Adam C. Siepel, Ph.D.

PERSONAL DETAILS

Present position:	Chair, Simons Center for Quantitative Biology Professor, Watson School of Biological Sciences Cold Spring Harbor Laboratory
Address:	Cold Spring Harbor Laboratory 2121 Koch Building One Bungtown Road Cold Spring Harbor, NY 11724–2209 (516) 367 6922 (516) 367 5501-Fax asiepel@cshl.edu http://SiepelLab.labsites.cshl.edu http://www.cshl.edu/scqb

EDUCATION AND PROFESSIONAL QUALIFICATIONS

Education

Cornell University, Ithaca, New York, B.S., Agricultural and Biological Engineering, 1994

University of New Mexico, Albuquerque, New Mexico, M.S., Computer Science, 2001

University of California, Santa Cruz, California, Ph.D., Computer Science, 2005

Professional Experience

2015 – present	Adjunct Professor, Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY
2014 – present	Professor, Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
2014 – present	Chair, Simons Center for Quantitative Biology, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
2014 – present	Adjunct Associate Professor, Department of Computer Science, Stony Brook University, Stony Brook, NY
2014 – present	Adjunct Professor of Computational Genomics, Institute for Computational Biomedicine, Weill Cornell Medical College of Cornell University, New York, NY
2010 – 2014	Associate Professor (with tenure), Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY
2012 – 2013	Visiting Scientist (during sabbatic leave from Cornell), European Bioinformatics Institute (EBI), Hinxton, UK
2012 – 2013	Visiting By-Fellow in Residence (during sabbatic leave from Cornell), Churchill College, Cambridge University, Cambridge, UK
2006 – 2010	Assistant Professor, Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY
2002 – 2005	Graduate Student Researcher, University of California, Santa Cruz, CA

2000 – 2001 Director of Information Technology (acting), National Center for Genome Resources (NCGR), Santa Fe, NM

1997 – 2001 Software Development Group Leader, National Center for Genome Resources (NCGR), Santa Fe, NM

1996 – 1997 Bioinformatics Programmer, National Center for Genome Resources (NCGR), Santa Fe, NM

1994 – 1996 Graduate Research Assistant, HIV Database Group, Los Alamos National Laboratory (LANL), Los Alamos, NM

Honors and Awards

Distinguished Graduate Student Alumnus Award, Baskin School of Engineering, University of Santa Cruz, California, 2017.

Guggenheim Fellowship in Molecular and Cellular Biology, John Simon Guggenheim Memorial Foundation, 2012–2013.

Sloan Research Fellowship in Computational & Evolutionary Molecular Biology, Alfred P. Sloan Foundation, 2009–2011.

Packard Fellowship for Science and Engineering, David & Lucile Packard Foundation, 2007.

Microsoft Research Faculty Fellowship Program, 2007.

National Science Foundation (NSF) CAREER Award, 2007.

Graduate Research and Education in Adaptive Bio-technology (GREAT) Training Program Fellowship, University of California System-wide Biotechnology Research and Education Program (UC BREP), 2004–2005.

Achievement Rewards for College Scientists (ARCS) Scholar Award, Sponsored by ARCS Northern California Chapter, ARCS Foundation, Inc., 2003–2004.

Chancellor’s Fellowship, University of California, Santa Cruz, 2002.

Editing and Review

NIH Special Emphasis Panel: Maximizing Investigators’ Research Award (MIRA) Review, National Institute of General Medical Sciences (NIGMS), Bethesda, MD, 2016.

Peer Reviewer: 24th FNP Prize, Foundation for Polish Science, Warsaw, Poland, 2015.

NIH Special Emphasis Panel: Genomic Resources Review, National Human Genome Research Institute (NHGRI), 2015.

NIH Study Section: Genomics, Computational Biology, and Technology (GCAT), 2010, 2014.

Editorial Board: *PLOS Computational Biology*, 2008–2012.

Program Committee: Research in Computational Molecular Biology (RECOMB), 2008–2010.

Editorial Board: *Genome Research*, 2007–2010.

Program Committee: Workshop on Algorithms in Bioinformatics (WABI), 2006–2008.

NSF Panelist: Directorate for Biological Sciences (BIO), 2007.

Program Committee: International Conference on Machine Learning (ICML), 2007.

Program Committee: Intelligent Systems for Molecular Biology (ISMB) / European Conference on Computational Biology (ECCB), 2004–2007.

PUBLICATIONS

Journal Articles

1. Huang Y-F, **Siepel A**. Estimation of allele-specific fitness effects across human protein-coding sequences and implications for disease. Preprint *bioRxiv* 441337. doi: https://doi.org/10.1101/441337, 2018.
2. Gulko B, **Siepel A**. An evolutionary framework for measuring epigenomic information and estimating cell-type-specific fitness consequences. In Press. *Nat. Gen.*, 2018.
3. Ramani R, Krumholz K, Huang YF, **Siepel A**. phastWeb: a web interface for evolutionary conservation scoring of multiple sequence alignments using phastCons and phyloP. Epub ahead of print. *Bioinformatics*. doi: https://doi.org/10.1093/bioinformatics/bty966, 2018.
4. Danko CG, Choate LA, Marks BA, Rice EJ, Zhong W, Chu T, Martins AL, Dukler N, Coonrod SA, Wojno EDT, Lis JT, Kraus WL, **Siepel A**. Dynamic evolution of regulatory element ensembles in primate CD4+ T-cells. *Nat. Ecol. Evol.* **2**(3): 537-548, 2018.
5. Fang H, Huang YF, Radhakrishnan A, **Siepel A**, Lyon GJ, Schatz MC. Scikit-ribo enables accurate estimation and robust modeling of translation dynamics at codon resolution. *Cell Syst.* **6**(2): 180-191.e4, 2018.
6. Mohammed J, Flynt AS, Panzarino AM, Mondal MMH, DeCruz M, **Siepel A**, Lai EC. Deep experimental profiling of microRNA diversity, deployment, and evolution across the *Drosophila* genus. *Genome Res.* **28**(1): 52-65, 2018.
7. Dukler N, Booth GT, Huang Y-F, Tippens N, Waters CT, Danko CG, Lis JT, and **Siepel A**. Nascent RNA sequencing reveals a dynamic global transcriptional response at genes and enhancers to the natural medicinal compound celastrol. *Genome Res.* **27**: 1816-1829, 2017.
8. Kondo S, Vedanayagam J, Mohammed J, Eizadshenass S, Pang N, Kan L, Aradhya R, **Siepel A**, Steinhauer J, Lai EC. New genes often acquire male-specific functions but rarely become essential in *Drosophila*. *Genes. Dev*. **31**(18): 1841-1846, 2017.
9. Huang YF, Gulko B, **Siepel A**. Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. *Nat. Genet.* **49**(4): 618–624, 2017.
10. Freedman AH, Schweizer RM, Ortega-Del Vecchyo D, Han E, Davis BW, Gronau I, **Siepel A**, et al. (30 co-authors). Demographically-based evaluation of genomic regions under selection in domestic dogs. *PLOS Genet.* **12**(3): e1005851, 2016.
11. Kuhlwilm M, Gronau I, Hubisz MJ, de Filippo C, Prado J, Kircher M, Fu Q, Burbano HA, Lalueza-Fox C, de la Rasilla M, Rosas A, Rudan P, Brajkovic D, Kucan Z, Gusic I, Marques-Bonet T, Andres AM, Viola B, Paabo S, Meyer M, **Siepel A**, and Castellano S. Ancient gene flow from early modern humans into eastern neanderthals. *Nature*. **530**(7591): 429–433, 2016.
12. Campagna L, Gronau I, Silveira LF, **Siepel A**, Lovette IJ. Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Mol. Ecol.* **24**(16): 4238–4251, 2015.
13. Danko CG, Hyland SL, Core LJ, Martins AL, Waters CT, Lee HW, Cheung VG, Kraus WL, Lis JT, **Siepel A**. Identification of active transcriptional regulatory elements from GRO-seq data. *Nat. Methods.* **12**(5): 433–438, 2015.

