Elizabeth Purdom, Ph.D.

Department of Statistics, UC Berkeley
367 Evans Hall #3860
Berkeley, CA 94720-3860

(510) 642-6154 (510) 642-7892
epurdom@stat.berkeley.edu
http://www.stat.berkeley.edu/~epurdom

Research Interests

My research interests lie in developing statistical methods for high-dimensional data arising in the field of biology and genetics. I focus on questions of robust estimation and hypothesis testing for high-throughput biological experiments, in particular sequencing technologies. My application areas include the areas of gene expression analysis, single cell sequencing, cancer biology, microbiome, methylation (bisulfite) sequencing, normalization of sequencing data, and integration of multiple 'omics data experiments. I am also interested in integration of heterogeneous sources of data, where the data can be multiple experimental platforms or, more generally, arbitrary forms of preexisting biological knowledge such as networks or trees. Statistically, I am interested in questions of high-dimensional inference and multivariate analysis – problems that arise naturally in trying to create a unified understanding of this type of data.

Education

08/2006 **Ph.D. Statistics, Stanford University**
Thesis title: *Multivariate Kernel Methods in the Analysis of Graphical Structures*
Advisor: Susan Holmes

05/2000 **B.S. Applied Mathematics and History, Yale University** (Double major)
Summa Cum Laude
Distinction in the major of Applied Mathematics
Distinction in the major of History

Post-Graduate Positions

06/2022 – present **Director** Center for Computational Biology, UC Berkeley

07/2016 – present **Associate Professor**
Department of Statistics, UC Berkeley

07/2009 – 06/2016 **Assistant Professor**
Department of Statistics, UC Berkeley

03/2007 – 07/2009 **NSF Post-doctoral Fellow**
Division of Biostatistics, UC Berkeley
Advisor: Terry Speed

08/2006 – 02/2007 **Post-Doctoral Scholar**
Mathematical Sciences Research Institute, Berkeley, CA
Advisor: Terry Speed

Awards and Honors

02/2022 Chan Zuckerberg Biohub Investigator

09/2013 Hellman Family Faculty Fund Award

09/2008 Stellar Abstract Award, Conference on Emerging Quantitative Issues in Parallel Sequencing, Harvard School of Public Health

Awards and Honors (continued)

08/2007	NSF Post-Doctoral Fellowship in Biological Informatics
06/2006	Departmental Teaching Award (Statistics Department, Stanford University)
06/2003	Departmental Teaching Award (Statistics Department, Stanford University)
05/2001	Gabilan Stanford Graduate Fellowship in Science and Engineering (Stanford University)

Grant and Funding Activites

Active Awards

07/2024-12/2025 *Post-Selection Inference Challenges in scRNA-Seq data from Patient Cohorts*
France-Berkeley Fund, UC Berkeley
Role: Co-PI
Award: $12,000

09/2022-08/2026 *A statistical framework for disease classification with scRNA-Seq data*
NIH R01 GM144493
Role: PI
Total Direct Costs: $1,868,963

07/2022 – 12/2023 *Benchmarking of scRNA-Seq from patient populations*
Data Insights Grant, Chan Zuckerberg Initiative
Role: PI
Award: $200,000

02/2022-01/2027 *Investigator Award*
CZ BioHub
Award: $1,000,000

Completed Awards

09/2017-06/2022 *A Comprehensive Whole-Brain Atlas of Cell Types in the Mouse*
NIH U19 MH114830
Role: Co-Investigator (PI: John Ngai)
Total Direct Costs: $908,808

Statistical Analysis and Comprehension of the Human Cell Atlas in R / Bioconductor: Access and Scalable Infrastructure
Chan Zuckerberg Inititive
Role: Collaborator (PI: Davide Risso)
Award: $45,000

08/2015-09/2021 *Epigenetic Control of Drought Response in Sorghum (EPICON)*
DOE DE-SC0014081
Role: Co-PI
Total Direct Costs: $4,690,531

09/2014-06/2017 *Classification of Cortical Neurons by Single Cell Transcriptomics*
NIH U01 MH105979
Role: Co-Investigator (PI: John Ngai)
Total Direct Costs: $2,778,372

Grant and Funding Activites (continued)

09/2016-06/2017	*UC Berkeley Presidential Chair Fellows – Curriculum Enrichment Grant* Role: Co-PI Award: $20,000
09/2010-08/2014	*SCREMS: Building a Statistical Computing Environment to Support Scientific Research* NSF DMS-1026441 Role: Co-PI Total Direct Costs: $101,213
09/2009-07/2014	*TCGA Data Analysis Center At Berkeley* NIH U24CA143799 Role: Collaborator (PI: Paul Spellman) Total Direct Costs: $724,564

Grant Review Panel and Other Activities

06/2022	Participant in NIH Review Panel (ASPA Study Section)
11/2016	Participant in NSF Review Panel
10/2015	Participant in NSF Genomes to Phenomes (G2P) Working Group

Research Publications

1. Roux de Bézieux, H., Street, K., Fischer, S., Van den Berge, K., Chance, R., Risso, D., Gillis, J., Ngai, J., Purdom, E., & Dudoit, S. (2024). Improving replicability in single-cell rna-seq cell type discovery with dune. *BMC Bioinformatics*, *25*, Article Article 198.

2. Wu, D., Hailer, A. A., Wang, S., Yuan, M., Chan, J., Kurdi, A. E., Rahim, M., Kondo, A., Han, D., Ali, H., D'Angio, B., Mayer, A., Klufas, D., Kim, E., Shain, A. H., Choi, J., Bhutani, T., Simpson, G., Grekin, R. C., … Cho, R. J. (2024). A single-cell atlas of IL-23 inhibition in cutaneous psoriasis distinguishes clinical response. *Science Immunology*, *9*(91), eadi2848.

3. Caddell, D. F., Pettinga, D., Louie, K., Bowen, B. P., Sievert, J. A., Hollingsworth, J., Rubanowitz, R., Dahlberg, J., Purdom, E., Northen, T., & Coleman-Derr, D. (2023). Drought shifts sorghum root metabolite and microbiome profiles and enriches for pipecolic acid. *Phytobiomes Journal*, *7*(4), 449–463.

4. Gao, C., Courty, P.-E., Varoquaux, N., Cole, B., Montoya, L., Xu, L., Purdom, E., Vogel, J., Hutmacher, R. B., Dahlberg, J. A., Coleman-Derr, D., Lemaux, P. G., & Taylor, J. W. (2023). Successional adaptive strategies revealed by correlating arbuscular mycorrhizal fungal abundance with host plant gene expression. *Molecular Ecology*, *32*(10), 2674–2687.

5. Van den Berge, K., Chou, H.-J., Kunda, D., Risso, D., Street, K., Purdom, E., Dudoit, S., Ngai, J., & Heavner, W. (2023). *A latent activated olfactory stem cell state revealed by single cell transcriptomic and epigenomic profiling.* [bioRxiv].

6. Wang, H., Torous, W., Gong, B., & Purdom, E. (2023b). *Visualizing scRNA-Seq data at population scale with GloScope.* [bioRxiv].

