposium on Biocomputing (PSB 2018) Keynote, Hilo, HI

## 2017

1. Health 2.0 Keynote Speaker, San Francisco, CA
2. 2017 Distinguished Presidential Scholar, University of Miami, FL
3. CSHL’s Second Annual Double Helix Day, NY, NY
4. Stanford Symposium for Genomics & Personalized Medicine, Menlo Park, CA
5. Stanford Executive Summer Program, Stanford University Graduate School of Business
6. Big Data in Biomedicine Conference 2017, Stanford, CA
7. Stanford-China Cardiovascular Research and Medicine Symposium, Stanford, CA
8. FH 2017 Global Summit, Miami, FL
9. Genetic Health Information Network Summit, Nashville, TN
10. IMSS XXVI National Health Research Forum, San Luis Potosí, Mexico

## 2016

1. 2016 Darwin Day Keynote Speaker, University of Iowa, Iowa City, IA
2. University of Iowa Medical School Distinguished Speaker, Iowa City, IA
3. University of Iowa Department of Biology Seminar, Iowa City, IA
4. 2<sup>nd</sup> Annual NIH-wide Pi Day Speaker, National Institutes of Health, Bethesda, MD
5. Personalized Medicine World Conference (PMWC 2016), Silicon Valley, CA
6. Rockefeller University Friday Seminar Speaker, New York, NY
7. UCSF Parnassus Institute for Human Genetics Seminar, San Francisco, CA
8. UCB Supernetwork Conference, London, UK
9. New York City Medical Examiner’s Office, NY, NY

10. SAP Sapphire 2016, Orlando, FL
11. Stanford Institute for Computational and Mathematical Engineering (ICME) Expo, Stanford, CA
12. Miami Children's Hospital Speaker Series, Miami, FL
13. Stanford Pediatric Grand Rounds and Retreat Keynote Speaker, Stanford, CA
14. Big Data in Biomedicine Conference 2016, Stanford, CA
15. 2016 CARDIA Conference Keynote Speaker, Napa, CA
16. Perspectives in Genomics Speaker, Cancun, Mexico
17. UCLA Institute for Pure and Applied Mathematics Seminar, Los Angeles, CA
18. Immunogenomics 2016, HudsonAlpha Institute, Huntsville, AL
19. Baylor College of Medicine Molecular & Human Genetics Seminar, Houston, TX

## 2015

1. Cornell University Department of Biological Statistics and Computational Biology Seminar, Ithaca, NY
2. Festival of Genomics, San Francisco, CA
3. Human Genome Variation and Complex Genome Analysis (HGV 2015) Conference, San Francisco, CA
4. Member, Iconiq Capital Patients Like You Panel, San Francisco, CA
5. Invited lecture at San Francisco Mexican Museum, San Francisco, CA
6. Gordon Conference on Human Genetics and Genomics Speaker and Vice Chair, Newport, RI
7. Inaugural lecture on "Genomics and Health Disparities" at the National Institutes of Health, Bethesda, MD
8. Graduate Student invited Speaker, Department of Human Genetics, University of Chicago, Chicago, IL
9. Keynote speaker, UCLA/IPAM "Latinos in Mathematics" conference, Los Angeles, CA
10. Panelist at "Big Data, Genomics, and Precision Medicine," Tech Museum, San Jose, CA
11. University of Chicago Department of Genetics Seminar, Chicago, IL
12. Keynote speaker, "Human Genome Meeting 2015", Kuala Lumpur, Malaysia
13. Advances in Genome Biology and Technology (AGBT), Marco Island, FL
14. "Genomics and Sequencing Data Integration, Analysis and Visualization" Symposium, San Francisco, CA
15. CIENCIAS—a forum for presenting science in Spanish, Stanford University, Stanford, CA
16. Plant and Animal Genome Conference, San Diego, CA

## 2014

1. TEDMED 2014 inaugural cohost and invited speaker, San Francisco, CA
2. New York Genome Center Seminar, New York, NY
3. MacArthur Foundation Fellows Forum, Chicago, IL
4. Why We Can't Wait: Conference to Eliminate Health Disparities, Chair and Speaker, Washington D.C.
5. Human Genome Meeting on "Genome Variation and Human Health" (HUGO 2014), Geneva, Switzerland
6. Universidad Nacional del Altiplano, Puno, Peru
7. Illumina Users Meeting, South San Francisco, CA
8. "1000 Genomes and Beyond" Community Meeting, Cambridge, UK
9. Thermo Fisher Next Generation Sequencing seminar, South San Francisco, CA
10. University of Minnesota / Illumina "Understanding Your Genome", Minneapolis, MN
11. SAP Sapphire, Orlando, CA
12. Harvard Medical School Department of Genetics, Student-Invited Speaker, Boston, MA

