
BIOGRAPHICAL SKETCH

NAME Bustamante, Carlos D., Ph.D.	POSITION TITLE Professor of Genetics and (by courtesy) Biology Founding Director, Stanford Center for Computational, Evolutionary, and Human Genomics
eRA COMMONS USER NAME (credential, e.g., agency login) CDBUSTAMANTE	

EDUCATION/TRAINING			
INSTITUTION AND LOCATION	DEGREE (if applicable)	MM/YY	FIELD OF STUDY
Harvard University, Cambridge, MA	B.A.	1997	Biology, <i>magna cum laude</i>
Harvard University, Cambridge, MA	M.S.	2001	Statistics
Harvard University, Cambridge, MA	Ph.D.	2001	Biology
University of Oxford, Oxford, UK	Post-doc	2001-2002	Mathematical Genetics

Contact information

Department of Genetics
Stanford School of Medicine
300 Pasteur Drive
Lane Building, Room L-301
Stanford, CA 94305-5120

Assistant: +1-650-723-1074
Direct: +1-650-723-6330
Fax: +1-650-723-3667
Email: cdbustam@stanford.edu

Personal Statement

I am a population geneticist whose research focuses on analyzing genome wide patterns of variation within and between species to address fundamental questions in biology, anthropology, and medicine. As a faculty member at Cornell and Stanford, I have trained 43 post-doctoral fellows and graduate students as primary advisory. My group works on a variety of organisms and model systems ranging from humans and other primates to domesticated plant and animals. Much of our research is at the interface of computational biology, mathematical genetics, and genomics. Examples of our research accomplishments:

- Developing selection maps of the human genome which pinpoint rapidly evolving genes as well as genomic regions subject to strong selective constraint;
- Developing a high-density map of genetic variation in the dog genome and using it to identify genomic regions underlying morphological differences among domestic dog breeds;
- Investigating the fine scale genetic structure of human populations and its implication for genomic medicine.

Our current research focuses on human population genomics and global health including developing statistical, computational, and genomic resources for enabling trans- and multi-ethnic genome-wide association and medical sequencing studies of Mendelian, oligogenic, and complex biomedical traits. My lab has both dry and wet lab expertise and students and post-docs from diverse fields including: applied math, biostatistics/statistics, computational biology/bioinformatics, anthropology/evolutionary biology, medicine, and genetics/genomics.

I am also Founding Director (with Marc Feldman) of the Stanford Center for Computational, Evolutionary, and Human Genomics and former Associate Director of the Cornell Center for Comparative and Population Genomics (3CPG). I am also Director of Informatics at the Stanford Center for Genomics and Personalized Medicine. 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, and veterinary and plant genomics.

Positions and Employment

2012-present Founding Director, Stanford Center for Computational, Evolutionary, and Human Genomics
2012-present Director of Informatics, Stanford Center for Genomics and Personalized Medicine
2010-present Professor, Department of Genetics and by, courtesy, Biology, Stanford University
2007-8 Associate Director, Cornell Center for Comparative and Population Genomics
2008 Professor, Department of Biological Statistics and Computational Biology and Department of Statistical Sciences, Cornell University
2008 Associate Professor, Department of Biological Statistics and Computational Biology and Department of Statistical Sciences, Cornell University
2004 Visiting Scholar, NSF Institute for Pure and Applied Mathematics, Los Angeles
2002-2008 Assistant Professor, Department of Biological Statistics and Computational Biology and Department of Statistical Sciences, Cornell University

Fellowships and Honors

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

Other Experience and Professional Memberships

2012-present Co-Chair, 2013 Conference on Genomics and Health Disparities (w/ Margaret Pericak-Vance)
2012-present Organizer, 2014 Meeting of the Society for Molecular Biology and Evolution
2012-present Organizer, Cold Spring Harbor Biology of Genomes Conference
2010-present Editorial Board, *Human Biology*
2009-present Member, The Center for Academic Research and Training in Anthropogeny (UCSD)
2008-present Editorial Board, *Genome Research*
2008-2012 Editorial Board, *Molecular Biology and Evolution*
2007-2009 Sloan Research Fellow in Computational and Evolutionary Molecular Biology
2005-2009 Editorial Board, *Bioinformatics*
2004-2005 Fellow of NSF Institute for Pure and Applied Mathematics
2001-2002 Marshall-Sherfield Fellowship
1998-2001 Howard Hughes Medical Institute Graduate Fellowship

