UCLA Computational Genomics Summer Institute

2016 Program

Visit the CGSI Website for an up-to-date archive of program videos, slides, papers, and more:
http://computationalgenomics.bioinformatics.ucla.edu/

Week One

Monday, July 18

09:00-09:15	**Introduction to CGSI 2016**
09:15-10:00	**Tutorial: Brian Browning** An Introduction to Genotype Imputation [1] Marchini J, Howie B. (2010). Genotype imputation for genome-wide association studies. *Nat Rev Genet.* 2010 Jul;11(7):499-511. DOI:10.1038/nrg2796. [2] Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing.*Nat Genet.* 2012 Jul 22;44(8):955-9. PMCID: PMC3696580 DOI: 10.1038/ng.2354. [3] Browning BL, Browning SR (2016). Genotype Imputation with Millions of Reference Samples. *Am J Hum Genet.* 2016 Jan 7;98(1):116-26. PMCID: PMC4716681 DOI: 10.1016/j.ajhg.2015.11.020
10:00-10:45	**Tutorial: William Wen** Bayesian Statistics and its Application to Integrative Statistical Genomics

	[1] Stephens, M., & Balding, D. J. (2009). Bayesian statistical methods for genetic association studies. *Nature Reviews Genetics*, 10(10), 681-690. doi: 10.1038/nrg2615. Review. PMID: 19763151. [2] Wakefield, J. (2009). Bayes factors for genome-wide association studies: comparison with P-values. *Genetic Epidemiology*, 33(1), 79-86. doi: 10.1002/gepi.20359. PMID:18642345. [3] Wen, X., Lee, Y., Luca, F., & Pique-Regi, R. (2016).Efficient Integrative Multi-SNP Association Analysis via Deterministic Approximation of Posteriors. *The American Journal of Human Genetics*, 98(6), 1114-1129. doi: 10.1016/j.ajhg.2016.03.029. PMID: 27236919. [4] Wen, X. Molecular QTL Discovery Incorporating Genomic Annotations using Bayesian False Discovery Rate Control. *Annals of Applied Statistics* (in press).
10:45-11:15	Coffee Break
11:15-12:00	Journal Club
12:00-14:00	Lunch Break
14:00-14:45	**Research Talk: Leonid Kruglyak** Genetic Basis of Complex Traits [1] Sadhu, M. J., Bloom, J. S., Day, L., & Kruglyak, L. (2016). CRISPR-directed mitotic recombination enables genetic mapping without crosses. *Science*, 352(6289), 1113-1116. [2] Bloom, J. S., Kotenko, I., Sadhu, M. J., Treusch, S., Albert, F. W., & Kruglyak, L. (2015). Genetic interactions contribute less than additive effects to quantitative trait variation in yeast. *Nature communications*, 6. [3] Albert, F. W., & Kruglyak, L. (2015). The role of regulatory variation in complex traits and disease. *Nature Reviews Genetics*, 16(4), 197-212. [4] Albert, F. W., Treusch, S., Shockley, A. H., Bloom, J. S., & Kruglyak, L. (2014). Genetics of single-cell protein abundance variation in large yeast populations. *Nature*, 506(7489), 494. [5] Bloom, J. S., Ehrenreich, I. M., Loo, W. T., Lite, T. L. V., & Kruglyak, L. (2013). Finding the sources of missing heritability in a yeast cross. *Nature*, 494(7436), 234-237.
14:45-15:15	Coffee Break
15:15-16:15	**Research Talk: Carlos Bustamante** Interpreting Human Variation in the Personal Genome Era

[1] Shringarpure, S. S., & Bustamante, C. D. (2015). Privacy risks from genomic data-sharing beacons. *The American Journal of Human Genetics*, 97(5), 631-646.
[2] Mendez, F. L., Poznik, G. D., Castellano, S., & Bustamante, C. D. (2016). The divergence of Neandertal and modern human Y chromosomes. *The American Journal of Human Genetics*, 98(4), 728-734.