7. Wu, D., Hailer, A., Liu, Y., Wang, S., Purdom, E., North, J., Cho, R., & Cheng, J. (2023). Disruption of a pathologic, skin-resident T17 cell identity in clinically effective il23 blockade of psoriasis. *The Journal of investigative dermatology.*, *143*(5). https://doi.org/10.1016/j.jid.2023.03.1562

8. Cook, C. P., Taylor, M., Liu, Y., Schmidt, R., Sedgewick, A., Kim, E., Hailer, A., North, J. P., Harirchian, P., Wang, H., Kashem, S. W., Shou, Y., McCalmont, T. C., Benz, S. C., Choi, J., Purdom, E., Marson, A., Ramos, S. B. V., Cheng, J. B., & Cho, R. J. (2022). A single-cell transcriptional gradient in human cutaneous memory t cells restricts th17/tc17 identity. *Cell Reports Medicine*, *3*(8) doi: 10.1016/j.xcrm.2022.100715.

9. Gao, C., Xu, L., Montoya, L., Madera, M., Hollingsworth, J., Chen, L., Purdom, E., Singan, V., Vogel, J., Hutmacher, R. B., Dahlberg, J. A., Coleman-Derr, D., Lemaux, P. G., & Taylor, J. W. (2022). Co-occurrence networks reveal more complexity than community composition in resistance and resilience of microbial communities. *Nature Communications*, *13*, Article Article 3867, 3867.

10. Hailer, A., Liu, Y., Wang, H., Taylor, M., Cook, C., North, J., Mauro, T., Purdom, E., Cheng, J., & Cho, R. (2022). RashX: Immune single-cell transcriptional classification of human chronic inflammatory skin disease. *The Journal of investigative dermatology.*, *142*(8). https://doi.org/10.1016/j.jid.2022.05.876

11. Liu, Y., Wang, H., Cook, C., Taylor, M. A., North, J. P., Hailer, A., Shou, Y., Sadik, A., Kim, E., Purdom, E., Cheng, J. B., & Cho, R. J. (2022). Defining patient-level molecular heterogeneity in psoriasis vulgaris based on single-cell transcriptomics. *Frontiers in Immunology*, *13*.

12. Liu, Y., Wang, H., Taylor, M., Cook, C., Martínez-Berdeja, A., North, J. P., Harirchian, P., Hailer, A. A., Zhao, Z., Ghadially, R., Ricardo-Gonzalez, R. R., Grekin, R. C., Mauro, T. M., Kim, E., Choi, J., Purdom, E., Cho, R. J., & Cheng, J. B. (2022). Classification of human chronic inflammatory skin disease based on single-cell immune profiling. *Science Immunology*, *7*(70), eabl9165.

13. Taylor, M., Cook, C., Liu, Y., Schmidt, R., Hailer, A., North, J., Wang, H., Kashem, S., Purdom, E., Marson, A., Ramos, S., Cho, R., & Cheng, J. (2022). A single-cell transcriptional gradient in human cutaneous memory t cells suppresses pathogenic th17 inflammation. *The Journal of investigative dermatology.*, *142*(8). https://doi.org/10.1016/j.xcrm.2022.100715

14. Callaway, E. M., Dong, H.-W., Ecker, J. R., Hawrylycz, M. J., Huang, Z. J., Lein, E. S., Ngai, J., Osten, P., Ren, B., Tolias, A. S., White, O., Zeng, H., Zhuang, X., Ascoli, G. A., Behrens, M. M., Chun, J., Feng, G., Gee, J. C., Ghosh, S. S., … management, P. (2021). A multimodal cell census and atlas of the mammalian primary motor cortex. *Nature*, *598*(7879), 86–102.

15. Gong, B., Zhou, Y., & Purdom, E. (2021). Cobolt: Integrative analysis of multimodal single-cell sequencing data. *Genome Biology*, *22*, 351.

16. Hicks, S. C., Liu, R., Ni, Y., Purdom, E., & Risso, D. (2021). mbkmeans: Fast clustering for single cell data using mini-batch k-means. *PLoS Computational Biology*, *17*(1), e1008625.

17. Scavuzzo-Duggan, T., Varoquaux, N., Madera, M., Vogel, J. P., Dahlberg, J., Hutmacher, R., Belcher, M., Ortega, J., Coleman-Derr, D., Lemaux, P., Purdom, E., & Scheller, H. V. (2021). Cell wall compositions of sorghum bicolor leaves and roots remain relatively constant under drought conditions. *Frontiers in Plant Science*, *12*.

18. Xu, L., Dong, Z., Chiniquy, D., Pierroz, G., Deng, S., Gao, C., Diamond, S., Simmons, T., Wipf, H. M. .-., Caddell, D., Varoquaux, N., Madera, M. A., Hutmacher, R., Deutschbauer, A., Dahlberg, J. A., Guerinot, M. L., Purdom, E., Banfield, J. F., Taylor, J. W., … Coleman-Derr, D. (2021). Genome-resolved metagenomics reveals role of iron metabolism in drought-induced rhizosphere microbiome dynamics. *Nature Communications*, *12*, 3209.

19. Yao, Z., Liu, H., Xie, F., Fischer, S., Adkins, R. S., Aldridge, A. I., Ament, S. A., Bartlett, A., Behrens, M. M., Van den Berge, K., Bertagnolli, D., de Bézieux, H. R., Biancalani, T., Booeshaghi, A. S., Bravo, H. C., Casper, T., Colantuoni, C., Crabtree, J., Creasy, H., … Mukamel, E. A. (2021). A transcriptomic and epigenomic cell atlas of the mouse primary motor cortex. *Nature*, *598*(7879), 103–110.

20. Brann, D. H., Tsukahara, T., Weinreb, C., Lipovsek, M., Van Den Berge, K., Gong, B., Chance, R., Macaulay, I. C., Chou, H. J., Fletcher, R. B., Das, D., Street, K., De Bezieux, H. R., Choi, Y. G., Risso, D., Dudoit, S., Purdom, E., Mill, J., Hachem, R. A., … Datta, S. R. (2020). Non-neuronal expression of SARS-CoV-2 entry genes in the olfactory system suggests mechanisms underlying COVID-19-associated anosmia. *Science Advances*, *6*(31), 1–20.

21. Gao, C., Montoya, L., Xu, L., Madera, M., Hollingsworth, J., Purdom, E., Singan, V., Vogel, J., Hutmacher, R. B., Dahlberg, J. A., Coleman-Derr, D., Lemaux, P. G., & Taylor, J. W. (2020). Fungal community assembly in drought-stressed sorghum shows stochasticity, selection, and universal ecological dynamics. *Nature Communications*, *11*, 34.

22. Purdom, E. *Statistical methods for data science*. Online book. 2020.
https://epurdom.github.io/Stat131A/book/index.html

23. Varoquaux, N., & Purdom, E. (2020a). A pipeline to analyse time-course gene expression data. *F1000Research*, *9*(1447).