Non-profit and Government Service

2013-present National Institutes of Health Council of Councils (pending)
2011-present National Human Genome Research Institute Council
2011-present External Evaluation committee NIDDK T2D GENES project
2010-present 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
2002-2012 *Ad hoc* reviewer for NIH and NSF study sections (~1-2 per year)

Consulting and Industry Service

2012-present Scientific Advisory Board, Personalis, Inc. (Menlo Park, CA)
2012-present Medical Advisory Board, Med-Tek (Miami, FL)
2012-present Scientific Advisory Board, InVitae (formerly Locus Development, Inc.) (San Francisco, CA)
2011-present Scientific Advisory Board, 23andme "Roots into the Future" (Mountain View, CA)
2011-present Scientific Advisory Board, Ancestry.com (Provo, UT)
2009-2012 Scientific Advisory Board, Nature Source Genetics (Ithaca, NY)
2012-present Consultant in Statistical Genomics, Mars, Inc. (McClean, VA)
2011-2012 Scientific Advisory Board, Mubadala Healthcare (Abu Dhabi, UAE)
2007-2010 Consultant in Statistical Genomics, Mars Symbioscience (McClean, VA)

2008 Consultant in Statistical Genomics, Knome, Inc. (Cambridge, MA)
2008 Consultant in Statistical Genomics, Marshall BioResources (North Rose, NY)

Publications (119 total)

Journal Article (116)

2012

116. 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.
115. Cornejo OE, Lefebure 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. (in press). *Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans**. **Mol Biol Evol**. PMID: 23228887.
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**. (in press) 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. PMID: 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. PMID: 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. PMID: 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**. *SnapShot: Human biomedical genomics*. **Cell.** 2011;147(1):248- e1.
96. **Bustamante CD**, Burchard EG, De la Vega FM. *Genomics for the world*. **Nature.** 2011;475(7355):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. *Phased whole-genome genetic risk in a family quartet using a major allele reference sequence*. **PLoS Genet.** 2011;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*. **Nature**. 2010;468(7327):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, Kretzschmar 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**, Korbelt 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, Elkahoun AG, Castelhamo 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, Elkhouloun 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, Elkhouloun 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, Elkhouloun 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, Bruskiewich 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. Blekhan 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.

Pre-2006

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)2D3 protects from Cytoxan-induced alopecia without protecting the leukemic cells from Cytoxan*. **Am J Med Sci**.

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

2R01HG003229 (Clark, Bustamante, Nielsen, PIs) 9/18/2008–6/30/2014 DHHS/NIH/NHGRI
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.

DMS 1201234 (Bustamante, PI) 09/01/2012--08/31/2016 NSF/DMS
Statistical Methods for Enabling Medical and Population Genomics of Admixed Human Populations
The major goal of this project is development of population genetic theory and statistical genomic approaches for modeling the demographic history of admixed human populations, including minority populations in the U.S.

1R01GM090087 (Bustamante, Martin, PIs) 01/01/2012–12/31/2016 DHHS/NIH/NIGMS
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.

58-6631-0-100 (Bustamante, PI) 07/1/2010–06/30/2013 USDA/ARS
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.

U01HG005715 (Bustamante, PI) 09/1/2010–06/30/2013 (NCX) DHHS/NIH/NHGRI
Population structure, admixture and selection across the 1000 genome 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.

COMPLETED RESEARCH (including grants with No-Cost Extensions expiring by 9/1/2012)

1R01GM083606 (Bustamante, PI) 08/01/2007–07/31/2012 (NCE) NIH/NIGMS/CBCB
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) 09/28/2008–07/31/2012 (NCE) DHHS/NIH/NIMH
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) 9/1/2007–8/31/2012 (NCE) NSF/DBI
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 of within and between species genomic variation data.

