Adam C. Siepel, Ph.D.

PERSONAL DETAILS

Present position: Professor, Simons Center for Quantitative Biology
Cold Spring Harbor Laboratory

Address: Cold Spring Harbor Laboratory
2121 Koch Building
One Bungtown Road
Cold Spring Harbor, NY 11724–2209
Tel: 516-367-6922
516-367-5501-fax
asiepel@cshl.edu
http://SiepelLab.labsites.cshl.edu
http://www.cshl.edu/scqb
https://www.ncbi.nlm.nih.gov/myncbi/adam.siepel.3/bibliography/public/

EDUCATION AND PROFESSIONAL QUALIFICATIONS

Education

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

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

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

Professional Experience

2014 – Present Professor, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

2025 – Present Adjunct Professor, Department of Human Genetics, McGill University, Montreal, Canada

2015 – Present Adjunct Professor, Department of Computational Biology, Cornell University, Ithaca, NY

2014 – Present Adjunct Associate Professor, Department of Computer Science, Stony Brook University, Stony Brook, NY

2014 – 2024 Chair, Simons Center for Quantitative Biology (SCQB), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

2014 – 2019 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 (EMBL-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, ARCS Foundation, Inc., Northern California Chapter, 2003–2004.

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

RESEARCH SUPPORT

Current Research Support

Project Name	PI(s), Affiliation(s)	Funding Source, Grantor #	Total Award	Award Period
A unified probabilistic model and software implementation for analysis of nascent RNA sequencing data	Siepel, CSHL Co-I: Koo, CSHL Co-I: Pai, UMass	NIH / NHGRI R01 HG012944	$2,309,517	9/16/24 – 9/15/28
Evolutionary Human Genomics: demography, natural selection, and transcriptional regulation	Siepel, CSHL	NIH / NIGMS R35 GM127070	$2,880,000	3/1/23 – 2/28/28
CSHL Cancer Center Support Grant – Cancer Genetics and Genomics (CCG) Program	Tuveson, CSHL Co-Leaders: Siepel, CSHL and Janowitz, CSHL	NIH / NCI P30 CA045508	$132,926 CCG Program Only	8/1/16 – 7/31/26
Novel mouse models for quantitative understanding of baseline and therapy-driven evolution of prostate cancer metastasis	Nowak, Weill-Cornell Co-I: Siepel, CSHL	NIH / NCI R01 CA272466	$465,374 (subaward only)	8/1/16 – 7/31/26

Completed Research Support

Project Name	PI(s), Affiliation(s)	Funding Source Grantor #	Total Award	Award Period
Defining the cellular & molecular effects of aging and age of pregnancy on breast tissue homeostasis & cancer initiation	Dos Santos, CSHL Co-I: Siepel, CSHL	NIH / NIA R01 AG069727	$353,891	9/30/20 – 5/31/25
Determining and Targeting Evolutionary Trajectories Driving Bladder Cancer	Siepel, CSHL	Starr Foundation I16-0060	$150,000	1/1/23 – 12/31/24
Evolutionary reconstruction of metastasis in new model of prostate cancer	Siepel, CSHL	CSHL CC Pilot Study	$100,000	1/01/24 – 12/31/24
Evolutionary Human Genomics: demography, natural selection, and transcriptional regulation	Siepel, CSHL	NIH / NIGMS R35 GM127070	$2,951,205	3/1/18 – 2/28/23
Evolution of chromatin architecture and transcriptional regulation in mammals	Danko, Cornell Siepel, CSHL	NIH / NHGRI R01 HG010346	$1,253,628	5/1/19 – 2/28/23
RESEARCH-PGR: PanAnd - Harnessing convergence and constraint to predict adaptations to	Buckler, Cornell Co-PI: Siepel, CSHL	NSF PGRP IOS-1822330	$394,429 subawd only	9/1/18 – 8/31/22

abiotic stress for maize and sorghum				
Graphical Processing Units and a Large-Memory Compute Node for Applications in Genomics, Neuroscience, and Structural Biology	Siepel, CSHL	NIH / OD S10 OD028632	$436,882	7/1/20 – 6/30/21
New York Center for collaborative research in common disease genomics	Maniatis, NYGC Co-I: Siepel, CSHL	NYGC / NHGRI UM1 HG008901	$148,279 subawd only	1/14/16 – 11/30/19
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
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 Fndn. 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	Lis, Cornell Co-I: Siepel, Cornell	NIH / NIGMS R01 GM025232	$1,137,000 administrative supplement	4/1/09 – 3/31/13
What made us human?	Pollard, UCSF Gladstone Co-I: Siepel, Cornell	NIH / NIGMS R01 GM082901	$274,000 subawd only	9/1/08 – 8/31/12
NSF Faculty Early Career Development (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

Development of a web interface for Mammalian Gene Collection program data. Subawd: Prediction of novel human protein-coding genes using cross-species alignments and phylogenetic hidden Markov models	Haussler, UCSC Siepel, Cornell	NIH / NCI subawd 22XS013A	$159,000 subawd only	5/15/06 – 11/14/07