24. Cole, M. B., Risso, D., Wagner, A., DeTomaso, D., Ngai, J., Purdom, E., Dudoit, S., & Yosef, N. (2019). Performance Assessment and Selection of Normalization Procedures for Single-Cell RNA-Seq. *Cell Systems*, *8*(4), 315–328.e8.

25. El Karoui, N., & Purdom, E. (2019). The non-parametric bootstrap and spectral analysis in moderate and high-dimension. In K. Chaudhuri & M. Sugiyama (Eds.), *Proceedings of the twenty-second international conference on artificial intelligence and statistics* (pp. 2115–2124, Vol. 89). PMLR.
https://proceedings.mlr.press/v89/karoui19a.html

26. Gao, C., Montoya, L., Xu, L., Madera, M., Hollingsworth, J., Purdom, E., Hutmacher, R., Dahlberg, J., Coleman-Derr, D., Lemaux, P., & Taylor, J. (2019). Strong succession in arbuscular mycorrhizal fungal communities. *ISME Journal*, *13*(1).

27. Gong, B., & Purdom, E. (2019). MethCP: Differentially methylated region detection with change point models. In L. J. Cowen (Ed.), *Research in computational molecular biology* (pp. 68–84). Springer International Publishing.
https://doi.org/10.1007/978-3-030-17083-7_5

28. Harirchian, P., Lee, J., Hilz, S., Sedgewick, A. J., Perez White, B. E., Kesling, M. J., Mully, T., Golovato, J., Gray, M., Mauro, T. M., Purdom, E., Kim, E. A., Sbitany, H., Bhutani, T., Vaske, C. J., Benz, S. C., Cho, R. J., & Cheng, J. B. (2019). A20 and ABIN1 Suppression of a Keratinocyte Inflammatory Program with a Shared Single-Cell Expression Signature in Diverse Human Rashes. *Journal of Investigative Dermatology*, *139*(6), 1264–1273.

29. Varoquaux, N., Cole, B., Gao, C., Pierroz, G., Baker, C. R., Patel, D., Madera, M., Jeffers, T., Hollingsworth, J., Sievert, J., Yoshinaga, Y., Owiti, J. A., Singan, V. R., DeGraaf, S., Xu, L., Blow, M. J., Harrison, M. J., Visel, A., Jansson, C., … Purdom, E. (2019). Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(52), 27124–27132.

30. Cheng, J. B., Sedgewick, A. J., Finnegan, A. I., Harirchian, P., Lee, J., Kwon, S., Fassett, M. S., Golovato, J., Gray, M., Ghadially, R., Liao, W., Perez White, B. E., Mauro, T. M., Mully, T., Kim, E. A., Sbitany, H., Neuhaus, I. M., Grekin, R. C., Yu, S. S., … Cho, R. J. (2018). Transcriptional Programming of Normal and Inflamed Human Epidermis at Single-Cell Resolution. *Cell Reports*, *25*(4), 871–883.

31. Cho, R. J., Alexandrov, L. B., Den Breems, N. Y., Atanasova, V. S., Farshchian, M., Purdom, E., Nguyen, T. N., Coarfa, C., Rajapakshe, K., Prisco, M., Sahu, J., Tassone, P., Greenawalt, E. J., Collisson, E. A., Wu, W., Yao, H., Su, X., Guttmann-Gruber, C., Hofbauer, J. P., … South, A. P. (2018). APOBEC mutation drives early-onset squamous cell carcinomas in recessive dystrophic epidermolysis bullosa. *Science Translational Medicine*, *10*(455), 1–12.

32. El Karoui, N., & Purdom, E. (2018). Can we trust the bootstrap in high-dimensions? the case of linear models. *Journal of Machine Learning Research*, *19*(5), 1–66. http://jmlr.org/papers/v19/17-006.html

33. Risso, D., Purvis, L., Fletcher, R. B., Das, D., Ngai, J., Dudoit, S., & Purdom, E. (2018). clusterExperiment and RSEC: A Bioconductor package and framework for clustering of single-cell and other large gene expression datasets. *PLoS Computational Biology*, *14*(9), 1–16.

34. Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., Purdom, E., & Dudoit, S. (2018). Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics*, *19*(1), 477.

35. Fletcher, R. B., Das, D., Gadye, L., Street, K. N., Baudhuin, A., Wagner, A., Cole, M. B., Flores, Q., Choi, Y. G., Yosef, N., Purdom, E., Dudoit, S., Risso, D., & Ngai, J. (2017). Deconstructing Olfactory Stem Cell Trajectories at Single-Cell Resolution. *Cell Stem Cell*, *20*(6), 817–830.e8.

36. Gadye, L., Das, D., Sanchez, M. A., Street, K., Baudhuin, A., Wagner, A., Cole, M. B., Choi, Y. G., Yosef, N., Purdom, E., Dudoit, S., Risso, D., Ngai, J., & Fletcher, R. B. (2017). Injury Activates Transient Olfactory Stem Cell States with Diverse Lineage Capacities. *Cell Stem Cell*, *21*(6), 775–790.e9.

37. Goin, D., Smed, M., Pachter, L., Purdom, E., Nelson, J., Kjærgaard, H., Olsen, J., Hetland, M., Zoffmann, V., Ottesen, B., & Jawaheer, D. (2017). Pregnancy-induced gene expression changes in vivo among women with rheumatoid arthritis: A pilot study. *Arthritis Research and Therapy*, *19*, Article Article 104.

38. Johnson, M., & Purdom, E. (2017). Clustering of mRNA-Seq data based on alternative splicing patterns. *Biostatistics*, *18*(2), 295–307.

39. Naylor, D., Degraaf, S., Purdom, E., & Coleman-Derr, D. (2017). Drought and host selection influence bacterial community dynamics in the grass root microbiome. *ISME Journal*, *11*(12).

40. Perraudeau, F., Risso, D., Street, K., Purdom, E., & Dudoit, S. (2017). Bioconductor workflow for single-cell RNA sequencing: Normalization, dimensionality reduction, clustering, and lineage inference. *F1000Research*, *6*.

41. Ruddy, S., Johnson, M., & Purdom, E. (2016). Shrinkage of dispersion parameters in the binomial family, with application to differential exon skipping. *Annals of Applied Statistics*, *10*(2).

42. Sanborn, J., Chung, J., Purdom, E., Wang, N., Kakavand, H., Wilmott, J., Butler, T., Thompson, J., Mann, G., Haydu, L., Saw, R., Busam, K., Lo, R., Collisson, E., Hur, J., Spellman, P., Cleaver, J., Gray, J., Huh, N., … Cho, R. (2015). Phylogenetic analyses of melanoma reveal complex patterns of metastatic dissemination. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(35).

43. Zheng, C. L., Wang, N. J., Chung, J., Moslehi, H., Sanborn, J. Z., Hur, J. S., Collisson, E. A., Vemula, S. S., Naujokas, A., Chiotti, K. E., Cheng, J. B., Fassihi, H., Blumberg, A. J., Bailey, C. V., Fudem, G. M., Mihm, F. G., Cunningham, B. B., Neuhaus, I. M., Liao, W., … Cho, R. J. (2014). Transcription Restores DNA Repair to Heterochromatin, Determining Regional Mutation Rates in Cancer Genomes. *Cell Reports*, *9*(4), 1228–1234.