5U01HL084706 -03 (Clark, PI, Bustamante, Co-PI) 06/15/2006 – 05/31/2010 NIH/NHLBI
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) 08/15/2009–7/31/2011 (NCE) NSF/DBI/EAGER
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) 10/1/2009-4/30/2012 (NCE) NHLBI
Northwest Genomics 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 as well as estimating 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), Assistant Professor, Albert Einstein College of Medicine
2. Dr. Adam Boyko, Ph. D. (9/2005—8/2011), Assistant Professor, Cornell University
3. Dr. Jake Byrnes, Ph. D. (9/2010—10/2011), Computational Biologists, Ancestry.com
4. Dr. Meredith Carpenter, Ph. D. (7/2011—current)
5. Dr. Omar Cornejo, Ph. D. (8/2009—current)
6. Dr. Matthew J. Dimmic, Ph. D. (3/2003—2005), Computational Biology Group Lead at Monsanto
7. Dr. Ryan Gutenkunst, Ph. D. (1/2008--12/2008), Assistant Professor, U. of Arizona
8. Dr. Simon Gravel, Ph.D. (10/2009—current) soon to be Assistant Professor, McGill University
9. Dr. Brenna Henn, Ph. D. (12/2009—current) soon to be Assistant Professor, SUNY Stony Brook

10. Dr. Eimear Kenny, Ph. D. (10/2010—10/2012), Assistant Professor, Mount Sinai School of Medicine
11. Dr. Joanna Kelley, Ph. D. (8/2010—current)
12. Dr. Jeffrey Kidd, Ph. D. (2/2010—12/2011), Assistant Professor, University of Michigan School of Medicine
13. Dr. Carolin Kosiol, Ph. D. (9/2006—1/2009), Group Leader at the University of Veterinary Medicine (vetmeduni) Vienna, Austria
14. Dr. Andres Moreno, M.D./Ph. D. (3/2009—current)
15. Dr. Marina Muzzio, Ph. D. (8/2011—current)
16. Dr. Badri Padhukasahasram, Ph. D. (9/2006—1/2008), Investigator at Henry Ford Health System (Detroit, Michigan)
17. Dr. Arindam RoyChoudhury, Ph. D. (7/2008—8/2009), Assistant Professor of Biostatistics at Columbia University
18. Dr. Suyash Shringarpure, Ph.D. (9/2012—current)
19. Dr. Sean Myles, Ph.D. (1/2008—7/2010), Assistant Professor, Dalhousie University
20. Dr. Karla Sandoval Mendoza, Ph.D. (4/10—current)
21. Dr. Martin Sikora, Ph. D. (7/10—current)
22. Dr. Scott Williamson, Ph. D. (6/2003—6/2006), deceased (was Assistant Professor at Cornell University)
23. Dr. Keyan Zhao, Ph.D. (9/2007—12/2010), Post-doc Scholar in Dept. of Internal Medicine at University of Iowa

Graduate Students Trained

1. Abra Brisbin (Ph. D. program in Applied Math) (2006 – 2010), joint with J. Mezey), Assistant Professor of Mathematics, University of Wisconsin-Eau Claire
2. Kasia Bryc (Ph. D. program in Biometry) (2006 — 2010), Post-doctoral fellow w/ David Reich at Harvard
3. Thomas Cooke (Ph. D. program in Genetics) (2011—current)
4. Helio Costa (Ph. D. program in Genetics) (2011—current)
5. Jeremiah Degenhardt (Ph. D. program in Biometry) (2005 — 2010), Senior Scientist Bioinformatics at Onyx Pharmaceuticals
6. Kirsten Eilertson (Ph. D. program in Statistics) (2007—2011), Research Biostatistician, Gladstone Institute of UCSF
7. Hong Gao (Ph. D. program in Genetics) (2004—2008), Bioinformatics Scientist at Affymetrix
8. Ryan Hernandez (Ph. D. program in Biometry) (2003—2008), Assistant Professor, UCSF
9. Sandra Indarraga (M.S./Ph. D. program in Genetics) (2009—2011)
10. Emilia Huarte-Sanchez (Ph. D. program in Applied Math, joint with R. Durrett) (2005—2009), Post-doctoral fellow at UC Berkeley
11. Jeffrey Jensen (Ph. D. program in Genetics & Development, joint with C.F. Aquadro) (2002—2006), Head of Laboratory and Professor, Ecole Polytechnique Federale de Laussane
12. Lin Lin (Ph. D. program in Computational Biology) (2006—2010), Post-doc at Harvard University
13. Kirk Lohmueller (Ph. D. program in Genetics & Development) (2005 – 2009, joint. with A. G. Clark), Assistant Professor, UCLA
14. Alicia Martin (Ph.D. program in Genetics) (2011—current)
15. Brian Maples (Ph.D. program in Biomedical Informatics) (2011—current)
16. Shaila Musharoff (Ph. D. program in Genetics) (2009—current)
17. Xin (Maria) Ma (Ph. D. program in Statistics) (2008—2011), Post-doc at Stanford University
18. Patricia Ortiz (M.D./Ph.D. program in Genetics) (2010—current)
19. David Poznik (Ph.D. program in Biomedical Informatics) (2011—current)
20. Mark Wright (Ph. D. program in Genetics & Development) (2007—2010), Post-doc at Cornell University
21. Fouad Zakharia (Ph. D. program in Genetics) (2010—2012), Analyst at Citibank
22. Lan Zhu (M.S. /Ph. D. program in Biometry) (2002—2006), Assistant Professor, Oklahoma State University