PUBLICATIONS

Total citations: 44,264 **h-index: 70** [*Google scholar*]

Journal Articles

1. Struck TJ, Vaughn AH, Daigle A, …, **Siepel A**, Gutenkunst RN (17 co-authors). GHIST 2024: The first Genomic History Inference Strategies Tournament. *Mol Biol Evol.* 42(11):msaf257, 2025.
2. Ciccone MF, Anandan D, Chatterjee D, Chen C, …, Dos Santos CO (20 co-authors). Loss of BPTF restores estrogen response and suppresses metastasis of mammary tumors. *Nat Commun.* 16(1):9168, 2025.
3. Ju XC, Lee SY, Ågren R, Machado LC, Xing J, Azama C, Roy MC, Endo T, Huttner W, **Siepel A**, Fukunaga I, Zeberg H, Pääbo S. The activity and expression of adenylosuccinate lyase were reduced during modern human evolution, affecting brain and behavior. *Proc Natl Acad Sci U S A.* 122(32):e2508540122, 2025.
4. Liu L, Zhao Y, Hassett R, Toneyan S, Koo PK, **Siepel A.** Probabilistic and machine-learning methods for predicting local rates of transcription elongation from nascent RNA sequencing data. *Nucleic Acids Res* 53(4):gkaf092, 2025.
5. Tajima Y, Vargas CDM, Ito K, Wang W, Luo J-D, Xing J, Kuru N, Machado LC, **Siepel A**, Carroll TS, Jarvis ED, Darnell RB. A humanized *NOVA1* splicing factor alters mouse vocal communications. *Nature Commun* 16(1):1542, 2025.
6. Berube B, Ernst E, Cahn J, Roche B, de Santis Alves C, Lynn J, Scheben A, Grimanelli D, **Siepel A**, Ross-Ibarra J, Kermicle J, Martienssen RA. Teosinte Pollen Drive guides maize diversification and domestication by RNAi. *Nature* 633(8029):380-388, 2024.
7. Serio RN, Scheben A, Lu B, Gargiulo DV, Patruno L, Buckholtz CL, Chaffee RJ, Jibilian MC, Persaud SG, Staklinski SJ, Hassett R, Brault LM, Ramazzotti D, Barbieri CE, **Siepel AC**, Nowak DG. Clonal Lineage Tracing with Somatic Delivery of Recordable Barcodes Reveals Migration Histories of Metastatic Prostate Cancer. *Cancer Discov* 14(10):1990-2009, 2024.
8. Ortiz JR, Lewis SM, Ciccone M, Chatterjee D, Henry S, **Siepel A**, Dos Santos CO. Single-Cell Transcription Mapping of Murine and Human Mammary Organoids Responses to Female Hormones. *J Mammary Gland Biol Neoplasia*. 29(1):3, 2024.

9. Kuderna LFK *et al*. Identification of constrained sequence elements across 239 primate genomes. *Nature* 625, 735–742, 2024.

10. Mo Z, **Siepel A**. Domain-adaptive neural networks improve supervised machine learning based on simulated population genetic data. *PLOS Genet.* 19, e1011032, 2023.

11. Scheben, A, Ramos OM, Kramer M, Goodwin S, Oppenheim S, Becker DJ, Schatz MC, Simmons NB, **Siepel A**, McCombie W R. Long-read sequencing reveals rapid evolution of immunity- and cancer-related genes in bats. *Genome Biol. Evol.* 15, evad148, 2023.

12. Zhao Y, Liu L, Hassett R, **Siepel A**. Model-based characterization of the equilibrium dynamics of transcription initiation and promoter-proximal pausing in human cells. *Nucleic Acids Res.* 51(21) e106, 2023. doi:10.1093/nar/gkad843.

13. Campagna L, Mo Z, **Siepel A**, Uy JAC. Selective sweeps on different pigmentation genes mediate convergent evolution of island melanism in two incipient bird species. *PLOS Genetics* 18 (11), e1010474, 2022

14. Dukler N, Mughal MR, Ramani R, Huang YF, **Siepel A**. Extreme purifying selection against point mutations in the human genome. *Nature Commun* 13 (1), 1-12, 2022

15. Wu Y, Johnson L, Song B, Romay C, Stitzer M, **Siepel A**, Buckler E, Scheben A. A multiple alignment workflow shows the effect of repeat masking and parameter tuning on alignment in plants. *The Plant Genome*, e20204, 2022

16. Zhao Y, Dukler N, Barshad G, Toneyan S, Danko CG, **Siepel A**. Deconvolution of expression for nascent RNA-sequencing data (DENR) highlights pre-RNA isoform diversity in human cells. *Bioinformatics* 37 (24), 4727-4736 doi: 10.1093/bioinformatics/btab582