44. Purdom, E., Restall, C., Busuttil, R., Schluter, H., Boussioutas, A., Thompson, E., Anderson, R., Speed, T., & Haviv, I. (2013). Determining epithelial contribution to in vivo mesenchymal tumour expression signature using species-specific microarray profiling analysis of xenografts. *Genetics Research*, *95*(1).

45. Purdom, E., Ho, C., Grasso, C. S., Quist, M. J., Cho, R. J., Spellman, P., & Kelso, J. (2013). Methods and challenges in timing chromosomal abnormalities within cancer samples. *Bioinformatics*, *29*(24), 3113–3120. https://doi.org/10.1093/bioinformatics/btt546

46. Couvillion, M. T., Bounova, G., Purdom, E., Speed, T. P., & Collins, K. (2012). A Tetrahymena Piwi Bound to Mature tRNA 3' Fragments Activates the Exonuclease Xrn2 for RNA Processing in the Nucleus. *Molecular cell*, *48*(4), 509–520. http://dx.doi.org/10.1016/j.molcel.2012.09.010

47. Heiser, L., Sadanandam, A., Kuo, W.-L., Benz, S., Goldstein, T., Ng, S., Gibb, W., Wang, N., Ziyad, S., Tong, F., Bayani, N., Hu, Z., Billig, J., Dueregger, A., Lewis, S., Jakkula, L., Korkola, J., Durinck, S., Pepin, F., … Spellman, P. (2012). Subtype and pathway specific responses to anticancer compounds in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(8).

48. Purdom, E. (2012a). *Annotation-free estimates of gene-expression from mRNA-Seq* (tech. rep. No. 825). Department of Statistics, UC Berkeley. http://statistics.berkeley.edu/tech-reports/825

49. Bell, D., Berchuck, A., Birrer, M., Chien, J., Cramer, D., Dao, F., Dhir, R., Disaia, P., Gabra, H., Glenn, P., Godwin, A., Gross, J., Hartmann, L., Huang, M., Huntsman, D., Iacocca, M., Imielinski, M., Kalloger, S., Karlan, B., … Thomson, E. (2011). Integrated genomic analyses of ovarian carcinoma. *Nature*, *474*(7353).

50. Durinck, S., Ho, C., Wang, N. J., Liao, W., Jakkula, L. R., Collisson, E. A., Pons, J., Chan, S. W., Lam, E. T., Chu, C., Park, K., Hong, S. w., Hur, J. S., Huh, N., Neuhaus, I. M., Yu, S. S., Grekin, R. C., Mauro, T. M., Cleaver, J. E., … Cho, R. J. (2011). Temporal Dissection of Tumorigenesis in Primary Cancers. *Cancer Discovery*, *1*(2), 137–143. http://cancerdiscovery.aacrjournals.org/cgi/doi/10.1158/2159-8290.CD-11-0028

51. Durinck, S., Ho, C., Wang, N. J., Liao, W., Jakkula, L. R., Collisson, E. A., Pons, J., Chan, S.-W., Lam, E. T., Chu, C., Park, K., Hong, S.-w., Hur, J. S., Huh, N., Neuhaus, I. M., Yu, S. S., Grekin, R. C., Mauro, T. M., Cleaver, J. E., … Cho, R. J. (2011). Timing chromosomal abnormalities using mutation data. *Genome Biology*, *12*, P39.

52. Purdom, E. (2011). Analysis of a data matrix and a graph: Metagenomic data and the phylogenetic tree. *Annals of Applied Statistics*, *5*(4), 2326–2358.

53. Purdom, E., & Mukherjee, S. (2011). Transcriptomic Technologies and Statistical Data Analysis [Issue: 7]. In *Handbook of Statistical Systems Biology* (pp. 133–162). Wiley Online Library. http://onlinelibrary.wiley.com/doi/10.1002/9781119970606.ch7/summary

54. Wang, X. V., Verhaak, R. G. W., Purdom, E., Spellman, P. T., & Speed, T. P. (2011). Unifying gene expression measures from multiple platforms using factor analysis. *PLoS ONE*, *6*(3), e17691. https://doi.org/10.1371/journal.pone.0017691

55. Bullard, J. H., Purdom, E., Hansen, K. D., & Dudoit, S. (2010). Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC bioinformatics*, *11*(1), 94. http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-94

56. Lapuk, A., Marr, H., Jakkula, L., Pedro, H., Bhattacharya, S., Purdom, E., Hu, Z., Simpson, K., Pachter, L., Durinck, S., Wang, N., Parvin, B., Fontenay, G., Speed, T., Garbe, J., Stampfer, M., Bayandorian, H., Dorton, S., Clark, T. A., … Gray, J. W. (2010). Exon-Level Microarray Analyses Identify Alternative Splicing Programs in Breast Cancer. *Molecular Cancer Research*, *8*(7), 961–74.

57. Verhaak, R. G. W., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., Miller, C. R., Ding, L., Golub, T., Mesirov, J. P., Alexe, G., Lawrence, M., O'Kelly, M., Tamayo, P., Weir, B. A., Gabriel, S., Winckler, W., Gupta, S., Jakkula, L., … Network, C. G. A. R. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*, *17*(1), 98–110. https://doi.org/10.1016/j.ccr.2009.12.020

58. McLendon, R., Friedman, A., Bigner, D., Van Meir, E., Brat, D., Mastrogianakis, G., Olson, J., Mikkelsen, T., Lehman, N., Aldape, K., Yung, W., Bogler, O., Weinstein, J., VandenBerg, S., Berger, M., Prados, M., Muzny, D., Morgan, M., Scherer, S., … Thomson, E. (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature*, *455*(7216).

59. Purdom, E., Simpson, K. M., Robinson, M. D., Conboy, J. G., Lapuk, A. V., & Speed, T. P. (2008). FIRMA: A method for detection of alternative splicing from exon array data. *Bioinformatics*, *24*(15), 1707–1714. https://doi.org/10.1093/bioinformatics/btn284

60. Bik, E., Eckburg, P., Gill, S., Nelson, K., Purdom, E., Francois, F., Perez-Perez, G., Blaser, M., & Relman, D. (2006). Molecular analysis of the bacterial microbiota in the human stomach. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(3).

61. Purdom, E. (2006). *Multivariate Kernel Methods in the Analysis of Graphical Structures* [Doctoral dissertation, Stanford University].

62. Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the human intestinal microbial flora. *Science*, *308*(5728), 1635–8.