14. Fuda NJ, Guertin MJ, Sharma S, Danko CG, Martins AL, **Siepel A**, Lis, JT. GAGA factor maintains nucleosome-free regions and has a role in RNA polymerase II recruitment to promoters. *PLOS Genet.* **11**(3): e1005108, 2015.
15. Gulko B, Hubisz MJ, Gronau I, **Siepel A**. A method for calculating probabilities of fitness consequences for point mutations across the human genome. *Nat. Genet.* **47**(3): 276–283, 2015.
16. Mohammed J, **Siepel A**, Lai EC. Diverse modes of evolutionary emergence and flux of conserved microRNA clusters. *RNA.* **20**(12): 1850–1863, 2014.
17. Core LJ, Martins AL, Danko CG, Waters CT, **Siepel A**, Lis JT. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* **46**(12): 1311–1320, 2014.
18. Peng X, Alföldi J, Gori K, Eisfeld AJ, Tyler SR, …, **Siepel A**, et al. (38 co-authors). The draft genome sequence of the ferret (*Mustela putorius furo*) facilitates study of human respiratory disease. *Nat. Biotechnol.* **32**(12): 1250–1255, 2014.
19. Couger MB, Pipes L, Squina F, Prade R, **Siepel A**, Palermo R, Katze MG, Mason CE, Blood PD. Enabling large-scale next-generation sequence assembly with Blacklight. *Concurr. Comput.* **26**(13): 2157–2166, 2014.
20. Mohammed J, Bortolamiol-Becet D, Flynt AS, Gronau I, **Siepel A**, Lai EC. Adaptive evolution of testis-specific, recently evolved, clustered miRNAs in *Drosophila*. *RNA.* **20**(8): 1195–1209, 2014.
21. Wen J, Mohammed J, Bortolamiol-Becet D, Tsai H, Robine N, Westholm JO, Ladewig E, Dai Q, Okamura K, Flynt AS, Zhang D, Andrews J, Cherbas L, Kaufman TC, Cherbas P, **Siepel A**, Lai EC. Diversity of miRNAs, siRNAs, and piRNAs across 25 *Drosophila* cell lines. *Genome Res.* **24**(7): 1236–1250, 2014.
22. Arbiza L, Gottipati S, **Siepel A**, Keinan A. Contrasting X-linked and autosomal diversity across 14 human populations. *Am. J. Hum. Genet.* **94**(6): 827–844, 2014.
23. Freedman AH, Gronau I, Schweizer RM, Ortega-Del Vecchyo D, Han E, …, **Siepel A**, et al. (30 co-authors). Genome sequencing highlights the dynamic early history of dogs. *PLOS Genet.* **10**(1): e1004016, 2014.
24. Rasmussen MD, Hubisz MJ, Gronau I, **Siepel A**. Genome-wide inference of ancestral recombination graphs. *PLOS Genet.* **10**(5): e1004342, 2014.
25. Mohammed J, Flynt AS, **Siepel A**, Lai EC. The impact of age, biogenesis, and genomic clustering on *Drosophila* microRNA evolution. *RNA.* **19**(9): 1295–1308, 2013.
26. Arbiza L, Gronau I, Aksoy BA, Hubisz MJ, Gulko B, Keinan A, **Siepel A**. Genome-wide inference of natural selection on human transcription factor binding sites. *Nat. Genet.* **45**(7): 723–729, 2013.
27. Gronau I, Arbiza L, Mohammed J, **Siepel A.** Inference of natural selection from interspersed genomic elements based on polymorphism and divergence. *Mol. Biol. Evol.* **30**(5): 1159–1171, 2013.
28. Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, **Siepel A**, Kraus WL. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol. Cell* **50**(2): 212–222, 2013.
29. Cornejo OE, Lefbure 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. Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans*. *Mol. Biol. Evol.* **30**(4): 881–893, 2013.
30. Zeng L, Choi SC, Danko CG, **Siepel A**, Stanhope MJ, Burne RA. Gene regulation by CcpA and catabolite repression explored by RNA-Seq in *Streptococcus mutans*. *PLOS One* **8**(3): e60465, 2013.