17. Abuhashem A, Chivu AG, Zhao Y, Rice EJ, **Siepel A**, Danko CG, et. al. RNA Pol II pausing facilitates phased pluripotency transitions by buffering transcription. *Genes Dev* 2022 doi: 10.1101/gad.349565.122

18. Hejase HA, Mo Z, Campagna L, **Siepel A**. A deep-learning approach for inference of selective sweeps from the ancestral recombination graph. *Molecular Biology and Evolution*, 2022 Jan; 39 (1) msab332

19. Hutton ER, Vakoc CR, **Siepel A**. ACE: a probabilistic model for characterizing gene-level essentiality in CRISPR screens. *Genome Biol.,* 22:278, 2021

20. Henry S, Trousdell MC, Cyrill SL, Zhao Y, Feigman MJ, Bouhuis JM, Aylard DA, **Siepel A**, Dos Santos CO. Characterization of gene expression signatures for the identification of cellular heterogeneity in the developing mammary gland. *J Mammary Gland Biol Neoplasia*, 26:43-66, 2021.

21. Blumberg A, Zhao Y, Huang Y-F, Dukler N, Rice EJ, Krumholz K, Danko CG, **Siepel A**. Characterizing RNA stability genome-wide through combined analysis of PRO-seq and RNA-seq data. *BMC Biol.,* 19:30, 2021.

22. Hejase HA, Salman-Minkov A, Campagna L, Hubisz MJ, Lovette IJ, Gronau I, **Siepel A**. Genomic islands of differentiation in a rapid avian radiation have been driven by recent selective sweeps. *Proc Natl Acad Sci USA,* 117:30554-30565, 2020.

23. Hubisz MJ, Williams AL, **Siepel A**. Mapping gene flow between ancient hominins through demography-aware inference of the ancestral recombination graph. *PLOS Genet.* 16:e1008895, 2020.

24. Dukler N, Huang Y-F, **Siepel A**. Phylogenetic modeling of regulatory element turnover based on epigenomic data. *Mol Biol Evol.* 37:2137-2152, 2020.

25. Adrion JR, Cole CB, Dukler N, Galloway JC, Gladstein AL, …, **Siepel A** et al. (28 co-authors). A community-maintained standard library of population genetic models. *eLife* 9:e54967, 2020.

26. Joly-Lopez Z, Platts AE, Gulko B, Choi JY, Groen SC, Zhong X, **Siepel A**, Purugganan MD. An inferred fitness consequence map of the rice genome. *Nat Plants*. 6:119-130, 2020.

27. Huang Y-F, **Siepel A**. Estimation of allele-specific fitness effects across human protein-coding sequences and implications for disease. *Genome Res*. 29:1310-1321, 2019.

28. Ramani R, Krumholz K, Huang Y-F, **Siepel A**. PhastWeb: a web interface for evolutionary conservation scoring of multiple sequence alignments using phastCons and phyloP. *Bioinformatics*. 35:2320-2322, 2019.

29. Gulko B, **Siepel A**. An evolutionary framework for measuring epigenomic information and estimating cell-type-specific fitness consequences. *Nat Genet*. 51:335-342, 2019.

30. Danko CG, Choate LA, Marks BA, Rice EJ, Wang Z, …, **Siepel A.** (13 co-authors). Dynamic evolution of regulatory element ensembles in primate CD4+ T-cells. *Nat Ecol Evol.* 2:537-548, 2018.

31. 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*. 6:180-191.e4, 2018.

32. 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.* 28:52-65, 2018.

33. Dukler N, Booth GT, Huang Y-F, Tippens N, Waters CT, Danko CG, Lis JT, **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.

34. Kondo S, Vedanayagam J, Mohammed J, Eizadshenass S, Pang N, Kan L, Aradhya R, **Siepel**, Steinhauer J, Lai EC. New genes often acquire male-specific functions but rarely become essential in *Drosophila*. *Genes Dev*. 31:1841-1846, 2017.

35. Huang YF, Gulko B, **Siepel A**. Fast, scalable prediction of deleterious noncoding variants from functional and population genomic data. *Nat Genet.* 49:618–624, 2017.

36. Freedman AH, Schweizer RM, Ortega-Del Vecchyo D, Han E, …, **Siepel A** et al. (30 co-authors). Demographically-based evaluation of genomic regions under selection in domestic dogs. *PLOS Genet.* 12:e1005851, 2016.

37. Kuhlwilm M, Gronau I, Hubisz MJ, de Filippo C, Prado J, …, **Siepel A**, Castellano S. (22 co-authors). Ancient gene flow from early modern humans into Eastern Neanderthals. *Nature.* 530:429–433, 2016.

38. 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.* 24:4238–4251, 2015.

39. 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.* 12:433–438, 2015.

40. 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:e1005108, 2015.

41. 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.* 47:276–283, 2015.

42. 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.* 46:1311–1320, 2014.

43. 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*. 32:1250–1255, 2014.

44. Mohammed J, **Siepel A**, Lai EC. Diverse modes of evolutionary emergence and flux of conserved microRNA clusters. *RNA.* 20:1850–1863, 2014.