63. Purdom, E., & Holmes, S. P. (2005). Error distribution for gene expression data. *Statistical applications in genetics and molecular biology*, *4*(1), Article16. https://www.degruyter.com/document/doi/10.2202/1544-6115.1070/html

64. Xu, T., Shu, C.-T., Purdom, E., Dang, D., Ilsley, D., Guo, Y., Weber, J., Holmes, S. P., & Lee, P. P. (2004). Microarray analysis reveals differences in gene expression of circulating CD8(+) T cells in melanoma patients and healthy donors. *Cancer Research*, *64*(10), 3661–3667. https://doi.org/10.1158/0008-5472.CAN-03-3396

Software Development

- GloScope (2023): Visualizing scRNA-Seq Data at Population Scale
  (available on Bioconductor)
- Cobolt (2022): Joint analysis of multimodal single-cell sequencing data (python)
  (python package available on Github, `https://github.com/epurdom/cobolt`)
- moanin (2020): An R Package for Time Course RNASeq Data Analysis
  (available on Bioconductor)
- Dune (2020): Improving replicability in single-cell RNA-Seq cell type discovery
  (available on Bioconductor)
- mbkmeans: Mini-batch K-means Clustering for Single-Cell RNA-seq
  (available on BioConductor)
- methCP (2018): An R package for Differentially Methylated Region Detection
  (available on Github `https://github.com/boyinggong/MethCP`)
- Slingshot (2018): Lineage Inference for Single-cell Data
  (available on Bioconductor)
- clusterExperiment (2015): a package for clustering of single-cell sequencing data and other high-throughput experiments with large numbers of samples
  (available on Bioconductor)
- DoubleExpSeq (2014): Differential Exon Usage Test for RNA-Seq data via Empirical Bayes Shrinkage of the Dispersion Parameter
  (available on CRAN).
- cancerTiming (2012): Estimation of temporal ordering of cancer abnormalities
  (available on CRAN).
- FIRMA: A Method for Detection of Alternative Splicing from Exon Array Data
  (available in R package aroma.affymetrix, part of the aroma project (`http://www.aroma-project.org`)).

Research Presentations

Invited Presentations (Conferences and Workshops)

1. Analyzing variability in single-cell patient cohorts
   CZI Single-Cell Biology 2022 Annual Meeting
   San Jose, November 14th-18th, 2022

2. Harnessing Multimodal Single-Cell Sequencing Data for Integrative Analysis
   Invited Session, Joint Statistical Meetings
   Washington, DC, August 8, 2022

3. Single Cell mRNA-Seq: Subtype discovery and lineage estimation
   ISU BCB Symposium
   Iowa State, May 1, 2019

4. Single Cell mRNA-Seq: Subtype discovery and lineage estimation
   Keynote Speaker
   Journées Ouvertes Biologie, Informatique et Mathématiques (JOBIM2018)
   Marseilles, France, July 6, 2018

5. Lineage Estimation with Single-Cell mRNA-Seq Data
   Statistical Methods for post-genomic data (SMPG2018)
   Montpellier, France, January 12, 2018

6. Statistical Challenges in Analyzing High Dimensional Experiments in Molecular Biology: The example of estimating lineages in single-cell sequencing data
   Special afternoon of Data Science in ENS
   Paris, France, December 8, 2017

7. Lineage Estimation with Single-Cell mRNA-Seq Data
   Statistical and Computational Challenges in Large Scale Molecular Biology
   Bamff International Research Station, Mar 28, 2017

8. Lineage Estimation with Single-Cell mRNA-Seq Data
   DASHU 2017: Data Science and Computational Precision Health
   Feb 22, 2017

9. Expression Analysis of Tumors Based on Patterns in Alternative Splicing
   Workshop on Computational Cancer Biology
   Simons Institute, Berkeley, CA, Feb 1-5, 2016

10. Differential Gene Expression Analysis
    Asia Pacific Bioinformatics Conference
    South San Francisco, CA, January 10, 2016

11. Comparing Temporal Orderings of Chromosomal Abnormalities Using Mutation Data
    Colloque de Bioinformatique Moléculaire
    Institut Pasteur, Paris, France, October 7-8, 2015

12. Practices for Reproducible Research in Bioinformatics
    Panel Discussion, Opening Workshop, SAMSI Program on Beyond Bioinformatics
    Research Triangle Park, NC, September 8-12, 2014

13. Clustering mRNA-Seq Samples for Alternative Splicing Detection
    Invited Session, Joint Statistical Meetings (JSM)

Boston, August 2-7, 2014

14. Comparing Temporal Orderings of Chromosomal Abnormalities Using Mutation Data
Information Theory and Applications Workshop
San Diego, February 9-14, 2014

15. Timing Chromosomal Abnormalities in Tumors using Mutation Data
Invited Session, Joint Statistical Meetings (JSM)
Montreal, August 7, 2013

16. Measuring Subclonity in Tumors Using NGS
Invited Session, Annual Meeting of the Western North American Region of The International Biometric Society (WNAR)
Los Angeles, June 16-19, 2013

17. Timing Chromosomal Abnormalities in Tumors using Mutation Data
Biostatistics Symposium
San Francisco Bay Area Chapter, ASA
Stanford, June 1, 2013

18. Timing Chromosomal Abnormalities in Tumors using Mutation Data
Invited Session, Annual Meeting of the Western North American Region of The International Biometric Society (WNAR)
Fort Collins, June 17-20, 2012

19. Invited speaker, unable to attend due to birth of child
Transcriptomics and Epigenomics
Los Angeles, October 25-28, 2011

20. Statistical Methods for normalization and DE in mRNA-Seq experiments
Illumina Expert Panel on Sequencing
Toronto, July 22, 2010

21. Detecting alternative splicing with RNAseq: Some Thoughts on the Pipeline
Joint Statistical Meetings
Vancouver, July 31 – August 5, 2010

22. Connecting Microarray and Sequencing Experiments
Invited Session, Annual Meeting of the Western North American Region of The International Biometric Society (WNAR)
Seattle, June 20-23, 2010

23. Exon Array Data for Tumor Profiling Approaches for Analysis of mRNA-Seq on Tumor Samples
Genomic Aspects of Glioblastoma Multiforme
CEIT, San Sebastian, Spain. June 14-16, 2010

24. Data Analysis with Graphs
Massive Modern Data Sets
Stanford University, June 2008

25. Use of the Graph Laplacian to Analyze Network Data
Invited Session, Joint Statistical Meetings
Denver, August 5, 2008

26. Estimating Transcription Using Next Generation Sequencing

Statistical and Computational Challenges in Next-Generation Sequencing
MSRI, October 9, 2008

27. Introduction to Exon Arrays
Detection of Alternative Splicing Using Exon Arrays
Mathematical Systems Biology of Cancer II
MSRI, November, 2007

Invited Presentations (Seminars)

1. Analyzing single-cell sequencing data from patient cohorts (May 3, 2023)
Columbia Genomics Seminar
Columbia University

2. Analyzing single-cell sequencing data from patient cohorts (Jan 26, 2023)
Investigator Meeting
CZ Biohub, San Francisco, CA

3. Topics in the Analysis of High-Throughput Molecular Data and the Bootstrap in High Dimensions (Nov 11, 2019)
Department of Statistics
U. of Chicago

4. Analyzing Epigenetic Time Course Data for Drought Response (Mar 26, 2018)
MIA Sémenaire
Agro Paris Tech, Paris, France