31. Ma X, Kelley JL, Eilertson K, Musharoff S, Degenhardt JD, Martins AL, Vinar T, Kosiol C, **Siepel A**, Gutenkunst RN, Bustamante CD. Population genomic analysis reveals a rich speciation and demographic history of orangutans (*Pongo pygmaeus* and *Pongo abelii*). *PLOS One* **8**(10): e77175, 2013.
32. Capra JA, Hubisz MJ, Kostka D, Pollard KS, **Siepel A**. A model-based analysis of GC-biased gene conversion in the human and chimpanzee genomes. *PLOS Genet.* **9**(8): e1003684, 2013.
33. Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, **Siepel A**. Replacing and additive horizontal gene transfer in *Streptococcus*. *Mol. Biol. Evol.* **29**(11): 3309–3320, 2012.
34. Kostka DA, Hubisz MJ, **Siepel A**, Pollard KS. The role of GC-biased gene conversion in shaping the fastest evolving regions of the human genome. *Mol. Biol. Evol.* **29**(3): 1047–1057, 2012.
35. Guertin MJ, Martins AL, **Siepel A**, Lis JT. Accurate prediction of inducible transcription factor binding intensities in vivo. *PLOS Genet.* **8**(3): e1002610, 2012.
36. Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, …, **Siepel A**, et al. (86 co-authors). A high-resolution map of human evolutionary constraint using 29 mammals. *Nature.* **478**(7370): 476–482, 2011.
37. Gronau I, Hubisz MJ, Gulko B, Danko CG, **Siepel A**. Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**(10): 1031–1034, 2011.
38. Lowe CB, Kellis M, **Siepel A**, Raney B, Clamp M, Salama SR, Kingsley D, Lindblad-Toh K, Haussler D. Three periods of regulatory innovation during vertebrate evolution. *Science* **333**(6045): 1019–1024, 2011.
39. Gottipati S, Arbiza L, **Siepel A**, Clark AG, Keinan A. Analyses of X-linked and autosomal genetic variation in population-scale whole genome sequencing. *Nat. Genet.* **43**(8): 741–743, 2011.
40. Hah N, Danko CG, Core L, **Siepel A**, Lis JT, Kraus WL. A Rapid, Extensive, and Transient Transcriptional Response to Estrogen Signaling in Breast Cancer Cells. *Cell* **145**(4): 622–634, 2011.
41. Hubisz MJ, Lin MF, Kellis M, **Siepel A**. Error and error mitigation in low-coverage genome assemblies. *PLOS One* **6**(2): e17034, 2011.
42. Yang J-S, Phillips MD, Betel D, Mu P, Sander C, Ventura A, **Siepel AC**, Chen KC, Lai EC. Widespread regulatory activity of vertebrate microRNA* species. *RNA* **17**(2): 312–326, 2011.
43. Orangutan Genome Sequencing and Analysis Consortium. Comparative and demographic analysis of orang-utan genomes. *Nature* **469**(7331): 529–533, 2011.
44. Suzuki H, Lefebure T, Hubisz MJ, Bitar PP, Lang P, **Siepel A**, Stanhope MJ. Comparative genomic analysis of the *Streptococcus dysgalactiae* species group: gene content, molecular adaptation, and promoter evolution. *Genome Biol. Evol.* **3**: 168–185, 2011.
45. Hubisz MJ, Pollard KS, **Siepel A**. PHAST and RPHAST: Phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**(1): 41–51, 2011.
46. Vinar T, Brejova B, Song G, **Siepel A**. Reconstructing histories of complex gene clusters on a phylogeny. *J. Comput. Biol.* **17**(9): 1267–1279, 2010.
47. Boyko AR, Quignon P, Li L, Schoenebeck JJ, Degenhardt JD, …, **Siepel A**, et al. (24 co-authors). A simple genetic architecture underlies morphological variation in dogs. *PLOS Biol.* **8**(8): e1000451, 2010.
48. da Fonseca RR, Kosiol C, Vinar T, **Siepel A**, Nielsen R. Positive selection on apoptosis related genes. *FEBS Lett.* **584**(3): 469–476, 2010.

49. Pollard KS, Hubisz MJ, Rosenbloom KR, **Siepel A.** Detection of non-neutral substitution rates on Mammalian phylogenies. *Genome Res.* **20**(1): 110–121, 2010.
50. The MGC Project Team. The completion of the Mammalian Gene Collection (MGC). *Genome Res.* **19**(12): 2324–2333, 2009.
51. Zhang Y, Song G, Vinar T, Green ED, **Siepel A**, Miller W. Evolutionary history reconstruction for Mammalian complex gene clusters. *J. Comput. Biol.* **16**(8): 1051–1070, 2009.
52. Holloway A, Begun D, **Siepel A**, Pollard KS. Accelerated sequence divergence of conserved genomic elements in *Drosophila melanogaster*. *Genome Res.* **18**(10): 1592–1601, 2008.
53. Wang Y, Diehl A, Wu F, Vrebalov J, Giovannoni J, **Siepel A**, Tanksley SD. Sequencing and comparative analysis of a conserved syntenic segment in the Solanaceae. *Genetics* **180**(1): 391–408, 2008.
54. 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.* **4**(8): e1000144, 2008.
55. Miller W, Rosenbloom K, Hardison RC, Hou M, Taylor J, …, **Siepel A**, et al. (27 co-authors). 28-way vertebrate alignment and conservation track in the UCSC genome browser. *Genome Res.* **17**(12): 1797–1808, 2007.
56. **Siepel A**, Diekhans M, Brejova B, Langton L, Stevens M, Comstock CLG, Davis C, Ewing B, Oommen S, Lau C, Yu H-C, Li J, Roe BA, Green P, Gerhard DS, Temple G, Haussler D, Brent MR. Targeted discovery of novel human exons by comparative genomics. *Genome Res.* **17**(12): 1763–1773, 2007.
57. Kininis M, Chen BS, Diehl AG, Isaacs GD, Zhang T, **Siepel AC**, Clark AG, Kraus WL. Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters. *Mol. Cell Biol.* **27**(14): 5090–5104, 2007.
58. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**(7146): 799–816, 2007.
59. Margulies EH, Cooper GM, Asimenos G, Thomas DJ, Dewey CN, **Siepel A**, et al. (42 co-authors). Analyses of deep mammalian sequence alignments and constraint predictions for 1% of the human genome. *Genome Res.* **17**(6): 760–774, 2007.
60. Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**(5822): 222–234, 2007.
61. Kuhn RM, Karolchick D, Zweig AS, Trumbower H, Thomas DJ, Thakkapallayil A, Sugnet CW, Stanke M, Smith KE, **Siepel A**, et al. (25 co-authors). The UCSC genome browser database: update 2007. *Nucleic Acids Res.* **35**: D668–D673, 2007.
62. Pollard KS, Salama SR, King B, Kern AD, Dreszer T, Katzman S, **Siepel A**, Pedersen JS, Bejerano G, Baertsch R, Rosenbloom KR, Kent J, and Haussler D. Forces shaping the fastest evolving regions in the human genome. *PLOS Genet.* **2**(10): e168, 2006.
63. Pollard KS, Salama SR, Lambert N, Lambot M-A, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, **Siepel A**, Kern AD, Dehay C, Igel H, Ares M, Vanderhaeghen P, and Haussler D. An RNA gene expressed during cortical development evolved rapidly in humans. *Nature* **443**(7108): 167–172, 2006.
64. Bejerano G, Lowe CB, Ahituv N, King B, **Siepel A**, Salama SR, Rubin EM, Kent WJ, and Haussler D. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* **441**(7089): 87–90, 2006.