45. 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.* 26:2157–2166, 2014.

46. 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*. 20:1195–1209, 2014.

47. Wen J, Mohammed J, Bortolamiol-Becet D, Tsai H, Robine N, …, **Siepel A**, Lai EC (17 co-authors). Diversity of miRNAs, siRNAs, and piRNAs across 25 *Drosophila* cell lines. *Genome Res*. 24:1236–1250, 2014.

48. Arbiza L, Gottipati S, **Siepel A**, Keinan A. Contrasting X-linked and autosomal diversity across 14 human populations. *Am J Hum Genet.* 94:827–844, 2014.

49. Rasmussen MD, Hubisz MJ, Gronau I, **Siepel A**. Genome-wide inference of ancestral recombination graphs. *PLOS Genet.* 10:e1004342, 2014.

50. 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:e1004016, 2014.

51. Ma X, Kelley JL, Eilertson K, Musharoff S, Degenhardt JD, …, **Siepel A** et al. (11 co-authors). Population genomic analysis reveals a rich speciation and demographic history of orangutans (*Pongo pygmaeus* and *Pongo abelii*). *PLOS One.* 8:e77175, 2013.

52. Mohammed J, Flynt AS, **Siepel A**, Lai EC. The impact of age, biogenesis, and genomic clustering on *Drosophila* microRNA evolution. *RNA.* 19:1295–1308, 2013.

53. 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:e1003684, 2013.

54. 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.* 45:723–729, 2013.

55. Gronau I, Arbiza L, Mohammed J, **Siepel A.** Inference of natural selection from interspersed genomic elements based on polymorphism and divergence. *Mol Biol Evol.* 30:1159–1171, 2013.

56. 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.* 50:212–222, 2013.

57. Cornejo OE, Lefbure T, Bitar PD, Lang P, Richards VP, …, **Siepel A** et al. (14 co-authors). Evolutionary and population genomics of the cavity causing bacteria *Streptococcus mutans*. *Mol Biol Evol.* 30:881–893, 2013.

58. 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:e60465, 2013.

59. Choi SC, Rasmussen MD, Hubisz MJ, Gronau I, Stanhope MJ, **Siepel A**. Replacing and additive horizontal gene transfer in *Streptococcus*. *Mol Biol Evol.* 29:3309–3320, 2012.

60. Guertin MJ, Martins AL, **Siepel A**, Lis JT. Accurate prediction of inducible transcription factor binding intensities in vivo. *PLOS Genet.* 8:e1002610, 2012.

61. 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.* 29:1047–1057, 2012.

62. 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.* 478:476–482, 2011.

63. Gronau I, Hubisz MJ, Gulko B, Danko CG, **Siepel A**. Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet.* 43:1031–1034, 2011.

64. Lowe CB, Kellis M, **Siepel A**, Raney BJ, Clamp M, Salama SR, Kingsley DM, Lindblad-Toh K, Haussler D. Three periods of regulatory innovation during vertebrate evolution. *Science.* 333:1019–1024, 2011.

65. 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.* 43:741–743, 2011.

66. 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,* 145:622–634, 2011.

67. Hubisz MJ, Lin MF, Kellis M, **Siepel A**. Error and error mitigation in low-coverage genome assemblies. *PLOS One,* 6:e17034, 2011.

68. Yang J-S, Phillips MD, Betel D, Mu P, Ventura A, **Siepel AC**, Chen KC, Lai EC. Widespread regulatory activity of vertebrate microRNA* species. *RNA,* 17:312–326, 2011.

69. 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.

70. Orangutan Genome Sequencing and Analysis Consortium. Comparative and demographic analysis of orang-utan genomes. *Nature.* 469:529–533, 2011.

71. Hubisz MJ, Pollard KS, **Siepel A**. PHAST and RPHAST: Phylogenetic analysis with space/time models. *Brief. Bioinform.* 12:41–51, 2011.

72. Vinar T, Brejova B, Song G, **Siepel A**. Reconstructing histories of complex gene clusters on a phylogeny. *J Comput Biol.* 17:1267–1279, 2010.

73. 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:e1000451, 2010.

74. da Fonseca RR, Kosiol C, Vinar T, **Siepel A**, Nielsen R. Positive selection on apoptosis related genes. *FEBS Lett.* 584:469–476, 2010.

75. Pollard KS, Hubisz MJ, Rosenbloom KR, **Siepel A.** Detection of non-neutral substitution rates on Mammalian phylogenies. *Genome Res.* 20:110–121, 2010.

76. MGC Project Team. The completion of the Mammalian Gene Collection (MGC). *Genome Res.* 19:2324–2333, 2009.

77. Zhang Y, Song G, Vinar T, Green ED, **Siepel A**, Miller W. Evolutionary history reconstruction for Mammalian complex gene clusters. *J Comput Biol.* 16:1051–1070, 2009.