Courses Taught

- CSRE 110: Comparative Studies of Race and Ethnicity (W2011, guest lecture)
 GENE 210: Personalized Genomics (S2010, S2011, guest lecture)
 GENE 203: Graduate Genetics (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)
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. to Bioinformatics and Statistical Genomics (F2004 w/ R. Nielsen, F2005 solo)
BTRY 652: Computationally Intensive Statistical Inference (S2003)

Invited Conferences, Keynote Lectures, and Departmental Seminars (135 total)

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. 50th Anniversary of NIGMS DeWitt Stetten Jr. Lecture
6. UCSF
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 Welcomme 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

18. Institute for Pure and Applied Math Keynote Speaker, Los Angeles, CA
19. Society for Advancement of Chicanos and Native Americans in Science (SACNAS) 2011 conference, San Jose, CA
20. International Congress of Human Genetics 2011 Satellite Conference on Populations of the New World
21. University of California, Santa Cruz
22. University of Washington, Genome Sciences
23. 3rd EMBO Conference on Host Genetics and Infectious Disease, Pasteur Institute
24. Human Genome Variation (HGV) Meeting 2011
25. American Genetic Association Meeting 2011
26. European Society of Human Genetics 2011
27. University of Oregon (student invited speaker)
28. MacArthur Fellows Meeting
29. Stanford University Medical Grand Grounds
30. Slim Initiative in Genomic Medicine in the Americas
31. Commonwealth Club of California
32. Broad Institute of Harvard/MIT (Population and Medical Genetics)
33. Yale University School of Medicine (Genetics)
34. Harvard University (Organismic and Evolutionary Biology)
35. Florida International University (Biology and Human Genetics)
36. 2011 Genome 10K Conference
37. 2011 Gordon Conference on Quantitative Genomics
38. U. California Berkeley (Computational Biology)
39. Advances in Genome Biology and Technology (AGBT)

2010

40. U. California San Francisco (Human Genetics)
41. National Science Foundation Invited Speaker
42. LifeTech Users Meeting 2011
43. Cold Spring Harbor Banbury Conference on "DNA and Human History"
44. CARTA Symposium on Evolution of Human Biodiversity (UCSD)
45. Princeton University Lewis Sigler Institute
46. UCLA Human Genetics Retreat Invited Speaker
47. Johns Hopkins University School of Medicine (Genetics)
48. Genetic Society of America (GSA) 2010 Model Organism Conference invited speaker
49. European Molecular Biology Lab (EMBL) Workshop
50. Institute Pasteur, Paris, France (Immunology)
51. Duke University Computational Biology Graduate Student Invited Speaker
52. Advances in Genome Biology and Technology (AGBT) 2010)
53. University of Alabama ALLELE Invited Speaker
54. Graduate University for Advanced Studies (Hayama, Japan)