65. Pedersen JK, Bejerano G, **Siepel A**, Rosenbloom K, Lindblad-Toh K, Lander ES, Kent J, Miller W, and Haussler D. Identification and classification of conserved RNA secondary structures in the human genome. *PLOS Comput. Biol.* **2**(4): e33, 2006.
66. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, Clawson H, Diekhans M, Furey TS, Harte RA, Hsu F, Hillman-Jackson J, Kuhn RM, Pedersen JS, Pohl A, Raney BJ, Rosenbloom KR, **Siepel A**, et al. (27 co-authors). The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* **34**: D590–598, 2006.
67. **Siepel A**, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, Weinstock GM, Wilson RK, Gibbs RA, Kent WJ, Miller W, and Haussler D. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**(8): 1034–1050, 2005.
68. Bejerano G, **Siepel AC**, Kent WJ, and Haussler D. Computational screening of conserved genomic DNA in search of functional noncoding elements. *Nat. Methods* **2**(7): 535–545, 2005.
69. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**(7018): 695–716, 2004.
70. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306**(5696): 636–640, 2004.
71. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431**: 931–945, 2004.
72. **Siepel A** and Haussler D. Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. *Mol. Biol. Evol.* **21**(3):468–488, 2004.
73. **Siepel A** and Haussler D. Combining phylogenetic and hidden Markov models in biosequence analysis. *J. Comput. Biol.* **11**(2-3): 413–428, 2004.
74. Thomas JW, Touchman JW, Blakesley RW, Bouffard GG, Beckstrom-Sternberg, SM, Margulies EH, Blanchette M, **Siepel AC**, et al (71 co-authors). Comparative analyses of multi-species sequences from targeted genomic regions. *Nature* **424**(6950): 788–793, 2003.
75. **Siepel AC**. An algorithm to enumerate sorting reversals for signed permutations. *J. Comput. Biol.* **10**(3-4): 575–597, 2003.
76. **Siepel AC**, Steenhuis TS, Rose CW, Parlange J-Y, McIsaac GF. A simplified hillslope erosion model with vegetation elements for practical applications. *J. Hydrol.* **258**(1-4): 111–121, 2002.
77. **Siepel A**, Tolopko A, Farmer A, Steadman P, Schilkey F, Perry BD, and Beavis W. An integration platform for heterogeneous bioinformatics software components. *IBM Systems Journal* **40**(2): 570–591, 2001.
78. **Siepel A**, Farmer A, Tolopko A, Zhuang M, Mendes P, Beavis W, and Sobral B. ISYS: a decentralized, component-based approach to the integration of heterogeneous bioinformatics resources. *Bioinformatics* **17**(1): 83–94, 2001.
79. Skupsk**i** MP, Booker M, Farmer A, Harpold M, Huang W, Inman J, Kiphart D, Kodira C, Root S, Schilkey F, Schwertfeger J, **Siepel A**, et al. (18 co-authors). The Genome Sequence DataBase: towards an integrated functional genomics resource. *Nucleic Acids Res.* **27**(1): 35–38, 1999.
80. Harger C, Skupski M, Bingham J, Farmer A, Hoisie S, Hraber P, Kiphart D, Krakowski L, McLeod M, Schwertfeger J, Seluja G, **Siepel A**, et al. (21 co-authors). The Genome Sequence DataBase (GSDB): improving data quality and data access. *Nucleic Acids Res.* **26**(1): 21–26, 1998.

81. Harger C, Skupski M, Allen E, Clark C, Crowley D, …, **Siepel A**, et al. (25 co-authors). The Genome Sequence DataBase version 1.0 (GSDB): from low pass sequences to complete genomes. *Nucleic Acids Res.* **25**(1): 18–23, 1997.
82. **Siepel AC**, Halpern AL, Macken C, and Korber BTM. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res. Hum. Retroviruses* **11**(11): 1413–1416, 1995.

Refereed Correspondence

83. Dukler N, Gulko B, Huang Y-F, **Siepel A**. Is a super-enhancer greater than the sum of its parts? Correspondence. *Nat. Genet.* **49**(1): 2–3, 2016.

Book Chapters and Review Articles

84. Hubisz MJ, **Siepel AC**. ARGweaver. In: Dutheil J (ed.), Statistical Population Genetics. Springer, 2018. Book Chapter. In Press.
85. **Siepel A**, Arbiza L. Cis-regulatory elements and human evolution. *Curr. Opin. Genet. Dev.* **29**C: 81–89. Review. 2014.
86. **Siepel A.** Phylogenomics of primates and their ancestral populations. *Genome Res.* **19**(11): 1929–1941. Review. 2009.
87. **Siepel A.** Darwinian alchemy: human genes from noncoding DNA. *Genome Res.* **19**(10): 1693–1695. Review. 2009.
88. **Siepel A**, Haussler D. Phylogenetic Hidden Markov Models. In: Nielsen R (ed.), Statistical Methods in Molecular Evolution. Book Chapter. Springer, 2005.

Refereed Conference Papers

89. Vinar T, Brejova B, Song G, **Siepel A.** Reconstructing histories of complex gene clusters on a phylogeny. In: Proceedings of RECOMB-CG 2009, Seventh Annual International Workshop on Comparative Genomics. LNCS Vol. 5817, Springer, 2009.
90. Zhang Y, Song GT, Vinar T, Green ED, **Siepel A**, Miller W. Reconstructing the evolutionary history of complex human gene clusters. In: Proceedings of RECOMB 2008, the Twelfth Annual International Conference on Research in Computational Molecular Biology. LNCS Vol. 4955. Springer-Verlag, 2008.
91. Sobral BWS, Mangalam H, **Siepel A**, Mendes P, Pecherer R, McLaren G. Bioinformatics for rice resources. In: Proceedings of Novartis Foundation Symposium 236. Rice Biotechnology: Improving Yield, Stress Tolerance and Grain Quality. John Wiley & Sons, 2001.
92. **Siepel A**, Pollard KS, and Haussler D. New methods for detecting lineage-specific selection. In: Proceedings of RECOMB 2006, Tenth Annual International Conference Research on Research in Computational Molecular Biology. LNBI Vol. 3909. Springer-Verlag, 2006.
93. Jojic V, Jojic N, Meek C, Geiger D, **Siepel A**, Haussler D, and Heckerman D. Efficient approximations for learning phylogenetic HMM models from data. In: Proceedings of ISMB / ECCB 2004, 12th International Conference on Intelligent Systems for Molecular Biology / 3rd European Conference on Computational Biology. *Bioinformatics*, 20, Suppl.1:i161-i168, 2004.

94. **Siepel A** and Haussler D. Computational identification of evolutionarily conserved exons. In: Proceedings of RECOMB 2004, Eighth Annual International Conference on Research in Computational Molecular Biology. ACM Press, 2004.
95. **Siepel A** and Haussler D. Combining phylogenetic and hidden Markov models in biosequence analysis. In: Proceedings of RECOMB 2003, Seventh Annual International Conference on Research in Computational Molecular Biology. ACM Press, 2003.
96. **Siepel A**. An algorithm to enumerate all sorting reversals. In: Proceedings of RECOMB 2002, Sixth Annual International Conference on Computational Molecular Biology. ACM Press, 2002.
97. Moret BME, **Siepel AC**, Tang J, Liu T. Inversion Medians Outperform Breakpoint Medians in Phylogeny Reconstruction from Gene-Order Data. In: Proceedings of WABI 2002, Second International Workshop on Algorithms in Bioinformatics. LNCS Vol. 2452. Springer, 2002.
98. **Siepel AC** and Moret BME. Finding an Optimal Inversion Median: Experimental Results. In: Proceedings of WABI 2001, First International Workshop on Algorithms in Bioinformatics LNCS Vol. 2149. Springer-Verlag, 2001.