5. Analysis of a High-Throughput Experiment: the example of single-cell sequencing (Feb 13, 2018)
Data Science Colloquium
Ecolé Normale Supérieure, Paris, France

6. Can We Trust the Bootstrap in (Moderately) High Dimensions? (Feb 5, 2018)
SMILE Workshop
Paris, France

7. Lineage Estimation with Single-Cell mRNA-Seq Data (Nov 30, 2017)
Bioinformatics/Genetics Seminar
Institut Cochin, Paris, France

8. Robust Strategies for the analysis of Single-Cell mRNA-Seq Data (October 20, 2017)
Laplace Data Science Reading Group
Ecolé Normale Supérieure, Paris France

9. Robust Strategies for Analysis of Single-Cell mRNA-Seq Data (Dec 1, 2016)
Statistics and Genomics Seminar
UC, Berkeley

10. Robust Strategies for Analysis of Single-Cell mRNA-Seq Data (Nov 10, 2016)
Biostatistics Seminar
Stanford University

11. Robust Strategies for Analysis of Single-Cell mRNA-Seq Data (Nov 2, 2016)
Visiting Lectures Series
Department of Bioinformatics and Computational Biology

MD Anderson Cancer Center

12. Robust strategies for analysis of single cell mRNA-Seq data (June 2, 2016)
    Science & Technology Seminar
    Joint Genome Institute

13. Comparing Temporal Orderings of Chromosomal Abnormalities Using Mutation Data (November 14, 2014)
    Department of Biostatistics and Computational Biology
    Dana-Farber Cancer Institute

14. Shrinkage of Dispersion Estimates for Analysis of Exon Usage in mRNA-Seq Experiments (May 13, 2014)
    Department of Statistics
    UC, Davis

15. Timing Chromosomal Abnormalities using Mutation Data (October 15, 2013)
    Department of Statistics
    UC, Riverside

16. Timing Chromosomal Abnormalities using Mutation Data (May 1, 2013)
    Department of Epidemology and Biostatistics
    UC, San Francisco

17. Timing Chromosomal Abnormalities using Mutation Data (July 26, 2012)
    Laboratoire Statistique et Génome, CNRS

18. Timing Chromosomal Abnormalities using Mutation Data (June 29, 2012)
    Institute Curie, Paris, France

19. Simplifying mRNA-Seq (October 28, 2010)
    Biostatistics Workshop
    Stanford

20. Normalizing mRNA-Seq Data (March 3, 2010)
    Mathematical and Computational Biology Seminar
    UC Berkeley

21. Simplifying mRNA-Seq (February 24, 2010)
    Statistics and Genomics Seminar
    UC, Berkeley

22. Estimating Transcription Using Next Generation Sequencing (November 19, 2009)
    Stanford Genome Technology Center

23. Statistical Problems in Estimating Alternative Splicing (March 19, 2009)
    Division of Biostatistics
    New York University

24. Statistical Problems in Estimating Alternative Splicing (March 18, 2009)
    Department of Biostatistics
    Columbia University

25. Statistical Problems in Estimating Alternative Splicing (March 9, 2009)
    Department of Bioinformatics
    Emory University

26. Statistical Problems in Estimating Alternative Splicing (February 26, 2009)
    The Broad Institute

27. Statistical Problems in Estimating Alternative Splicing (February 10, 2009)
    Program in Biostatistics and Biomathematics
    Fred Hutchinson Cancer Research Center

28. Statistical Problems in Estimating Alternative Splicing (February 6, 2009)
    Department of Statistics
    Columbia University

29. Statistical Problems in Estimating Alternative Splicing (February 2, 2009)
    Department of Bioinformatics and Computational Biology
    MD Anderson

30. Estimating Splicing Events Using Sequencing Data (December 5, 2008)
    Department of Statistics
    Stanford, University

31. Estimating Transcription Using Next Generation Sequencing (September 22, 2008)
    Correlated and High-Dimensional Data Seminar
    Biostatistics, Harvard Medical School

32. Detection of Alternative Splicing Using Exon Arrays (March 6, 2008)
    Statistics and Genomics Seminar
    UC, Berkeley

33. Data Analysis Using Graphical Structures (January 8, 2007)
    Applied Mathematics Seminar
    Institut National de la Recherche Agronomique, Jouy-en-Josas, France

34. A Look at Data Analysis with Graphs: By Way of Spectral Decompositions of Graphs (March 7, 2007)
    Neyman Seminar, Department of Statistics
    UC Berkeley

35. Data Analysis Using Graphical Structures (May 23, 2006)
    Department of Statistics
    Stanford University

36. Data Analysis Using Graphical Structures (February 15, 2006)
    Division of Biostatistics
    UC Berkeley.

37. Analysis of Ecological Data: Use of Phylogenetic Trees with Diversity Measurements (December 8, 2005)
    Statistics and Genomics Seminar
    UC Berkeley

Contributed Presentations

1. Clustering Tumors Based on Patterns in Alternative Splicing
   Beyond the Genome 2014
   Boston, October 8-10, 2014

2. Estimating Transcription Using Next Generation Sequencing

Emerging Quantitative Issues in Parallel Sequencing
Harvard Medical School, September 25, 2008

Posters Presented

1. Purdom E, Ho C, Grasso CS, Quist MJ, Cho RJ, and Spellman P (November 25, 2013) "Timing Chromosomal Abnormalities in Tumors Using Mutation Data" Impact of Large-Scale Genomic Data on Statistical and Quantitative Genetics, U. Washington, Seattle, WA
2. Purdom E., Ho C., Huang H, Durinck S., Spellman PT, and Cho, RJ. (June 20-25, 2011) "Timing Chromosomal Abnormalities using Mutation Data" Changing Landscape of the Cancer Genome, Boston, MA
3. Purdom E., Bullard J., Hansen K., and Dudoit S. (October 6, 2009) "Improving Differential Expression Detection with mRNA-Seq" 12th International MGED Meeting, Phoenix, Arizona.
4. Purdom E. and Holmes S. (June 2005) "Multivariate Analysis of Data Collected on Species" Mathematics of Evolution and Phylogeny Conference, Paris, France.
5. Dethlefsen L, Purdom E, Eckburg PB, Bik EM, Bernstein CN, Sargent M, Gill SR, Nelson KE, and Relman DA (April 2005) "Comparisons of 16S rDNA clone libraries using multiple methods to investigate within- and between-subject variability of the colonic bacterial communities in 3 human individuals" ASM Beneficial Microbes Conference, Lake Tahoe, Nevada.
6. Purdom, E. and Holmes, S (August 2004) "Asymmetric Laplace Distribution for Gene Expression Data" BIRS Statistical Science for Genome Biology Workshop, Banff, Canada.