Tuesday, July 19

09:15-10:00	**Tutorial: Alexander Schönhuth** Snakemake: Reproducible and Scalable Data Analysis [1] Köster, J., & Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*,*28*(19), 2520-2522. [2] Köster, J. (2014). Parallelization, Scalability, and Reproducibility in Next-Generation Sequencing Analysis, PhD thesis, TU Dortmund.
10:00-10:45	**Tutorial: Jo Hardin** Tutorial on RNASeq Normalization and Differential Expression [1] Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57-63. [2] 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), 1. [3] Dillies, M. A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., … & Guernec, G. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis.*Briefings in Bioinformatics*, *14*(6), 671-683. [4] Lovén, J., Orlando, D. A., Sigova, A. A., Lin, C. Y., Rahl, P. B., Burge, C. B., … & Young, R. A. (2012). Revisiting global gene expression analysis. *Cell*, *151*(3), 476-482.
10:45-11:15	Coffee Break
11:15-12:00	Journal Club
12:00-14:00	Lunch Break
14:00-14:45	**Tutorial: Sohini Ramachandran** Genomic Reconstructions of Deep Human History

[1] Palacios, J. A., Wakeley, J., & Ramachandran, S. (2015).Bayesian Nonparametric Inference of Population Size Changes from Sequential Genealogies. *Genetics*, genetics-115.
[2] Rasmussen, M. D., Hubisz, M. J., Gronau, I., & Siepel, A. (2014). Genome-wide inference of ancestral recombination graphs. *PLoS Genet*, 10(5), e1004342.
[3] Li, H., & Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. *Nature*, 475(7357), 493-496.

14:45-15:15 Coffee Break

15:15-16:00 **Tutorial: Alex Zelikovski**

High-Throughput Sequencing Applications to Molecular Epidemiology

[1] Wertheim, J. O., Brown, A. J. L., Hepler, N. L., Mehta, S. R., Richman, D. D., Smith, D. M., & Pond, S. L. K. (2014).The global transmission network of HIV-1. *Journal of Infectious Diseases*, 209(2), 304-313.
[2] Jombart, Thibaut, et al. (2014). Bayesian reconstruction of disease outbreaks by combining epidemiologic and genomic data. *PLoS Computational Biology* 10(1), e1003457.
[3] Artyomenko, A., Wu, N. C., Mangul, S., Eskin, E., Sun, R., & Zelikovsky, A. (2016). Long single-molecule reads can resolve the complexity of the Influenza virus composed of rare, closely related mutant variants. In International Conference on Research in Computational Molecular Biology (pp. 164-175). Springer International Publishing.
[4] Skums, P., Artyomenko, A., Glebova, O., Ramachandran, S., Mandoiu, I., Campo, D. S., … & Khudyakov, Y. (2014). Computational framework for next-generation sequencing of heterogeneous viral populations using combinatorial pooling. *Bioinformatics*, btu726.
[5] GHOST Makes Connections in Hepatitis C Virus Transmission. Web resource. Page last updated: October 16, 2015. Centers for Disease Control and Prevention.

Wednesday, July 20

09:15-10:00 **Tutorial: Fereydoun Hormozdiari**

Detecting Structural Variation

[1] Hormozdiari, F., Alkan, C., Eichler, E. E., & Sahinalp, S. C. (2009). Combinatorial algorithms for structural variation detection in high-throughput sequenced genomes.*Genome research*, 19(7), 1270-1278.
[2] Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature Reviews Genetics*, 12(5), 363-376.
[3] Handsaker, R. E., Korn, J. M., Nemesh, J., & McCarroll, S. A. (2011). Discovery and genotyping of genome structural polymorphism by sequencing on a population scale. *Nature genetics*, 43(3), 269-276.