Teaching

Quantitative Biology I, II, and Genetics & Genomics Course Lecturer, Watson School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 2014–present.

Graduate Special Topics in Biometry & Statistics Course: Genomics of Gene Regulation (with John Lis), Cornell University, Ithaca, NY, 2014.

Probabilistic Graphical Models Course, Cornell University, Ithaca, NY, 2008, 2010, 2013.

Computational Genetics & Genomics Course, Cornell University, Ithaca, NY, 2006–2007, 2009, 2011.

Graduate-level Biometry & Statistics Course: Applied Bioinformatics (with Jason Mezey, Alon Keinan, and Haiyuan Yu), Cornell University, Ithaca, NY, 2010–2011.

Summer Course in Evolutionary Genomics: Computational Genomics, Kunming Institute of Zoology (KIZ), Chinese Academy of Sciences, Kunming, China, August 2011.

Comparative Genomics/Molecular Evolution, Summer Institute in Statistical Genetics (with Spencer Muse), School of Public Health, University of Washington, Seattle, Washington, 2008–2010.

Graduate-level Biometry & Statistics Course: Topics in Computational Genomics, Cornell University, Ithaca, NY, 2006–2010.

Trainees

list	First	Last	Type	Joint Mentor	Start	End
1	Leonardo	Arbiza	Postdoc	A. Keinan	2010	2013
2	Amit	Blumberg	Postdoc	—	2016	Present
3	Bronislava	Brejova	Postdoc	—	2006	2008
4	Sang Chul	Choi	Postdoc	—	2010	2013
5	Omar	Cornejo	Postdoc	C.D. Bustamante	2009	2010
6	Charles	Danko	Postdoc	W.L. Kraus	2009	2014

7	Alexandra	Denby	Undergrad	—	2006	2008
8	Adam	Diehl	Grad student, Masters	—	2006	2010
9	Noah	Dukler	Grad student, PhD	—	2014	Present
10	Ilan	Gronau	Postdoc	—	2009	2014
11	Brad	Gulko	Grad student, PhD	—	2010	2017
12	Hussein	Hijazi	Postdoc	—	2018	Present
13	Yifei	Huang	Postdoc	—	2015	Present
14	Melissa	Hubisz	Grad student, PhD	—	2014	Present
15	Elizabeth	Hutton	Grad student, PhD	—	2015	Present
16	Carolin	Kosiol	Postdoc	C.D. Bustamante	2006	2008
17	Tytus	Mak	Undergrad	—	Spring 2006	Spring 2006
18	Alison	Marklein	Undergrad	—	Summer 2006	Summer 2006
19	Andre L.	Martins	Grad student, PhD	—	2007	2014
20	Jaaved	Mohammed	Grad student, PhD	E. Lai	2009	2016
21	Michael D.	Phillips	Grad student, Masters	E. Lai	2008	2010
22	Lenore	Pipes	Grad student, PhD	C. Mason	2012	2017
23	Joseph	Porter	Undergrad	—	Spring 2012	Spring 2012
24	Matthew	Rasmussen	Postdoc	—	2011	2013
25	Hirak	Sarkar	Summer Intern	—	2016	2016
26	Renee	Setter	Undergrad	—	Summer 2010	Fall 2010
27	Daniel	Sussman	Undergrad	—	Summer 2007	Summer 2008
28	Tomas	Vinar	Postdoc	—	2006	2008
29	Alexander	Xue	Postdoc	—	2018	Present
30	Yixin	Zhao	Postdoc	—	2018	Present

Thesis Committees

Current

Laraib Malik, Ph.D., Computer Science, Stony Brook University.

Kathryn O’Neil, (Committee Chair), Watson School of Biological Sciences, Cold Spring Harbor Laboratory.

Monica Ramstetter, Ph.D., Computational Biology, Cornell University.

Hirak Sarkar, Ph.D., Computer Science, Stony Brook University.

Avi Srivastava, Ph.D., Computer Science, Stony Brook University.

Nathaniel Tippens, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program.

Jonathan Werner, Ph.D., (Academic Mentor), Watson School of Biology Sciences, Cold Spring Harbor Laboratory.

Cole Wunderlich, Ph.D., (Academic Mentor), Watson School of Biology Sciences, Cold Spring Harbor Laboratory.

Completed

B. Arman Aksoy, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2015.

Samuel Arbesman, Ph.D., Computational Biology, Cornell University, 2008.

Tara Baxter, M.S., Genetics and Development, Cornell University, 2014.

Jeremiah Degenhardt, Ph.D., Computational Biology, Cornell University, 2010.

Talitha Forcier, Ph.D., (Committee Chair), Watson School of Biological Sciences, Cold Spring Harbor Laboratory, 2018.

Feng Gao, Ph.D., Computational Biology, Cornell University 2016.

Stephanie Hyland, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2018.

Gabriel Hoffman, Ph.D., Genetics and Development, Cornell University, 2013.

Haley Hunter-Zinck, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2014.

Helgi Ingolfsson, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2010.

Hayan Lee, Ph.D., Computer Science, Stony Brook University, 2015.

Aaron Lenfestey, Ph.D., Computer Science, Cornell University, qualified 2009.

Maria Nattestad, Ph.D., Watson School of Biological Sciences, Cold Spring Harbor Laboratory, 2017.

Yogesh Saletore, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2015.

Solomon Shenker, Ph.D., Tri-Institute Computational Biology and Medicine (CBM) program, 2016.

Molly Shook, Ph.D., Genetics and Development, Cornell University, qualified 2009.

Nick Stepankiw, Ph.D., Biochemistry, Molecular & Cell Biology (BMCB), Cornell University, 2016.

Paula Tataru, Ph.D., Aarhus University, Aarhus, Denmark, 2015.

Xu Wang, Ph.D., Genetics and Development, Cornell University, 2011.

Katherine Wilkins, Ph.D., Computational Biology, Cornell University, 2016.

Erin Wissink, Ph.D., Genetics and Development, Cornell University, left committee 2014.

Wendy Wong, Ph.D., Biometry, Cornell University, 2006.

Chun-Nam Yu, Ph.D., Computer Science, Cornell University, 2010.

Other Service

Co-organizer NYPG'19: Fifth New York Area Population Genomics Workshop (with Christopher Brown, Barbara Engelhardt, Eimear Kenny, Joe Pickrell), Mt. Sinai, January 2019.

Co-organizer ProbGen'18: 4th Probabilistic Modeling in Genomics Conference (with Gerton Lunter and Molly Przeworski), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, November 2018.

Co-organizer: Workshop on Standardized Simulated Data Sets for Population Genomics (PopSim), inaugural meeting (with Ryan Gutenkunst, Kelley Harris, and Dmitri Petrov), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, November 2018.

Affiliate Member, New York Genome Center, New York, NY, 2018–present.