13. Broad Institute of Harvard and MIT Medical and Population Genomics Seminar, Cambridge, MA
14. Advances in Genome Biology and Technology (AGBT) 2014, Marco Island, FL
15. University of California Riverside Department of Biology, Riverside, CA

### 2013

1. University of Miami Department of Human Genetics, Miami, FL
2. Willamette University Department of Biology, Portland, CA
3. Gordon Research Conference on Human Genetics and Genomics, RI
4. Arizona State University, MTBI program, Phoenix, AZ
5. Society for Mathematical Biology (SMB) Keynote Speaker 2013, Phoenix, AZ
6. 9<sup>th</sup> Molecular Biology of Hearing and Deafness Conference, Stanford, CA
7. 2013 Genome 10K Meeting Invited Speaker, Ft. Lauderdale, FL
8. Stanford/Oxford Big Data Conference, Stanford, CA

### 2012

1. Memorial Sloan Kettering Presidential Invited lecture
2. Oxford-Stanford Conference on Big Data, Oxford, UK
3. Illumina Users Meeting, American Society of Human Genetics 2012
4. Royal Society 2012 Galton and Fisher Lecture invited speaker, London, UK
5. 50<sup>th</sup> Anniversary of NIGMS DeWitt Stetten Jr. Lecture
6. UCSF Parnassus Seminar, San Francisco, CA
7. University of Helsinki Conference on Next-Generation Sequence Analysis
8. 2012 Green Mountain Conference on Forensic DNA Keynote Speaker
9. 2012 US/Indonesia Kavli Symposium (Solo, Indonesia)
10. University of Oxford Wellcome Trust Center for Human Genomics, Oxford, UK
11. Why We Can't Wait: 2012 Conference on Genomics and Health Disparities, Miami, FL
12. Boston College Landmarks in Biology Seminar, Boston, MA
13. North Carolina State Department of Genetics, Raleigh, NC
14. University of North Carolina Conference on Genetics of the Peoples of Africa and the Transatlantic African Diaspora, Chapel Hill, NC
15. University of California Davis, Davis, CA
16. University of California Berkeley, Berkeley, CA
17. Stanford University School of Law, Stanford, CA

### 2011

1. Institute for Pure and Applied Math Keynote Speaker, Los Angeles, CA
2. Society for Advancement of Chicanos and Native Americans in Science (SACNAS) 2011 conference, San Jose, CA
3. International Congress of Human Genetics 2011 Satellite Conference on Populations of the New World
4. University of California, Santa Cruz
5. University of Washington, Genome Sciences
6. 3<sup>rd</sup> EMBO Conference on Host Genetics and Infectious Disease, Pasteur Institute
7. Human Genome Variation (HGV) Meeting 2011
8. American Genetic Association Meeting 2011
9. European Society of Human Genetics 2011
10. University of Oregon (student invited speaker)
11. MacArthur Fellows Meeting
12. Stanford University Medical Grand Grounds

13. Slim Initiative in Genomic Medicine in the Americas
14. Commonwealth Club of California
15. Broad Institute of Harvard/MIT (Population and Medical Genetics)
16. Yale University School of Medicine (Genetics)
17. Harvard University (Organismic and Evolutionary Biology)
18. Florida International University (Biology and Human Genetics)
19. 2011 Genome 10K Conference
20. 2011 Gordon Conference on Quantitative Genomics
21. U. California Berkeley (Computational Biology)
22. Advances in Genome Biology and Technology (AGBT)

## 2010

1. U. California San Francisco (Human Genetics)
2. National Science Foundation Invited Speaker
3. LifeTech Users Meeting 2011
4. Cold Spring Harbor Banbury Conference on “DNA and Human History”
5. CARTA Symposium on Evolution of Human Biodiversity (UCSD)
6. Princeton University Lewis Sigler Institute
7. UCLA Human Genetics Retreat Invited Speaker
8. Johns Hopkins University School of Medicine (Genetics)
9. Genetic Society of America (GSA) 2010 Model Organism Conference invited speaker
10. European Molecular Biology Lab (EMBL) Workshop
11. Institute Pasteur, Paris, France (Immunology)
12. Duke University Computational Biology Graduate Student Invited Speaker
13. Advances in Genome Biology and Technology (AGBT) 2010)
14. University of Alabama ALLELE Invited Speaker
15. Graduate University for Advanced Studies (Hayama, Japan)