Service and Leadership

Department of Statistics

2023-24 (Co-Chair)	**Faculty Hiring Committee** Oversight of open-rank search for a faculty position in area of statistics, broadly
2018-23 2016-17	**Head Graduate Advisor (HGA)** Oversight of the PhD Graduate Program, including advising first year students, deciding on students for awards and fellowships, and general oversight of students meeting milestones. During my time as HGA, I developed a new mentoring system for our PhD program, initiated a new doctoral seminar for the first year students, revised our requirements for the first year program, and updated our PhD policies, and updated our funding strategy to be more competitive.
2018-20 (Chair) 2016-17 (Chair) 2014-16 (Member)	**PhD Admissions Committee** The main duties of this committee are the PhD Admissions and recruitment. We receive roughly 300 applications a year and admit around 20 students, and the faculty committee reviews all these applications. As chair, I would additionally coordinate the review process and oversee the recruitment of the admitted students
Summer 2017	**Ad hoc committee for the development of new data science major**
Fall 2015	**Ad hoc "Next Decade Committee"** Committee task to prepare for 10-year departmental review
2013-14	**MA Admissions Committee**

Service and Leadership (continued)

2012-13 **Development & Alumni relations**
This committee created the department's Industrial Alliance Program, organizing its inaugural meeting, and developing our alumni connections.

2011-12 **Course Curriculum Committee**
The committee was tasked with collecting and analyzing our past enrollment and GSI allocations to determine appropriate guidelines for future allocation, especially in the presence of huge growth in our undergraduate major. I also spearheaded our effort to support faculty in introducing new undergraduate course offerings, resulting in two new courses being proposed for the 2012-2013 academic year (one of which was proposed by myself). Both courses (STAT 154 and STAT 158) have been approved by COCI and are now part of our standard offering.

Center for Computational Biology (CCB)

2022-present **Director**
The Director of CCB is the faculty member who oversees the administration of CCB, with similar duties as a Chair of a Department, including overseeing staff, CCB budget, assignment of committee and teaching duties, and personnel cases. As Director, I was also included with the heads of units in the College of CDSS in leadership meetings and strategic planning with the Dean and CDSS administration.

2015-2022 **Executive Committee**
This committee makes strategic decisions for CCB and advises the Director of the CCB in the running of the center

2019-2022 **Program Chair & Head Graduate Advisor**
Oversee both the PhD and Designated Emphasis (PhD minor) graduate programs that are run through CCB. This involved biweekly meetings with the staff and Director of CCB, as well as advising and approving graduate student curriculum affairs

2018-2022 **Curriculum Committee**
This committee oversees decisions regarding curriculum for the graduate programs offered by CCB

2016-2017 (Chair)
2015-2016 **PhD Admissions Committee**

2016-2017 **PhD Program Advisor**

Campus or College Level

2022-2024 **Committee on Faculty Welfare (FWEL)**
UC Berkeley Senate Committee that works with administration on issues of both regular and emeriti faculty welfare.

2020-21 **CDSS Research Advisory Committee**
Ad hoc research advisory committee to help inform Associate Provost of the Division of Computing, Data Science and Society (CDSS)

Spring 2019 **University Fellowship Committee, Physical Sciences & Engineering section**
Campus-wide committee charged with reviewing and ranking nominations of admitted graduate students for University multi-year fellowships (Berkeley Fellowships & Chancellor's Fellowships).

Service and Leadership (continued)

Spring 2013	**RNA-Seq Data Analysis Workshop (Organizer)** Computational Genomics Resource Laboratory, UC Berkeley
Spring 2011	**Advisory committee for planning of the Computational Genomics Resource Laboratory (CGRL)**
2010-11	**Berkeley Weekly Sequencing Group (Organizer)**

Editorial Activities

Associate Editor

2009 – 2011	Statistical Applications in Genetics and Molecular Biology (www.bepress.com/sagmb)

Referee for the following journals:

Annals of Applied Statistics
Bioinformatics
BMC Genomics
Biometrics
Biostatistics
Computational Statistics and Data Analysis
Genetics
Genome Biology
Genome Medicine
Journal of the American Statistical Association
Journal of the Royal Statistical Society
Journal of Statistical Computation and Simulation
Nature Biotechnology
Nature Communications
Nature Scientific Reports
Nucleic Acid Research
Statistical Applications in Genetics and Molecular Biology

Conference Organizing

August 2-7, 2014	"Statistical Problems in Cancer Genomics" (Session Organizer) Joint Statistical Meetings (JSM), Boston
June 16-19, 2013	"Cancer Genomics" (Session Organizer) Annual Meeting of the Western North American Region of The International Biometric Society (WNAR), Los Angeles
June 20-23, 2010	"Statistical Methods for High-throughput Sequencing" (Session Organizer) Annual Meeting of the Western North American Region of The International Biometric Society (WNAR), Seattle, Washington
November 2007	"Mathematical Systems of Biology of Cancer II" (Conference Organizer) Mathematical Sciences Research Institute, Berkeley, CA

Teaching

Primary instructor in the following courses in the Department of Statistics, UC Berkeley:

STAT 201B	Introduction to Statistics at an Advanced Level (Graduate) (Fall 2020, Fall 2023)
STAT 28 / STAT 131A	Statistical Methods for Data Science (Undergraduate) (Spring 2017, Fall 2019, Spring 2023)
STAT 272	Statistical Consulting (Graduate) (Fall 2011, Fall 2013, Fall 2015, Fall 2016, Fall 2018)
STAT 215B	Applied Statistics (Graduate) (Spring 2010, Spring 2011, Spring 2020, Spring 2021, Spring 2022)
STAT 152	Survey Sampling (Undergraduate) (Spring 2011, Spring 2014, Spring 2019)
STAT 158	Design and Analysis of Experiments (Undergraduate) (Spring 2013, Fall 2013, Spring 2015, Spring 2016)
Stat 151A	Linear Models (Undergraduate) (Spring 2014)

A teaching assistant and co-Instructor as a graduate student for the following courses in the Department of Statistics, Stanford University:

Stat 60	Introduction to Statistical Methods (TA)
Stat 141	Biostatistics (co-Instructor, TA)
Stat 206	Multivariate Statistics (TA)
	Survival Analysis (TA)
Stat 305a	Linear Models (TA)
Stat 390	Consulting (TA)

Student Advising

Post-Doctoral Researchers

- Nelle Veroquaux (2016 – 2019)
  Department of Statistics and BIDS Fellow
  Project: Integrative analysis of epigenetic 'omics data in Sorghum (part of the EPICON project)
- Sean Ruddy (2015 – 2016)
  Department of Statistics
  Project: Normalization of mRNA-Seq data
- Gergana Bounova (2011)
  (co-advised with Terry Speed and Sandrine Dudoit)
  Department of Statistics, UC Berkeley
  Project: Analysis of data produced from sequenced sRNAs in Tetrahymena (collaboration with Kathleen Collins)

PhD Students

- Yilong Hao (2024 – present)
  Division of Biostatistics, UC Berkeley

Student Advising (continued)