Guest Member, 2019 Editorial Committee Planning Meeting for the *Annual Review of Genomics and Human Genetics*, New York, NY, January 2018.

Co-organizer NYPG'18: Fourth New York Area Population Genomics Workshop (with Casey Brown, Barbara Engelhardt, Joe Pickrell, Molly Przeworski and, Matt Rockman), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, January 2018.

Co-organizer NYQB’17: New York Area Meeting in Quantitative Biology: Making Use of Emerging Technologies (with Alex Krasnitz, Neville Sanjana, Rahul Satija, and Olga Troyanskaya), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, August 2017.

Advisory Board Member, School of Engineering Bioinformatics, New York University, New York, NY, 2015–present.

Admissions Committee, Watson School of Biological Sciences (WSBS), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 2015–present.

Graduate Field Membership, Genetics Program, Stony Brook University, Stony Brook, NY, 2015–present.

Graduate Field Membership, Physiology, Biophysics and Systems Biology Program, Weill Cornell Medical College of Cornell University, New York, NY, 2015–present.

Graduate Student Admissions Committee, Tri-Institutional MD-PhD Training Program in Computational Biology & Medicine, Cornell University, Ithaca, NY, 2006–present.

Developer of freely available software package: PHAST (PHylogenetic Analysis with Space/Time models), 2003–present.

Contributor: University of California Santa Cruz Genome Browser (Conservation, ExoniPhy, DLESS, and Positive Selection tracks), 2003–present.

Co-organizer NYQB’16: New York Area Meeting in Quantitative Biology: Computational and Statistical Genomics: Applications to Disease, inaugural meeting (with Christina Leslie, Bud Mishra, Joe Pickrell, and Olga Troyanskaya), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, July 2016.

Co-organizer: Measuring and Modeling Quantitative Sequence-function Relationships Meeting (with Justin B. Kinney and Douglas M. Fowler), The Banbury Center, Cold Spring Harbor Laboratory, Huntington, NY, July 2016.

Co-organizer ProbGen’15: First Probabilistic Modeling in Genomics Conference, inaugural meeting (with Barbara Engelhardt and Thomas Mailund), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, October 2015.

Co-organizer NYPG’15: First New York Area Population Genomics Workshop, inaugural meeting (with Barbara Engelhardt, Joe Pickrell, Molly Przeworski, Matt Rockman, et al.), New York Genome Center, New York, NY, January 2015.

Co-organizer: Biological Sequence Analysis and Probabilistic Models Workshop (with Anders Krogh, Gerton Lunter, Gilean McVean, Molly Przeworski, Richard Durbin, and Thomas Mailund), Merton College, Oxford, UK, July 2014.

Director of Graduate Studies, Graduate Field of Computational Biology, Cornell University, Ithaca, NY, 2011–2014.

Endorsement Committee, Cornell University Nominees of the Churchill Scholarship, Ithaca, NY, 2009–2014.

Associate Director, Center for Comparative and Population Genomics, Cornell University, Ithaca, NY, 2008–2014.

Advisory Committee, Weill Institute for Cell and Molecular Biology, Cornell University, Ithaca, NY, 2007–2014.

Graduate Field Membership, Cornell University, Ithaca, NY, 2006–2014:
  Statistics, 2013–2014.
  Applied Math, 2009–2014.
  Computational Biology, Biometry, Computer Science, and Genetics & Development, 2006–2014.

Faculty Search Committees, Cornell University, Ithaca, NY, 2007–2014:
  Biological Statistics and Computational Biology, 2008, Committee Chair 2012 and 2014.
  Computer Science, 2007–2009, 2012.
  Molecular Biology and Genetics, 2008, 2012.
  Department of Biological Statistics and Computational Biology (BSCB) / Weill Institute, 2008.

Faculty Advisor for Information Technology, Life Sciences Core Laboratories Center, Cornell University, Ithaca, NY, 2008–2012.

Faculty Senate for College of Agriculture and Life Sciences, Cornell University, Ithaca, NY, 2008–2011.

International Genomics Consortia, 2003–2011:
  Orangutan Sequencing and Analysis, 2008–2011.
  2x Mammals Sequencing and Analysis 2008–2011.
  Mammalian Gene Collection (MGC) Project, 2004–2009.
  Rhesus Macaque Sequencing and Analysis, 2006–2007.
  ENCyclopedia Of DNA Elements (ENCODE), 2004–2007.
  International Human Genome Sequencing, 2004.
  International Chicken Genome Sequencing, 2003–2004.

Strategic Planning Working Group on "Research, Scholarship and Creativity", Cornell University, Ithaca, NY, 2009.

NHGRI Advisory Committee for Primate Genome Sequencing, National Human Genome Research Institute, Bethesda, MD, 2007–2009.

Undergraduate Biology Curriculum Task Force, Cornell University, Ithaca, NY, 2007–2008.

Research Support

Current

Program and/or Project Title	PI, Institution	Source, Grantor #	Total Award	Award Period
RESEARCH-PGR: PanAnd – Harnessing convergence and constraint to predict adaptations to abiotic stress for maize and sorghum	Buckler, Cornell Co-I: Siepel, CSHL	NSF / BIO IOS-1822330	$ 394,429 Subawd only	9/1/18 – 8/31/22
Evolutionary Human Genomics: Demography, Natural Selection, and Transcriptional Regulation	Siepel, CSHL	NIH / NIGMS R35 GM127070	$ 2,396,075	3/1/18 – 2/28/23
Collaborative Research: NSFDEB-BSF: Quantifying genomic porosity in non-model radiations	Siepel, CSHL Lovette, Cornell	NSF / BIO DEB1555769	$ 197,334	3/1/16 – 2/28/19
Continued development and maintenance of the PHAST software for comparative genomics	Siepel, CSHL	NIH / NHGRI R01 HG008161	$ 576,000 NCE to 1/31/19	4/24/15 – 1/31/19
New York Center for Collaborative Research in Common Disease Genomics	Darnell, NYGC Subawd Co-I: Siepel, CSHL	NIH / NHGRI UM1 HG008901 *Renewal pending	$ 97,506 Subawd only	1/14/16 – 11/30/18*