	[4] Sindi, S. S., Önal, S., Peng, L. C., Wu, H. T., & Raphael, B. J. (2012). An integrative probabilistic model for identification of structural variation in sequencing data.*Genome biology*, 13(3), 1. [5] Rausch, T., Zichner, T., Schlattl, A., Stütz, A. M., Benes, V., & Korbel, J. O. (2012). DELLY: structural variant discovery by integrated paired-end and split-read analysis.*Bioinformatics*, 28(18), i333-i339. [6] Layer, R. M., Chiang, C., Quinlan, A. R., & Hall, I. M. (2014). LUMPY: a probabilistic framework for structural variant discovery. *Genome biology*, 15(6), 1. [7] Chaisson, M. J., Huddleston, J., Dennis, M. Y., Sudmant, P. H., Malig, M., Hormozdiari, F., … & Landolin, J. M. (2015). Resolving the complexity of the human genome using single-molecule sequencing. *Nature*, 517(7536), 608-611.
10:00-10:45	**Research Talk: Jason Ernst** Deciphering the Non-coding Human Genome [1] Ernst, J., & Kellis, M. (2010). Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature Biotechnology*, 28(8), 817-825. [2] Ernst, J., Kheradpour, P., Mikkelsen, T. S., Shoresh, N., Ward, L. D., Epstein, C. B., … & Ku, M. (2011). Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*, 473(7345), 43-49. [3] Ernst, J., & Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nature Methods*, 9(3), 215-216. [4] Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., … & Amin, V. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), 317-330. [5] Kheradpour, P., Ernst, J., Melnikov, A., Rogov, P., Wang, L., Zhang, X., … & Kellis, M. (2013). Systematic dissection of regulatory motifs in 2,000 predicted human enhancers using a massively parallel reporter assay. *Genome Research*, 23(5), 800-811.
10:45-11:15	Coffee Break
11:15-12:00	Journal Club
12:00-14:00	Lunch Break During this time, UCLA is hosting a Career Opportunities lunch on Wednesday July 20th at 12:30pm where we will describe the various programs at UCLA including our Bioinformatics Ph.D. Program, our Post-Doctoral Fellowships in Bioinformatics through the UCLA Collaboratory, faculty hiring opportunities through the Quantitative and Computational Biology Institute Faculty Search as well as describing our undergraduate programs. **Contact information from UCLA Career Opportunities Lunch** Hillary Coller: hcoller@ucla.edu Paivi Pajukanta: ppajukanta@mednet.ucla.edu

14:00-14:45	**Research Talk: Bogdan Pasaniuc** Methods to Understand the Polygenic Architecture of Complex Traits [1] Shi, H., Kichaev, G., & Pasaniuc, B. (2016). Contrasting the genetic architecture of 30 complex traits from summary association data. *bioRxiv*, 035907. [2] Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B. W., … & Sullivan, P. F. (2016). Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics*. [3] Kichaev, G., & Pasaniuc, B. (2015). Leveraging functional-annotation data in trans-ethnic fine-mapping studies. *The American Journal of Human Genetics*, *97*(2), 260-271. [4] Gusev, A., Shi, H., Kichaev, G., Pomerantz, M., Li, F., Long, H. W., … & Nemesure, B. (2016). Atlas of prostate cancer heritability in European and African-American men pinpoints tissue-specific regulation. *Nature Communications*, *7*.
14:45-15:15	Coffee Break
15:15-16:00	**Tutorial: Ben Raphael** Computational Analysis of Somatic Mutations in Cancer [1] Ding, L., Wendl, M. C., McMichael, J. F., & Raphael, B. J. (2014). Expanding the computational toolbox for mining cancer genomes. *Nature Reviews Genetics*, 15(8), 556-570. doi:10.1038/nrg3767. PubMed PMID: 25001846. [2] Raphael, B. J., Dobson, J. R., Oesper, L., & Vandin, F. (2014). Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine.*Genome Medicine*, 6(1), 1. doi: 10.1186/gm524. PMID: 24479672. [3] Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., & Kinzler, K. W. (2013). Cancer genome landscapes. *Science*, 339(6127), 1546-1558. doi: 10.1126/science.1235122. PubMed PMID: 23539594. [4] Oesper, L., Mahmoody, A., & Raphael, B. J. (2013).THetA: inferring intra-tumor heterogeneity from high-throughput DNA sequencing data. *Genome Biology*, 14(7), 1. doi: 10.1186/gb-2013-14-7-r80. PubMed PMID: 23895164. [5] El-Kebir, M., Oesper, L., Acheson-Field, H., & Raphael, B. J. (2015). Reconstruction of clonal trees and tumor composition from multi-sample sequencing data.*Bioinformatics*, 31(12), i62-i70. doi: 10.1093/bioinformatics/btv261. PubMed PMID: 26072510.
16:00-16:15	Coffee Break
16:15-17:00	**Research Talk: Alexander Schönhuth** Quantifying Uncertainties in Big Genome Data