2009

55. Plant and Animal Genome XVI Meeting Aquaculture Keynote speaker
56. Office of AIDS Research Workshop on "Genomic Analyses of HIV Infection"
57. Conservation Genetics (ConGen) Workshop Smithsonian Tropical Research Institute
58. University of Pennsylvania (Genetics)
59. University of Pennsylvania (Veterinary School)
60. University of Miami Miller School of Medicine (Human Genetics)
61. Mexican National Institute of Genomic Medicine (INMEGEN)
62. Inter American Development Bank (IADB) 50th Annual Meeting (Medellin, Colombia)
63. Broad Institute of Harvard and MIT Medical and Population Genetics Seminar Series
64. University of Massachusetts at Amherst (Biology)
65. University of Miami Miller School of Medicine, HIV Grand Rounds
66. University of Washington St. Louis School of Medicine and UW Genome Center
67. RECOMB 2009 Symposium Speaker
68. Stanford University School of Medicine (Genetics)
69. 74th Cold Spring Harbor Symposia on Quantitative Biology
70. Pritzker School of Medicine University of Chicago (Human Genetics)
71. Society for Molecular Biology and Evolution 2009 Meeting (Iowa State University)
72. Gordon Research Conference on Human Genomics invited speaker
73. NGS2009 Conference on Next Generation Sequencing: Challenges and Opportunities (Barcelona, Spain)
74. American Society of Human Genetics (ASHG) 2009 Special Symposium Speaker
75. Columbia University Computational Biology Seminar Series
76. University of Indiana and University of Oregon NSF IGERT Invited Speaker
77. PABEW: Second Biological Evolution Workshop in (Porto Alegre, Brazil)
78. National Academy of Sciences Sackler Symposium "In the light of Evolution IV"

2008

79. Plant and Animal Genome XVI Meeting
80. University of California Los Angeles (History and Language Seminar Series)
81. University of Washington (Genome Sciences)
82. Gordon Conference on Molecular Evolution
83. Cornell University (Baker Institute for Animal Health)
84. University of Michigan (Biostatistics)
85. Harvard University (Anthropology)
86. Cornell University Plant Breeding and Genetics)
87. Stanford University (Biology)

88. Third International Conference on Primate Genomics
89. Cold Spring Harbor Biology of Genomes Meeting
90. 70th Anniversary Celebration for Centro Nacional de Investigacion en Cafe (CENICAFE) Bogota (Colombia)
91. Miami Institute of Human Genomics
92. HGM 2008: HUGO's 13th Human Genome Meeting in Hyderabad, India
93. Banbury Conference on Kinship, Race, and Ancestry
94. Cold Spring Harbor Symposium on Personalized Genomics
95. NHGRI Workshop on "Ethics and the Study of Natural Selection"
96. American Society of Human Genetics (ASHG) 2008 Invited Speaker
97. Stanford University (Biology)
98. 23andMe
99. Cornell-Weill School of Medicine (Pediatric Grand Rounds)

2007

100. Recomb Satellite Workshop
101. University of Arizona Ecology and Evolutionary Biology
102. American Institute of Biological Sciences Annual Meeting
103. NIH Workshop on Macaque Genetics
104. Society for Molecular Biology and Evolution 2007 Annual Conference
105. Banbury Conference on Domestication Genomics
106. National Association of Biology Teachers (NABT)
107. Konrad Lorenz Institute Workshop on Selective Sweeps

2006

108. Institute for Pure and Applied Math Short Program
109. University of Chicago
110. Universitat Pompeu Fabra
111. Arizona State University
112. Society for Molecular Biology and Evolution 2006 Annual Conference)
113. SNP Meetings 2006
114. American Society of Human Genetics ASHG 2006
115. Cornell Center for Vertebrate Genomics

2005

116. Eastern Great Lakes Molecular Evolution Meeting invited speaker
117. Mathematical Biosciences Institute Conference on Evolutionary Genomics
118. Hunter College

2004

119. Max Planck Institute for Evolutionary Anthropology
120. Cornell University (Field of Genetics)
121. North Carolina State (Bioinformatics)
122. Cornell University (Plant Biology)
123. University of Southern California (Computational Biology)
124. University of California Los Angeles (Human Genetics)
125. NSF Institute for Pure and Applied Mathematics
126. ETH Zurich (Student Invited Speaker)
127. University of Bergen
128. Blackwell-Tapia Conference
129. University of Maryland (Ecology & Evolutionary Biology)

2001-2003

130. University California Los Angeles (Biology)

131. University California Los Angeles (Human Genetics)
132. University of Arizona (Biology)
133. California Institute of Technology (Biology)
134. University North Carolina Chapel Hill (Genetics)
135. Brown University (Biology)
136. Cornell University (Biometry)
137. University of California Berkeley