Completed

Program and/or Project Title	**PI, Institution**	**Source, Grantor #**	**Total Award**	**Award Period**
Computational methods for human genomic data integration: demography, selection, and functional potential	Siepel, CSHL	NIH / NIGMS R01 GM102192	$ 1,722,155	1/1/13 – 12/31/17
Factor-general characterization of dynamic transcriptional stress responses	Siepel, CSHL Lis, Cornell	NIH / NHGRI R01 HG007070	$ 1,011,632 NCE to 5/31/17	9/1/13 – 5/31/17
Packard Fellowship for Science and Engineering in Biological Sciences	Siepel, Cornell	David & Lucile Packard Foundation 2007–31762	$ 825,000 NCE to 8/31/14	9/1/07 – 8/31/14
Evolutionary genomics and population genetics of pathogenic streptococci	Stanhope, Cornell Co-I's: Siepel, Cornell Bustamante, Cornell Burne, U. Florida	NIH / NIAID R01 AI073368	$ 2,252,000	7/1/08 – 6/30/13
Coordinate gene regulation in animal cells (Administrative Supplement)	Lis, Cornell Co-I: Siepel, Cornell	NIH / NIGMS R01 GM025232	$ 1,137,000	4/1/09 – 3/31/13
What made us human?	Pollard, Gladstone Inst., UCSF Co-I: Siepel, Cornell	NIH / NIGMS R01 GM082901	$ 274,000 Siepel Subawd only	9/1/08 – 8/31/12
NSF CAREER Award: Models of evolution and discovery of functional elements in mammals and *Drosophila*	Siepel, Cornell	NSF / BIO DBI 0644111	$ 645,870	3/1/07 – 2/29/12
Sloan Research Fellowship in Molecular Biology	Siepel, Cornell	Alfred P. Sloan Foundation	$ 50,000	2009 – 2011
Microsoft Research Faculty Fellowship Program	Siepel, Cornell	Microsoft Research	$ 200,000 lump sum	2007
Prediction of novel human protein-coding genes using cross-species alignments and phylogenetic hidden Markov models	Haussler, UCSC Subawd PI: Siepel, Cornell	NIH / NCI Subawd 22XS013A	$ 159,000 Subawd only	5/15/06 – 11/14/07

Invited Conference and Symposium Lectures

SMBE 2018, 50th Anniversary of the Neutral Theory of Molecular Evolution, and Society for Molecular Biology and Evolution Annual Meeting, Yokohama, Japan, July 2018.

Thirteenth Conference on Systems Biology: Global Regulation of Gene Expression, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, March 2018.

Genome Informatics Conference, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, November 2017.

ProbGen17, Keynote Lecture, 3rd Probabilistic Modeling in Genomics Conference, Aarhus University, Denmark, September 2017.

Revolutionizing Next-Generation Sequencing (2nd Edition), Antwerp, Belgium, March 2017.

Genomics & Systems Biology VII Conference, NYU Abu Dhabi, Saadiyat Campus, Abu Dhabi, United Arab Emirates, February 2017.

66th Annual Meeting of the American Society of Human Genetics (ASHG), Vancouver, BC, Canada, October 2016.

ProbGen16, 2nd Probabilistic Modeling in Genomics Conference, University of Oxford, UK, September 2016.

ICQG5, 5th International Conference in Quantitative Genetics, Madison, WI, June 2016

Genome Informatics Conference, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, October 2015.

Society for Molecular Biology and Evolution Annual Meeting (SMBE 2015), Vienna, Austria, July 2015.

Twelfth RECOMB Satellite Conference on Comparative Genomics (RECOMB-CG), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, October 2014.

Keynote Lecture, Mathematical and Computational Evolutionary Biology Conference (MCEB2014), Montpellier Laboratory of Computer Science, Robotics, and Microelectronics; Montpellier, France, June 2014

11th Basel Computational Biology Conference [BC]2, Basel, Switzerland, July 2013.

26th Annual Meeting of The Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, May 2013.

24th Annual Meeting of The Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, May 2011.

Second RECOMB Satellite Conference on Bioinformatics Education (RECOMB-BE), University of California, San Diego, La Jolla, CA, May 2010.

Society for Molecular Biology and Evolution Annual Meeting (SMBE 2009), Iowa City, IA, June 2009.

22nd Annual Meeting on The Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, May 2009.

8th Annual Symposium on Computational Molecular Biology, Carnegie Mellon University, Pittsburgh, PA, April 2009.

10th Annual Meeting on Advances in Genome Biology and Technology (AGBT), Marco Island, FL, February 2009.

Frontiers of Evolutionary Biology Workshop, Pennsylvania State University, State College, PA, October 2008.

Packard Foundation 20th Anniversary Fellows Reunion, Park City, UT, September 2008.

Biomedical Engineering Society (BMES) Annual Meeting, Chicago, IL, October 2006.

8th International Bioinformatics Conference, Society for Bioinformatics in the Nordic Countries (SocBIN), Aarhus, Denmark, June 2006.

19th Annual Meeting of The Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, May 2006.

First International Conference on Phylogenomics, Sainte-Adele, Quebec, Canada, March 2006.

17th Annual Meeting of The Biology of Genomes, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, May 2004.

Invited Seminars and Workshops

McKusick-Nathans Institute of Genetic Medicine Lecture Series, Johns Hopkins University School of Medicine, Baltimore, MD, September 2018.

NIH Genome Sequencing Program Meeting, National Human Genome Research Institute (NHGRI), Bethesda, MD, April 2018.

Advanced Genomics Unit Seminar, National Laboratory of Genomics for Biodiversity (LANGEBIO Cinvestav), Irapuato, Mexico, February 2018.

Horizons in Genomic Research Seminar, International Laboratory for Human Genome Research, National University of Mexico (LIIGH-UNAM), Queretaro, Mexico, February 2018.

Charles Bronfman Institute for Personalized Medicine (IPM) Seminar Series, Icahn School of Medicine at Mt. Sinai, New York, NY, December 2017.

Challenges and Synergies in the Analysis of Large-Scale Population-Based Biomedical Data Workshop, Banff International Research Station for Mathematical Innovation and Discovery, Casa Matemática Oaxaca (BIRS-CMO), Oaxaca, Mexico, November 2017.

Laufer Center for Physical and Quantitative Biology Seminar Series, Stony Brook University, Stony Brook, NY, October 2017.

Computational Biology PhD Program Seminar Series, Carnegie Mellon-University of Pittsburgh, Pittsburgh, PA, September 2017.

Bioinformatics Seminar Series, University of California, Los Angeles, CA, February 2017.

Mathematical Biology Seminar Series, University of Pennsylvania, Philadelphia, PA, January 2017.

Quantitative Biology Seminar Series, Center for Computational Biology, University of California, Berkeley, CA, October 2016.

Mechanisms of Transcription: Pausing to Celebrate John Lis, Cornell University, Ithaca, NY, July 2016.

Measuring and Modeling Quantitative Sequence-function Relationships Meeting, The Banbury Center, Cold Spring Harbor Laboratory, Huntington, NY, July 2016.

April is Math Awareness Month Seminar Series, Departments of Mathematical Sciences and Computer Science, Worcester Polytechnic Institute, Worcester, MA, April 2016.

Monthly Colloquium Seminar Series, Renaissance Technologies, East Setauket, NY, January 2016.

Mathematics and the Quest for Fundamental Principles of Biology Workshop, Center for Quantitative Biology, University of Utah, Salt Lake City, UT, December 2015.

Biostatistics and Medical Informatics (BMI) Department Seminars, University of Wisconsin-Madison, Madison, WI, September 2015.