Thursday, July 21

09:15-10:00	**Research Talk/Tutorial hybrid: Jennifer Listgarten** Structured Populations in Genetics [1] Lippert, C., Listgarten, J., Liu, Y., Kadie, C. M., Davidson, R. I., & Heckerman, D. (2011). FaST linear mixed models for genome-wide association studies. *Nature Methods*, *8*(10), 833-835. [2] Listgarten, J., Lippert, C., Kadie, C. M., Davidson, R. I., Eskin, E., & Heckerman, D. (2012). Improved linear mixed models for genome-wide association studies. *Nature Methods*, *9*(6), 525-526. [3] Listgarten, J., Lippert, C., & Heckerman, D. (2013).FaST-LMM-Select for addressing confounding from spatial structure and rare variants. *Nature Genetics*, *45*(5), 470-471. [4] Zou, J., Lippert, C., Heckerman, D., Aryee, M., & Listgarten, J. (2014). Epigenome-wide association studies without the need for cell-type composition. *Nature Methods*, *11*(3), 309-11. [5] A New Method for Deducing the Attention Span of Workshop Attendees, *Journal of Visionary Research* 2017. [6] Also see the Fast LMM page for a full, annotated bibliography and pointers to software related to GWAS/EWAS.
10:00-10:45	**Tutorial: Saharon Rosset** Bootstrap – The Statistician's Magic Wand [1] Efron, B., & Tibshirani, R. J. (1994). An Introduction to the Bootstrap. Chapman & Hall/CRC. [2] Felsenstein, J. (1985). Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution*, 39(4), 783-791. [3] Efron, B., Halloran, E., & Holmes, S. (1996). Bootstrap confidence levels for phylogenetic trees. *Proceedings of the National Academy of Sciences*, 93(23), 13429-13429.
10:45-11:15	Coffee Break
11:15-12:00	Journal Club
12:00-14:00	Lunch Break
14:00-14:45	**Research Talk: Noah Zaitlen** Variances and Covariances in Recently Admixed Populations [1] Zou, J. Y., Park, D. S., Burchard, E. G., Torgerson, D. G., Pino-Yanes, M., Song, Y. S., … & Zaitlen, N. (2015). Genetic and socioeconomic study of mate choice in Latinos reveals novel assortment patterns. *Proceedings of the National Academy of Sciences*, *112*(44), 13621-13626. [2] Zaitlen, N., Huntsman, S., Hu, D., Spear, M., Eng, C., Oh, S. S., … & Brigino-Buenaventura, E. (2016). The Effects of Migration and Assortative Mating on Admixture Linkage Disequilibrium. *bioRxiv*, 056168.

	[3] Park, D., Eskin, I., Kang, E. Y., Gamazon, E. R., Eng, C., Gignoux, C. R., … & Eskin, E. (2016). An Ancestry Based Approach for Detecting Interactions. *bioRxiv*, 036640.
14:45-15:15	Coffee Break
14:45-15:30	**Research Talk: Brian Browning** Genotype Imputation with Millions of Reference Samples [1] Marchini J, Howie B. (2010). Genotype imputation for genome-wide association studies. *Nature Reviews Genetics*, 11(7), 499-511. DOI:10.1038/nrg2796. [2] Howie, B., Fuchsberger, C., Stephens, M., Marchini, J., & Abecasis, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nature Genetics*, 44(8), 955-959. PMCID: PMC3696580. DOI: 10.1038/ng.2354. [3] Browning, B. L., & Browning, S. R. (2016). Genotype Imputation with Millions of Reference Samples. *The American Journal of Human Genetics*, 98(1), 116-126. PMCID: PMC4716681 DOI: 10.1016/j.ajhg.2015.11.020.
15:30-16:00	Coffee Break
16:00-16:50	**Research Talk: Dan Geschwind** Integrative Genomics in Neuropsychiatric Diseases [1] Parikshak, N. N., Gandal, M. J., & Geschwind, D. H. (2015). Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders.*Nature Reviews Genetics*, 16(8), 441-458. doi: 10.1038/nrg3934. PMID: 26149713. [2] Parikshak, N. N., Luo, R., Zhang, A., Won, H., Lowe, J. K., Chandran, V., … & Geschwind, D. H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, 155(5), 1008-1021. doi: 10.1016/j.cell.2013.10.031. PMID: 24267887. [3] Geschwind, D. H., & Flint, J. (2015). Genetics and genomics of psychiatric disease. *Science*, 349(6255), 1489-1494. doi: 10.1126/science.aaa8954. PMID:26404826. [4] Stein, J. L., de la Torre-Ubieta, L., Tian, Y., Parikshak, N. N., Hernández, I. A., Marchetto, M. C., … & Wexler, E. M. (2014). A quantitative framework to evaluate modeling of cortical development by neural stem cells. *Neuron*, 83(1), 69-86. doi: 10.1016/j.neuron.2014.05.035. [5] Chandran, V., Coppola, G., Nawabi, H., Omura, T., Versano, R., Huebner, E. A., … & Blesch, A. (2016). A systems-level analysis of the peripheral nerve intrinsic axonal growth program. *Neuron*, 89(5), 956-970. doi: 10.1016/j.neuron.2016.01.034.