## 2009

1. Plant and Animal Genome XVI Meeting Aquaculture Keynote speaker
2. Office of AIDS Research Workshop on “Genomic Analyses of HIV Infection”
3. Conservation Genetics (ConGen) Workshop Smithsonian Tropical Research Institute
4. University of Pennsylvania (Genetics)
5. University of Pennsylvania (Veterinary School)
6. University of Miami Miller School of Medicine (Human Genetics)
7. Mexican National Institute of Genomic Medicine (INMEGEN)
8. Inter American Development Bank (IADB) 50th Annual Meeting (Medellin, Colombia)
9. Broad Institute of Harvard and MIT Medical and Population Genetics Seminar Series
10. University of Massachusetts at Amherst (Biology)
11. University of Miami Miller School of Medicine, HIV Grand Rounds
12. University of Washington St. Louis School of Medicine and UW Genome Center
13. RECOMB 2009 Symposium Speaker
14. Stanford University School of Medicine (Genetics)
15. 74th Cold Spring Harbor Symposia on Quantitative Biology
16. Pritzker School of Medicine University of Chicago (Human Genetics)
17. Society for Molecular Biology and Evolution 2009 Meeting (Iowa State University)
18. Gordon Research Conference on Human Genomics invited speaker
19. NGS2009 Conference on Next Generation Sequencing: Challenges and Opportunities (Barcelona, Spain)

20. American Society of Human Genetics (ASHG) 2009 Special Symposium Speaker
21. Columbia University Computational Biology Seminar Series
22. University of Indiana and University of Oregon NSF IGERT Invited Speaker
23. PABEW: Second Biological Evolution Workshop in (Porto Alegre, Brazil)
24. National Academy of Sciences Sackler Symposium "In the light of Evolution IV"

## 2008

1. Plant and Animal Genome XVI Meeting
2. University of California Los Angeles (History and Language Seminar Series)
3. University of Washington (Genome Sciences)
4. Gordon Conference on Molecular Evolution
5. Cornell University (Baker Institute for Animal Health)
6. University of Michigan (Biostatistics)
7. Harvard University (Anthropology)
8. Cornell University Plant Breeding and Genetics)
9. Stanford University (Biology)
10. Third International Conference on Primate Genomics
11. Cold Spring Harbor Biology of Genomes Meeting
12. 70th Anniversary Celebration for Centro Nacional de Investigacion en Cafe (CENICAFE) Bogota (Colombia)
13. Miami Institute of Human Genomics
14. HGM 2008: HUGO's 13th Human Genome Meeting in Hyderabad, India
15. Banbury Conference on Kinship, Race, and Ancestry
16. Cold Spring Harbor Symposium on Personalized Genomics
17. NHGRI Workshop on "Ethics and the Study of Natural Selection"
18. American Society of Human Genetics (ASHG) 2008 Invited Speaker
19. Stanford University (Biology)
20. 23andMe
21. Cornell-Weill School of Medicine (Pediatric Grand Rounds)

## 2007

1. Recomb Satellite Workshop
2. University of Arizona Ecology and Evolutionary Biology
3. American Institute of Biological Sciences Annual Meeting
4. NIH Workshop on Macaque Genetics
5. Society for Molecular Biology and Evolution 2007 Annual Conference
6. Banbury Conference on Domestication Genomics
7. National Association of Biology Teachers (NABT)
8. Konrad Lorenz Institute Workshop on Selective Sweeps

## 2006

1. Institute for Pure and Applied Math Short Program
2. University of Chicago
3. Universitat Pompeu Fabra
4. Arizona State University
5. Society for Molecular Biology and Evolution 2006 Annual Conference)
6. SNP Meetings 2006
7. American Society of Human Genetics ASHG 2006

## 8. Cornell Center for Vertebrate Genomics

### 2005

1. Eastern Great Lakes Molecular Evolution Meeting invited speaker
2. Mathematical Biosciences Institute Conference on Evolutionary Genomics
3. Hunter College, invited speaker

### 2004

1. Max Planck Institute for Evolutionary Anthropology
2. Cornell University (Field of Genetics)
3. North Carolina State (Bioinformatics)
4. Cornell University (Plant Biology)
5. University of Southern California (Computational Biology)
6. University of California Los Angeles (Human Genetics)
7. NSF Institute for Pure and Applied Mathematics
8. ETH Zurich (Student Invited Speaker)
9. University of Bergen
10. Blackwell-Tapia Conference
11. University of Maryland (Ecology & Evolutionary Biology)

### 2001-2003

1. University California Los Angeles (Biology)
2. University California Los Angeles (Human Genetics)
3. University of Arizona (Biology)
4. California Institute of Technology (Biology)
5. University North Carolina Chapel Hill (Genetics)
6. Brown University (Biology)
7. Cornell University (Biometry)
8. University of California Berkeley