- Van Hovenga (2024 – present)
  Department of Statistics, UC Berkeley
  (co-advised with Iain Carmichael, UNC-Chapel Hill)
- Zhiwei Xiao (2024 – present)
  Department of Statistics, UC Berkeley
- Keying (Maggie) Kuang (2023 – present)
  Division of Biostatistics, UC Berkeley
- William Torous (2022 – present)
  Department of Statistics, UC Berkeley
- Florica Constantine (2022 – present)
  Department of Statistics, UC Berkeley
- Hao Wong (2021-present)
  Division of Biostatistics, UC Berkeley
- Boying Gong (2017-2021)
  Division of Biostatistics, UC Berkeley
  Thesis Title: Statistical Analysis and Integration of Multi-modal Sequencing Data
- Stephanie DeGraaf (2016 – 2020)
  Department of Statistics, UC Berkeley
  Thesis Title: Time-Course Analysis and Clustering of Gene Expression Data Time-Course Analysis and Clustering of Gene Expression Data
- Marla Johnson (2010 – 2017)
  Division of Biostatistics, UC Berkeley
  Thesis Title: Clustering of mRNA-Seq Data for Detection of Alternative Splicing Patterns
- Christine Ho (2010 – 2016)
  (co-advised with Haiyan Huang)
  Department of Statistics, UC Berkeley
  Thesis Title: Statistical modeling and analysis for biomedical applications.
- Sean Ruddy (2010 – 2014)
  Department of Statistics, UC Berkeley
  Thesis Title: Shrinkage of dispersion parameters in the double exponential family of distributions with applications to genomic sequencing

MA Students

- Yilong Hao (2023 - 2024)
  Department of Statistics, UC Berkeley
- Keying (Maggie) Kuang (2022 – 2023)
  Division of Biostatistics, UC Berkeley
- Hao Wang (2020 – 2021)
  Division of Biostatistics, UC Berkeley
  Project: Analysis of patient-level single-cell sequencing data
- Namita Trikannad (Summer 2019 – Summer 2020)
  Division of Biostatistics, UC Berkeley
  Project: Analysis of gene expression time course data for drought effects

Student Advising (continued)

- Boying Gong (2016-2017)
  Department of Statistics, UC Berkeley
  Project: Analysis of differential methylation in plants using bisulfite sequencing data.
- Jonathan Ackerman (Fall 2016)
  Department of Statistics, UC Berkeley
  Project: Evaluation of Normalization of microbiome data

Undergraduate and Other Advising

- Ryan Nayebi (Spring 2024)
  Undergraduate, UC Berkeley
- Daisy Wang (Spring 2024)
  Undergraduate, UC Berkeley
- Arya Haghighi (Fall 2023)
  Undergraduate, UC Berkeley
- Yiyan Hao (2023-24)
  Undergraduate, UC Berkeley
- Chae Yeon Lee (Spring 2023)
  Undergraduate, UC Berkeley
- Zhe Chen (Summer 2019)
  Visiting Undergraduate from Southern University of Science and Technology, China
  Project: Comparison of doublet-finding methods
- Liam Purvis (Summer 2016)
  Undergraduate, UC Berkeley
  Project: Writing Shiny app for clusterExperiment package
- Tae-Kyung (Bryan) Jung (Independent Study, Spring 2010)
  Undergraduate, UC Berkeley
  Project: "Estimating the Number of Medical Data Breaches"
- Eric Chi (Summer 2009)
  Visiting student, LBL; PhD student, Rice University

Dissertation and Examination Committees

PhD Dissertation Committee

- Jinyan Zhao, Civil & Environmental Eng. (2023)
- Boying Gong, Biostatistics, UC Berkeley (Chair, 2021)
- Stephanie DeGraaf, UC Berkeley (Chair, 2020)
- Hector Roux de Bezieux, Biostatistics, UC Berkeley (201)
- Rebecca Sarto Basso, Industrial Eng & Operation Research, UC Berkeley (2020)
- Magali Jaillard Dancette, University of Lyon
- Amy Ko, Computational Biology, UC Berkeley
- Kelly Street, Biostatistics, UC Berkeley
- Fanny Perraudeau, Biostatistics, UC Berkeley
- Marla Johnson, Biostatistics, UC Berkeley

Student Advising (continued)

- Taryn M. Sumabat, Molecular and Cell Biology, UC Berkeley
- Marina P. Volegova, Molecular and Cell Biology, UC Berkeley
- Christine Ho, Statistics, UC Berkeley
- Sean Ruddy, Statistics, UC Berkeley

PhD Qualifying Examination Committee

- William Torous, Statistics (May 6, 2024)
- Kat Flug, MCB (April 19, 2024)
- Corrine Elliot, Statistics (Dec 5, 2023)
- Kushal Nimkar, Chemical Engineering (Sep 15, 2023)
- Shuni Li, Statistics (Aug 16, 2022)
- Hao Wang, Biostatistics (May 3, 2022)
- Wenxin Zhang, Biostatistics (Aug 13, 2021)
- Jinyan Zhao, Civil & Environmental Engineering (May 11, 2021)
- Yun Zhou, Statistics, UC Berkeley (May 8, 2019)
- Hector Roux de Bezieu, Biostatistics (Chair, Nov 2019)
- Stephanie DeGraaf, Statistics, UC Berkeley (Nov 14, 2018)
- Boying Gong, Biostatistics, UC Berkeley (Oct 30, 2018)
- Shinya Iguchi, MCB, UC Berkeley (May 10, 2017)
- Jonathan R. Fischer, Statistics, UC Berkeley (Dec 2, 2016)
- David M. DeTomaso, Computational Biology, UC Berkeley (Oct 2016)
- Kelly Street, Biostatistics, UC Berkeley (Sep 12, 2016)
- Fanny Perraudeau, Biostatistics, UC Berkeley (May 5, 2016)
- Amy Ko, Computational Biology, UC Berkeley (June 17, 2015)
- Marla Johnson, Biostatistics, UC Berkeley (Dec 11, 2014)
- Diya Das, Molecular and Cell Biology, UC Berkeley (May 16, 2014)
- Marcus Stoiber, Biostatistics, UC Berkeley (March 17, 2014)
- Angie Zhu, Statistics, UC Berkeley (Chair, Dec 17, 2013)
- Christine Ho, Statistics, UC Berkeley (May 10, 2012)
- Sean Ruddy, Statistics, UC Berkeley (April 20, 2012)
- Tessa Lorraine Childers-Day, Statistics, UC Berkeley (Chair, Nov 19, 2010)
- Daisy (Yan) Huang, Statistics, UC Berkeley (Jan 20, 2010).
- Cagla Meral, Civil and Environmental Engineering, School of Engineering, UC Berkeley (January 15, 2010)

Master's Thesis Committee

- Yilong Hao, Statistics, Spring 2024
- Keying (Maggie) Kuang, Biostatistics, Spring 2023
- Hao Wang, Biostatistics, Spring 2021
- Aidan McLoughlin, Biostatistics, Spring 2021
- Philippe Boileau, Biostatistics, Spring 2020
- Hector Roux de Bezieux, Biostatistics, Fall 2018

Student Advising (continued)

- Steven Pollack, Biostatistics, Spring 2014
- Nathan P. Boley, Biostatistics, Fall 2013
- Danielle Champney, Statistics, Spring 2010