Friday, July 22

09:15-10:00	**Tutorial: Sriram Sankararaman** Evolutionary Models in Population Genomics
10:00-10:45	**Tutorial: Wei Wang** Alignment-free RNASeq Analysis
10:45-11:15	Coffee Break
11:15-12:00	Journal Club
12:00-14:00	Lunch Break
14:00-14:45	**Research Talk: Jingyi (Jessica) Li** NMFP – Identifying mRNA Isoforms from RNASeq Data [1] Ye, Y., & Li, J. J. (2016). NMFP: a non-negative matrix factorization based preselection method to increase accuracy of identifying mRNA isoforms from RNA-seq data. *BMC Genomics*, 17(1), 127.
14:45-15:15	Coffee Break
15:15-16:00	**Research Talk: Kirk Lohmueller** The interplay between demography and selection in dogs, wolves, and foxes [1] Marsden, C. D., Ortega-Del Vecchyo, D., O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., … & Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. *Proceedings of the National Academy of Sciences*, 113(1), 152-157. PMID: 26699508. [2] Robinson, J. A., Ortega-Del Vecchyo, D., Fan, Z., Kim, B. Y., Marsden, C. D., Lohmueller, K. E., & Wayne, R. K. (2016). Genomic Flatlining in the Endangered Island Fox.*Current Biology*, 26(9), 1183-1189. doi: 10.1016/j.cub.2016.02.062. PMID: 27112291. [3] Henn, B. M., Botigué, L. R., Bustamante, C. D., Clark, A. G., & Gravel, S. (2015). Estimating the mutation load in human genomes. *Nature Reviews Genetics*, 16(6), 333-343. doi: 10.1038/nrg3931. PMID: 25963372. [4] Henn, B. M., Botigué, L. R., Peischl, S., Dupanloup, I., Lipatov, M., Maples, B. K., … & Excoffier, L. (2016).Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proceedings of the National Academy of Sciences*, 113(4), E440-E449. doi: 10.1073/pnas.1510805112. PMID: 26712023.

Week Two

Monday, July 25

Boyer Hall, Room 159

09:30-10:15	**Tutorial: Paul Medvedev** Applying CS tools to Biological questions: a case study in genome assembly
10:15-10:45	Coffee Break
10:45-11:00	**Mini Presentation: Elior Rahmani** Finding hidden signals in whole-genome epigenetic data
11:00-11:15	**Mini Presentation: Lara Urban** Integrative modelling of regulatory and phenotypic variation
11:15-11:30	**Mini Presentation: Regev Schweiger** Quantifying uncertainty in heritability estimation using linear mixed models
11:30-13:30	Lunch Break
13:30-13:45	**Mini Presentation: Eyal Fisher** Improving bootstrap confidence intervals via Quantile Regression
13:45-14:00	**Mini Presentation: Liat Shenhav** Quantifying replicability in systematic reviews: The r-value
14:00-14:15	**Mini Presentation: Adriana Sperlea** A novel annotation of the human genome through conservation states aids interpretation of disease associated genetic variation
14:15-15:15	**CGSI Long Course Orientation**