78. Holloway A, Begun D, **Siepel A**, Pollard KS. Accelerated sequence divergence of conserved genomic elements in *Drosophila melanogaster*. *Genome Res.* 18:1592–1601, 2008.

79. 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.* 180:391–408, 2008.

80. 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:e1000144, 2008.

81. 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.* 17:1797–1808, 2007.

82. **Siepel A**, Diekhans M, Brejova B, Langton L, Stevens M et al. (18 co-authors). Targeted discovery of novel human exons by comparative genomics. *Genome Res.* 17:1763–1773, 2007.

83. 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.* 27:5090–5104, 2007.

84. ENCODE Project Consortium. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447:799–816, 2007.
85. 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.* 17:760–774, 2007.
86. Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science.* 316:222–234, 2007.
87. Kuhn RM, Karolchick D, Zweig AS, Trumbower H, Thomas DJ, …, **Siepel A** et al. (25 co-authors). The UCSC genome browser database: update 2007. *Nucleic Acids Res.* 35:D668–D673, 2007.
88. Pollard KS, Salama SR, King B, Kern AD, Dreszer T, Katzman S, **Siepel A** et al. (13 co-authors). Forces shaping the fastest evolving regions in the human genome. *PLOS Genet.* 2:e168, 2006.
89. Pollard KS, Salama SR, Lambert N, Lambot M-A, Coppens S, …, **Siepel A** et al. (16 co-authors). An RNA gene expressed during cortical development evolved rapidly in humans. *Nature.* 443:167–172, 2006.
90. 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.* 441:87–90, 2006.
91. 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:e33, 2006.
92. Hinrichs AS, Karolchik D, Baertsch R, Barber GP, Bejerano G, …, **Siepel A** et al. (27 co-authors). The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* 34:D590–598, 2006.
93. **Siepel A**, Bejerano G, Pedersen JS, Hinrichs AS, Hou M et al. (16 co-authors). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 15:1034–1050, 2005.
94. Bejerano G, **Siepel AC**, Kent WJ, and Haussler D. Computational screening of conserved genomic DNA in search of functional noncoding elements. *Nat Methods*. 2:535–545, 2005.
95. Jojic V, Jojic N, Meek C, Geiger D, **Siepel A**, Haussler D, Heckerman D. Efficient approximations for learning phylogenetic HMM models from data. *Bioinformatics,* 20 Suppl.1:i161-168, 2004.
96. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature,* 432:695–716, 2004.

97. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science,* 306:636–640, 2004.
98. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature.* 431:931–945, 2004.
99. **Siepel A** and Haussler D. Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. *Mol Biol Evol.* 21:468–488, 2004.
100. **Siepel A** and Haussler D. Combining phylogenetic and hidden Markov models in biosequence analysis. *J Comput Biol.* 11:413–428, 2004.
101. Thomas JW, Touchman JW, Blakesley RW, Bouffard GG, Beckstrom-Sternberg SM, …, **Siepel AC** et al. (71 co-authors). Comparative analyses of multi-species sequences from targeted genomic regions. *Nature.* 424:788–793, 2003.
102. **Siepel AC**. An algorithm to enumerate sorting reversals for signed permutations. *J Comput Biol.* 10:575–597, 2003.
103. **Siepel AC**, Steenhuis TS, Rose CW, Parlange J-Y, McIsaac GF. A simplified hillslope erosion model with vegetation elements for practical applications. *J Hydrol.* 258:111–121, 2002.
104. **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.* 40:570–591, 2001.
105. **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.* 17:83–94, 2001.
106. Skupsk**i** MP, Booker M, Farmer A, Harpold M, Huang W, …, **Siepel A** et al. (18 co-authors). The Genome Sequence DataBase: towards an integrated functional genomics resource. *Nucleic Acids Res.* 27:35–38, 1999.
107. Harger C, Skupski M, Bingham J, Farmer A, Hoisie S, …, **Siepel A** et al. (21 co-authors). The Genome Sequence DataBase (GSDB): improving data quality and data access. *Nucleic Acids Res.* 26:21–26, 1998.
108. 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.* 25:18–23, 1997.
109. **Siepel AC**, Halpern AL, Macken C, Korber BT. A computer program designed to screen rapidly for HIV type 1 intersubtype recombinant sequences. *AIDS Res Hum Retroviruses.* 11:1413–1416, 1995.

Preprints

110. **Siepel A**, Hassett R, Staklinski SJ. Variational Inference with Node Embeddings (VINE) for Scalable Bayesian Phylogenetics. *bioRxiv* 2025.12.24.696405, 2025.

111. Staklinski SJ, Scheben A, Brault L, Hassett R, Serio R, Xing J, Nowak DG, **Siepel A**. Bayesian inference of tissue-migration histories in metastatic cancer from cell-lineage tracing data. *bioRxiv* 2025.09.09.668633, 2025

112. Harris M, Mo Z, **Siepel A**, Garud N. The persistence and loss of hard selective sweeps amid ancient human admixture. *bioRxiv* 2025.10.14.682443, 2025.