Evolutionary Biology and the Theory of Computing Reunion Workshop, Simons Institute for the Theory of Computing, University of California, Berkeley, CA, July 2015.

Invited Lecturer: Statistical Methods for Functional Genomics Course, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, June 2015.

Department of Computer Science, Columbia University, New York, NY, June 2015.

Computational Biology Center (cBio) Annual Scientific Retreat, Memorial Sloan-Kettering Cancer Center, Norwalk, CT, May 2015.

Paleobiology Seminar Series, Stony Brook School of Medicine, Stony Brook University, Stony Brook, NY, April 2015.

From Genome Function to Biomedical Insight: ENCODE and Beyond Workshop, NHGRI, Bethesda, MD, March 2015.

Understanding Non-coding DNA Through Intra- and Inter-species Epigenomic Variation Workshop, (organized by Guillaume Bourque and Tomi M. Pastinen), McGill University, Bellairs Research Institute, Holetown, Barbados, January 2015.

Genetics Department Seminar, Rutgers University, Piscataway, NJ, November 2014.

Computer Science Department Colloquium Series: Topics in Modern Computer Science, Stony Brook University, Stony Brook, NY, October 2014.

Organismic and Evolutionary Biology (OEB) Seminar Series, College of Natural Sciences, University of Massachusetts, Amherst, MA, September 2014.

Institute for Computational Biomedicine (ICB) Seminar Series, Weill Cornell Medical College of Cornell University, New York, NY, September 2014.

Co-Organizer: Biological Sequence Analysis and Probabilistic Models Workshop (with Gerton Lunter, Thomas Mailund, Gilean McVean, Richard Durbin, Anders Krogh, and Molly Przeworski), Merton College, Oxford, UK, July 2014.

Computational Biology Institute, University of Montpellier 2, Montpellier, France, June 2014.

Computational Biology & Bioinformatics Program, Duke University, Durham, NC, April 2014.

Department of Biostatistics Colloquium, Harvard School of Public Health, Boston, MA, March 2014.

Bioinformatics Forum, University of Pennsylvania, Philadelphia, PA, February 2014.

Cold Spring Harbor Laboratory Seminar, Cold Spring Harbor, NY, December 2013.

Department of Human Genetics Seminar, University of Chicago, Chicago, IL, November 2013.

Center for Bioinformatics Research Seminar, Indiana University, Bloomington, IN, September 2013.

Wellcome Trust Centre for Human Genetics Seminar, University of Oxford, UK, May 2013.

Cambridge Statistics Initiative Seminar, Centre for Mathematical Sciences, University of Cambridge, Cambridge, UK, April 2013.

Co-organizer: Biological Sequence Analysis and Probabilistic Models Workshop, (with Katherine Pollard and Sean Eddy), Janelia Farm Research Campus, Ashburn, VA, March 2013.

European Bioinformatics Institute (EMBL-EBI), Hinxton, UK, February 2013.

Microsoft Research Ltd., Cambridge, UK, February 2013.

Claude Bernard University Lyon 1, Villeurbanne, France, January 2013.

Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, December 2012.

Bioinformatics Research Centre, Aarhus University, Denmark, November 2012.

Bioinformatics Centre, Department of Biology, University of Copenhagen, Denmark, October 2012.

Population Biology, Ecology, and Evolution (PBEE) Graduate Program, Emory University, Atlanta, GA, February 2012.

Evolutionary Studies (EvoS) Program, SUNY Binghamton, NY, December 2011.

Evolutionary Genomics Seminar Series (online phylogenetics seminar), Phyloseminar.org, May 2011.

Vienna Graduate School of Population Genetics, Department of Biomedical Sciences, University of Veterinary Medicine, Vienna, Austria, March 2011.

Comenius University, Bratislava, Slovakia, March 2011.

University of Geneva, Switzerland, March 2011.

BioMaPS Summer School: New Directions in Evolutionary and Population Genetics, Rutgers University, Piscataway, NJ, June 2010.

Gene Histories Versus Species Histories Workshop, Bertinoro, Italy, May 2010.

Meeting of Twenty Leading Experts in Bioinformatics, Computational Center for Algorithmic and Systems Biology (CASB-20), University of California, San Diego, La Jolla, CA, May 2010.

Inference in Stochastic Models for Sequence Evolution Workshop, Mathematical Biosciences Institute, Columbus, OH, February 2010.

Physics Department, Syracuse University, Syracuse, NY, January 2010.

Cross-campus Epigenomics Seminar Series, Weill Cornell Medical College of Cornell University, New York, NY, January 2010.

Bioinformatics Seminar Series, UCLA, Los Angeles, CA, January 2010.

National Human Genome Research Institute (NHGRI), Bethesda, MD, March 2009.

Applied Math Colloquium, Cornell University, Ithaca, NY, October 2008.

Mammalian Gene Collection (MGC), Final External Steering Committee Meeting, Rockville, MD, September 2008.

ISMB 2008 Special Interest Group: Genome-scale Pattern Analysis in the Post-ENCODE Era, Toronto, Ontario, Canada, July 2008.

Microsoft Research Faculty Summit, Redmond, WA, July 2008.

Tree Thinking at Cornell Symposium, Laboratory of Ornithology, Cornell University, Ithaca, NY, May 2008.

Evolutionary Genetics Seminar Series, University of California, Davis, CA, May 2008.

Department of Developmental Biology, Stanford University, Stanford, CA, May 2008.

Department of Genome Sciences, University of Washington, Seattle, WA, March 2008.

Department of Statistics, University of Washington, Seattle, WA, March 2008.

Microsoft Research, Redmond, WA, March 2008.

Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Canada, October 2007.

Center for Comparative Genomics, University of Copenhagen, Denmark, June 2007.

3rd Barbados Workshop on Computational Gene Regulation, McGill University, Bellairs Research Institute, Holetown, Barbados, April 2007.

NESCent Workshop on *cis*-regulatory molecular evolution, National Evolutionary Synthesis Center, Durham, NC, March 2007.

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, September 2006.

Computational and Statistical Genomics Workshop, Banff International Research Station, Banff, Canada, July 2006.

1st Barbados Workshop on Paleogenomics, McGill University, Bellairs Research Institute, Holetown, Barbados, April 2006.

Centre for Bioinformatics, McGill University, Montreal, Quebec, Canada, March 2006.

Center for Comparative Genomics and Bioinformatics, Pennsylvania State University, University Park, PA, February 2006.

Department of Entomology, Cornell University, Ithaca, NY, January 2006.

Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, April 2005.

Department of Biology and Courant Institute, New York University, New York, NY, March 2005.

Department of Computer Science, UC Berkeley, Berkeley, CA, February 2005.

Department of Computer Science and Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY, February 2005.

Department of Statistics, Pennsylvania State University, University Park, PA, February 2005.

Department of Genome Sciences, University of Washington, Seattle, WA, January 2005.

Department of Computer Science, Stanford University, Stanford, CA, October 2003.