Tuesday, July 26

Boyer Hall, Room 159

09:30-10:15	**Tutorial: John Novembre** Models and methods for detecting population structure in humans

	[1] Schraiber, J. G., & Akey, J. M. (2015). Methods and models for unravelling human evolutionary history. *Nature Reviews Genetics*.
10:15-11:00	**Research Talk: Noah Rosenberg** Mathematical Bounds on Population-Genetic Statistics [1] Jakobsson, M., Edge, M. D., & Rosenberg, N. A. (2013).The relationship between FST and the frequency of the most frequent allele. *Genetics*, 193(2), 515-528.
11:00-11:30	Coffee Break
11:30-11:45	**Mini Presentation: Roni Wilentzik** Genome-Wide Analysis of the molecular function of eQTLs in regulatory networks
11:45-12:00	**Mini Presentation: Eliran Avni** Constructing the tree of life in light of extensive HGT events
12:00-12:15	**Mini Presentation: Christopher Robles** The impact of Neanderthal ancestry on depression in East Asians
	Wednesday, July 27 **Boyer Hall, Room 159**
09:30-10:15	**Tutorial: Kenneth Lange** MM Algorithms [1] Hunter, D. R., & Lange, K. (2004). A tutorial on MM algorithms. *The American Statistician*, 58(1), 30-37.
10:15-10:45	Coffee Break
10:45-11:00	**Mini Presentation: Siddharth Jain** Biological information channel
11:00-11:15	**Mini Presentation: Mohammed Alser** Enabling fast read alignment in genome analysis

11:15-11:30	**Mini Presentation: Kin Fai Au** Transcriptome analysis at the gene isoform level using hybrid sequencing
11:30-12:00	Coffee Break
12:00-12:15	**Mini Presentation: Diana Domanska** Genome analysis and visualization of genomic data
12:15-12:30	**Mini Presentation: Tanya Phung** Detecting mutagenic recombination using genome-wide divergence data
12:30-12:45	**Mini Presentation: Liangke Gou** Understanding the Genetic Basis of Mutation Rate Variation in Yeast
16:00-17:00	External (non-CGSI) Event **UCLA Bioinformatics Lecture: DNA and RNA Editing of Retrotransposons Accelerate Mammalian Genome Evolution** Erez Levanon, Ph.D., The Mina & Everard Faculty of Life Sciences, Bar-Ilan University Ramat Gan, Israel Erez Levanon Seminar Information

Thursday, July 28

Boyer Hall, Room 130

09:30-10:15	**Research Talk: Saharon Rosset** Mixed Modeling for Case-Control Genome-Wide Studies: A Major Challenge [1] Golan, D., Lander, E. S., & Rosset, S. (2014). Measuring missing heritability: inferring the contribution of common variants. *Proceedings of the National Academy of Sciences*, 111(49), E5272-E5281. [2] Golan, D., & Rosset, S. (2014). Effective genetic-risk prediction using mixed models. *The American Journal of Human Genetics*, 95(4), 383-393. doi: 10.1016/j.ajhg.2014.09.007. [3] Golan, D., & Rosset, S. (2016). Mixed Models for Case-Control Genome-Wide Association Studies: Major Challenges and Partial Solutions. Draft book chapter available at: http://www.tau.ac.il/~saharon/ccGWAS.pdf
10:15-10:45	Coffee Break
10:45-11:00	**Mini Presentation: Chloe Robins** Using DNA methylation data to understand the evolutionary basis of human aging

11:00-11:15	**Mini Presentation: Kristoffer Sahlin** Modeling illumina read pair data for assembly scaffolding and structural variation detection
11:15-11:30	**Mini Presentation: Farhad Hormozdiari** Fine mapping and post-GWAS Era
11:30-12:00	Coffee Break
12:00-12:15	**Mini Presentation: Olga Schubert** Surveying the complete space of genetic loci influencing protein abundance
12:15-12:30	**Mini Presentation: David Koslicki** Mathematical methods in metagenomics
12:30-12:45	**Mini Presentation: Jaehee Kim** The effect of assortative mating on population admixture
14:00-15:00	**RNAseq Methods Discussion** Professors Yi Xing, Jessica Li, Wei Wang (UCLA)