113. Stitzer MC, Seetharam AS, Scheben A, Hsu SK, Schulz AJ, AuBuchon-Elder TM, El-Walid M, Ferebee TH, Hale CO, La T, Liu ZY, McMorrow SJ, Minx P, Phillips AR, Syring ML, Wrightsman T, Zhai J, Pasquet R, McAllister CA, Malcomber ST, Traiperm P, Layton DJ, Zhong J, Costich DE, Dawe RK, Fengler K, Harris C, Irelan Z, Llaca V, Parakkal P, Zastrow-Hayes G, Woodhouse MR, Cannon EK, Portwood JL 2nd, Andorf CM, Albert PS, Birchler JA, **Siepel A**, Ross-Ibarra J, Romay MC, Kellogg EA, Buckler ES, Hufford MB. Extensive genome evolution distinguishes maize within a stable tribe of grasses. *bioRxiv* 2025.01.22.633974, 2025.

114. Staklinski SJ, Scheben A, **Siepel A**, Kilberg MS. Utility of AlphaMissense predictions in Asparagine Synthetase deficiency variant classification. *bioRxiv* 2023.10.30.564808, 2023.

115. **Siepel A**. A unified probabilistic modeling framework for eukaryotic transcription based on nascent RNA sequencing data. *bioRxiv* 2021.01.12.426408, 2021.

116. Mo Z, Scheben A, Steinberg J, **Siepel A**, Martienssen R. Circadian immunity, sunrise time and the seasonality of respiratory infections. *medRxiv* 2021.03.29.21254556, 2021.

Refereed Correspondence

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

Book Chapters, Review Articles, and Opinion Articles

118. Hejase HA, Dukler N, **Siepel A.** From summary statistics to gene trees: methods for inferring positive selection. *Trends Genet.* 36:243-258 [Review], 2020.

119. Hubisz MJ, **Siepel A**. Inference of Ancestral Recombination Graphs Using ARGweaver. In: Dutheil JY, ed. Statistical Population Genomics. Methods in Molecular Biology, vol 2090. Humana Press, New York, NY, p. 231-266 [Book Chapter], 2020.

120. **Siepel A**. Challenges in funding and developing genomic software: roots and remedies. *Genome Biol.* 20:147 [Opinion], 2019.

121. **Siepel A**, Arbiza L. *Cis*-regulatory elements and human evolution. *Curr Opin Genet Dev.* 29:81-89 [Review], 2014.

122. **Siepel A.** Phylogenomics of primates and their ancestral populations. *Genome Res.* 19:1929-1941 [Review], 2009.

123. **Siepel A.** Darwinian alchemy: human genes from noncoding. DNA. *Genome Res.* 19:1693-1695 [Review], 2009.

124. **Siepel A**, Haussler D. Phylogenetic Hidden Markov Models. In: Nielsen R, ed. Statistical Methods in Molecular Evolution. Statistics for Biology and Health. Springer, New York, NY, p. 325-351 [Book Chapter], 2005.

Refereed Conference Papers

125. Vinar T, Brejova B, Song G, **Siepel A.** Reconstructing histories of complex gene clusters on a phylogeny. In RECOMB-CG '09: Proceedings of the 7th international satellite workshop on comparative genomics. LNCS vol. 5817, Springer, 2009.
126. Zhang Y, Song GT, Vinar T, Green ED, **Siepel A**, Miller W. Reconstructing the evolutionary history of complex human gene clusters. In RECOMB '08: Proceedings of the twelfth annual international conference on research in computational molecular biology. LNCS vol. 4955. Springer-Verlag, 2008.
127. **Siepel A**, Pollard KS, Haussler D. New methods for detecting lineage-specific selection. In RECOMB '06: Proceedings of the tenth annual international conference on research in computational molecular biology. LNBI vol. 3909. Springer-Verlag, 2006.
128. **Siepel A** and Haussler D. Computational identification of evolutionarily conserved exons. In RECOMB '04: Proceedings of the eighth annual international conference on research in computational molecular biology. ACM Press, 2004.
129. **Siepel A** and Haussler D. Combining phylogenetic and hidden Markov models in biosequence analysis. In RECOMB '03: Proceedings of the seventh annual international conference on research in computational molecular biology. ACM Press, 2003.
130. **Siepel A**. An algorithm to enumerate all sorting reversals. In: RECOMB '02: Proceedings of the sixth international conference on computational molecular biology, 2002.
131. Moret BME, **Siepel AC**, Tang J, Liu T. Inversion Medians Outperform Breakpoint Medians in Phylogeny Reconstruction from Gene-Order Data. In: Guigo R, Gusfield D (eds). WABI '02: Proceedings of the second international workshop on algorithms in bioinformatics. Springer-Verlag, 2002.
132. Sobral BWS, Mangalam H, **Siepel A**, Mendes P, Pecherer R, McLaren G. Bioinformatics for rice resources. In: Goode JA, Chadwick D (eds). Proceedings of Novartis Foundation Symposium 236, "Rice biotechnology: improving yield, stress tolerance and grain quality". Wiley & Sons, Chichester, UK, 2001.
133. **Siepel, AC** & Moret, BME. Finding an Optimal Inversion Median: Experimental Results. In Gascuel O, Moret BME (eds). WABI '01: Proceedings of the first international workshop on algorithms in bioinformatics., LNCS vol. 2149, Springer, Berlin, Heidelberg, 2001.