Friday, July 29

IPAM Building

09:30-10:15	**Research Talk: Jonathan Flint** The Genetics of Depression
10:15-10:30	**Mini Presentation: Joel Mefford** Applications of genomic predictions from linear mixed models
10:30-10:45	**Mini Presentation: Pavel Avdeev** Genome rearrangements and reconstruction of ancestral genomes
10:45-11:15	Coffee Break
11:15-11:30	**Mini Presentation: Serghei Mangul** Dumpster diving in RNA-sequencing to find the source of every last read

11:30-11:45	**Mini Presentation: Pasquale Laise** Deciphering the Polycomb-dependent regulatory networks in gliomagenesis
11:45-12:30	**Research Talk: Yi Xing** Elucidating The Complexity of the Mammalian m6A Epitranscriptome [1] Molinie, B., Wang, J., Lim, K. S., Hillebrand, R., Lu, Z. X., Van Wittenberghe, N., … & Xing, Y. (2016). m6A-LAIC-seq reveals the census and complexity of the m6A epitranscriptome. *Nature Methods*.

Week Three

Monday, August 1

Boyer Hall, Room 159

09:30-10:15	**Tutorial: Jason Ernst** Epigenome Imputation [1] Ernst, J., & Kellis, M. (2015). Large-scale imputation of epigenomic datasets for systematic annotation of diverse human tissues. *Nature Biotechnology*, 33(4), 364-376. [2] Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., … & Amin, V. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), 317-330.
10:15-11:00	**Tutorial: Emilia Huerta-Sanchez** Evolutionary adaptation in Humans [1] Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics*, 39, 197-218. [2] Vitti, J. J., Grossman, S. R., & Sabeti, P. C. (2013).Detecting natural selection in genomic data. *Annual Review of Genetics*, 47, 97-120. [3] Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. *Nature Reviews Genetics*, 16(6), 359-371.
15:00-15:30	Tea Time

Tuesday, August 2

Boyer Hall, Room 159

10:30-11:15	**Tutorial: Bogdan Pasaniuc** Integrative fine-mapping for causal variants [1] Kichaev, G., & Pasaniuc, B. (2015). Leveraging functional-annotation data in trans-ethnic fine-mapping studies. *The American Journal of Human Genetics*, 97(2), 260-271. [2] Kichaev, G., Yang, W. Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A. L., … & Pasaniuc, B. (2014).Integrating functional data to prioritize causal variants in statistical fine-mapping studies. *PLoS Genet*, 10(10), e1004722. [3] Hormozdiari, F., Kichaev, G., Yang, W. Y., Pasaniuc, B., & Eskin, E. (2015). Identification of causal genes for complex traits. *Bioinformatics*, 31(12), i206-i213.
11:15-12:00	**Tutorial: Noah Zaitlen** Conditioning in association studies [1] Zaitlen, N., Lindström, S., Pasaniuc, B., Cornelis, M., Genovese, G., Pollack, S., … & Freedman, B. I. (2012).Informed conditioning on clinical covariates increases power in case-control association studies. *PLoS Genet*, 8(11), e1003032. [2] Aschard, H., Vilhjalmsson, B., Patel, C., Skurnik, D., Yu, J., Wolpin, B., … & Zaitlen, N. (2016). Playing Musical Chairs in Big Data to Reveal Variables Associations.*bioRxiv*, 057190. [3] Aschard, H., Vilhjálmsson, B. J., Joshi, A. D., Price, A. L., & Kraft, P. (2015). Adjusting for heritable covariates can bias effect estimates in genome-wide association studies.*The American Journal of Human Genetics*, 96(2), 329-339.
15:00-15:30	Tea Time
	Wednesday, August 3 **Boyer Hall, Room 159**
10:30-11:30	**Tutorial: Sagi Snir** Tutorial on Reconstructing Evolution from the Tiniest Fractions : Basics, Challenges, and Applications [1] Hinchliff, C. E., Smith, S. A., Allman, J. F., Burleigh, J. G., Chaudhary, R., Coghill, L. M., … & Gude, K. (2015).Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences*, 112(41), 12764-12769. [2] Snir, S., & Rao, S. (2010). Quartets maxcut: A divide and conquer quartets algorithm. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 7(4), 704-718. [3] Avni, E., Cohen, R., & Snir, S. (2015). Weighted quartets phylogenetics. *Systematic biology*, 64(2), 233-242.
15:00-15:30	Tea Time