TEACHING

Courses

Quantitative Biology I, II, and Genetics & Genomics, PhD Program, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 2014–2024.

Lead Instructor for qbio course, PhD Program, CSHL School of Biological Sciences, Cold Spring Harbor, NY, 2014-2018.

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

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

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

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

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

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

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

Trainees

First	Last	Position	Joint Mentor	Start	End
Shareef	Khalid	Grad Student, PhD	—	2025	Present
Jiawei	Xing	Postdoc	—	2024	Present
Xin	Zheng	Postdoc	—	2024	Present
Nurdan	Kuru	Postdoc	—	2024	Present
Stephen	Staklinski	Grad Student, PhD	—	2023	Present
Luiz	Machado de Oliveira	Grad Student, PhD	—	2022	Present
Mehreen	Mughal	Postdoc	—	2020	2022
Lingjie	Liu	Grad Student, PhD	—	2020	2024
Armin	Scheben	Postdoc	—	2019	2024
Ziyi	Mo	Grad Student, PhD	—	2019	2024
Hussein	Hijazi	Postdoc	—	2018	2020
Alexander	Xue	Postdoc	—	2018	2024
Yixin	Zhao	Postdoc	—	2018	2024
Amit	Blumberg	Postdoc	—	2016	2019
Hirak	Sarkar	Grad Student, Visiting	—	2016	2017
Yifei	Huang	Postdoc	—	2015	2019
Elizabeth	Hutton	Grad Student, PhD	—	2015	2020
Melissa	Hubisz	Grad Student, PhD	—	2014	2019
Noah	Dukler	Postdoc	—	2019	2020

		Grad Student, PhD		2014	2018
Lenore	Pipes	Grad Student, PhD	C. Mason	2012	2017
Joseph	Porter	Undergrad	—	2012	2012
Matthew	Rasmussen	Postdoc	—	2011	2013
Leonardo	Arbiza	Postdoc	A. Keinan	2010	2013
Sang Chul	Choi	Postdoc	—	2010	2013
Brad	Gulko	Grad Student, PhD	—	2010	2017
Renee	Setter	Undergrad	—	2010	2010
Omar	Cornejo	Postdoc	C.D. Bustamante	2009	2010
Charles	Danko	Postdoc	W.L. Kraus	2009	2014
Ilan	Gronau	Postdoc	—	2009	2014
Jaaved	Mohammed	Grad Student, PhD	E. Lai	2009	2016
Michael D.	Phillips	Grad Student, Masters	E. Lai	2008	2010
André L.	Martins	Grad Student, PhD	—	2007	2014
Daniel	Sussman	Undergrad	—	2007	2008
Bronislava	Brejova	Postdoc	—	2006	2008
Carolin	Kosiol	Postdoc	C.D. Bustamante	2006	2008
Tomas	Vinar	Postdoc	—	2006	2008
Adam	Diehl	Grad Student, Masters	—	2006	2010
Alexandra	Denby	Undergrad	—	2006	2008
Tytus	Mak	Undergrad	—	Spring 2006	Spring 2006
Alison	Marklein	Undergrad	—	Summer 2006	Summer 2006

Thesis Committees – Current

Diego Hernandez (Committee Chair), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory.

Vasilisa Kovaleva (Committee Chair), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory.

Dmitry Biba, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory.

David Zimmerman, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory.

Yijie Kang, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory.

Thesis Committees – Completed

Alexander Kwakye, Genetics, Stony Brook University, 2025.

Jose Salome-Correa, Biology, New York University, 2025.

Iacopo Gentile, Ph.D., (Academic Mentor), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2024.

Danielle Ciren, Ph.D., CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2023.

Jenelys Ruiz-Ortiz, Ph.D., (Committee Chair), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2023.

Elysia Cecilia Saputra, Ph.D., Computational Biology Program, Joint Carnegie Mellon University-University of Pittsburgh, 2023.

Jonathan Werner, Ph.D., (Academic Mentor), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2023.

Cole Wunderlich, Ph.D., (Academic Mentor), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2023.

Kathryn O'Neill, Ph.D., (Committee Chair), CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2022.

Laraib Malik, Ph.D., Computer Science, Stony Brook University, 2020.

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

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

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

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

Maria Nattestad, Ph.D., PhD Program, CSHL School of Biological Sciences, Cold Spring Harbor Laboratory, 2017.

Monica Ramstetter, Ph.D., Computational Biology, Cornell University, 2017.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

OTHER SERVICE

Special Advisor on Cancer AI, Cold Spring Harbor Laboratory Cancer Center, Cold Spring Harbor, NY, 2025–present.

Co-leader, Cancer Genetics and Genomics Program, Cold Spring Harbor Laboratory Cancer Center, Cold Spring Harbor, NY, 2020–2025.

Founding and lead organizer for ProbGen, Probabilistic Modeling in Genomics Conference, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (2015, 2018, 2021, 2023, 2025, 2027).

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

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

Board Member, NHGRI Board of Scientific Counselors (BSC), National Human Genome Research Institute, Bethesda, MD, 2020–2025.

Board Member, NCGR Board of Directors, National Center for Genome Resources, Santa Fe, NM, 2019–2024.

Co-chair (with Eimear Kenny then David Knowles), Population Genomics Working Group, New York Genome Center, New York, NY, 2018–2023.

Co-organizer, Annual New York Area Population Genomics Workshop, 2015–2019, 2023.

Advisory Board Member, School of Engineering Bioinformatics, New York University, New York, NY, 2015–2023.

US Regional Judge for the Blavatnik Awards for Young Scientists, The New York Academy of Sciences, New York, NY, 2021.

NIH Study Section for Genetic Variation and Evolution (GVE), 2020–2021.

NHGRI Strategic Planning Workshop: Perspectives in Comparative Genomics & Evolution, National Human Genome Research Institute, Bethesda, MD, August 2019.

Admissions Committee, Cold Spring Harbor Laboratory School of Biological Sciences, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 2015–2018.

Graduate Field Membership, Physiology, Biophysics and Systems Biology Program, Weill Cornell Medical College of Cornell University, New York, NY, 2014–2019.

NHGRI Strategic Planning Workshop: From Genome to Phenotype, National Human Genome Research Institute, Rockville, MD, January 2019.

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

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

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

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

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

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

Graduate Student Admissions Committee, Tri-Institutional MD-PhD Training Program in Computational Biology & Medicine, Cornell University, Ithaca, NY, 2006–2014.

Co-organizer, Biological Sequence Analysis and Probabilistic Models Workshop (with Anders Krogh, Gerton Lunter, Gilean McVean, Molly Przeworski et al.), 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, 2014).
- Computer Science (2007-2009, 2012).
- Molecular Biology and Genetics (2008, 2012).
- Department of Biological Statistics and Computational Biology (BSCB) / Weill Institute (2008).

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

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

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.

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

Editorial Board *Genome Research*, 2007–2010.

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.

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

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

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

Program Committee for the 24th International Conference on Machine Learning (ICML), 2007.

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

INVITED TALKS

Conference and Symposium Lectures

9th HGNYC, Human Genetics in NYC Conference [virtual], The Rockefeller University, New York, NY, July 2021.

Keynote Lecture, 11th RECOMB / ISCB Conference on Regulatory and Systems Genomics with DREAM Challenges, NYU Langone Health, New York, NY, December 2018.

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.

Keynote Lecture, ProbGen'17, 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.

ProbGen'16, 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, MCEB2014, Mathematical and Computational Evolutionary Biology Conference, Montpellier Laboratory of Computer Science, Robotics, and Microelectronics; Montpellier, France, June 2014.

11th $[BC]^2$, Basel Computational Biology Conference, 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, 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 Computational Molecular Biology Symposium, 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.

Seminars and Workshops

Department of Human Genetics, University of Chicago, Chicago, IL, October 2025.

Department of Human Genetics, McGill University, Montreal, Canada, July 2025.

UConn Health, Farmington, CT, February 2025.

Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, September 2024.

Cornell Center for Comparative and Population Genomics, Ithaca, NY, May 2024.

Flatiron Institute, Simons Foundation, New York, NY, October 2022.

RobFest: Seminar in Honor of Rob Martienssen. Cold Spring Harbor Laboratory, NY, May 2022

Lecture on Ancestral Recombination Graph inference [virtual], Harvard University, November 2021.

Evolutionary Genetics Talk, Genetics Society Seminar Series [virtual], University College London, UK, September 2021.

3rd Edition, Models and Inference in Population Genetics Workshop, University of Warwick, Coventry, UK, December 2019.

Program in Quantitative Genomics (PQG) Seminar Series, Harvard University, Boston, MA, November 2019.

Department of Genetics Seminar Series, Washington University School of Medicine, St. Louis, MO, September 2019.

Bioinformatics and Genomics Seminar Series, University of North Carolina at Charlotte, NC, March 2019.

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 of 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), Bellairs Research Institute of McGill University, 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 of 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, Montpellier, France, June 2014.

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

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

Penn 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 of Biological Sequence Analysis and Probabilistic Models Workshop, (with Katherine Pollard and Sean Eddy), Janelia Farm Research Campus, Howard Hughes Medical Institute (HHMI), 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, Bellairs Research Institute of McGill University, 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, Bellairs Research Institute of McGill University, 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.