Thursday, August 4

Boyer Hall, Room 130

10:30-11:15 **Research Talk: Sagi Snir**

Research Talk on Reconstructing Evolution from the Tiniest Fractions : Basics, Challenges, and Applications

[1] Steel, M. (1992). The complexity of reconstructing trees from qualitative characters and subtrees. *Journal of classification*, 9(1), 91-116.
[2] Snir, S., & Yuster, R. (2012). Reconstructing approximate phylogenetic trees from quartet samples.*SIAM Journal on Computing*, 41(6), 1466-1480.
[3] Alon, N., Snir, S., & Yuster, R. (2014). On the compatibility of quartet trees. *SIAM Journal on Discrete Mathematics*, 28(3), 1493-1507.
[4] Roch, S., & Snir, S. (2013). Recovering the treelike trend of evolution despite extensive lateral genetic transfer: a probabilistic analysis. *Journal of Computational Biology*, 20(2), 93-112.

11:15-12:15 **Research Talk: Kirk Lohmueller**

Comparison of the distribution of fitness effects across species using the Poisson Random Field framework

[1] Sawyer, S. A., & Hartl, D. L. (1992). Population genetics of polymorphism and divergence. *Genetics*, 132(4), 1161-1176.
[2] Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, 8(8), 610-618.
[3] Boyko, A. R., Williamson, S. H., Indap, A. R., Degenhardt, J. D., Hernandez, R. D., Lohmueller, K. E., … & White, T. J. (2008). Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet*, 4(5), e1000083.

15:00-15:30 Tea Time

Friday, August 5

Boyer Hall, Room 130

11:00-11:45 **Research Talk: Sriram Sankararaman**

Inferring the structure of archaic admixture in modern humans

[1] Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., … & Reich, D. (2012). Ancient admixture in human history. *Genetics*, 192(3), 1065-1093.
[2] Sankararaman, S., Patterson, N., Li, H., Pääbo, S., & Reich, D. (2012). The date of interbreeding between Neandertals and modern humans. *PLoS Genet*, 8(10), e1002947.

[3] Sankararaman, S., Mallick, S., Dannemann, M., Prüfer, K., Kelso, J., Pääbo, S., … & Reich, D. (2014). The genomic landscape of Neanderthal ancestry in present-day humans.*Nature*, 507(7492), 354-357.

15:00-15:30 Tea Time

Week Four

Monday, August 8

Boyer Hall, Room 159

10:30-11:30 **Tutorial: Barbara Engelhardt**

Dimension reduction methods to identify latent structure in genomic data

11:30-12:30 **Tutorial: Sriram Sankararaman**

Challenges in genomic privacy

[1] Erlich, Y., & Narayanan, A. (2014). Routes for breaching and protecting genetic privacy. *Nature Reviews Genetics*, 15(6), 409-421.
[2] Gymrek, M., McGuire, A. L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science*, 339(6117), 321-324.
[3] Sankararaman, S., Obozinski, G., Jordan, M. I., & Halperin, E. (2009). Genomic privacy and limits of individual detection in a pool. *Nature genetics*, 41(9), 965-967.
[4] Dwork, C. (2008, April). Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation* (pp. 1-19). Springer Berlin Heidelberg.

15:00-15:30 Tea Time

Tuesday, August 9

Boyer Hall, Room 159

10:30-11:30 **Research Talk: Barbara Engelhardt**

Exploring covariation in gene expression data

11:30-12:30 **Tutorial: Kenneth Lange**

Next generation statistical genetics: Modeling, penalization, and optimization in high-dimensional data

[1] Lange, K., Papp, J. C., Sinsheimer, J. S., & Sobel, E. M. (2014). Next generation statistical genetics: Modeling, penalization, and optimization in high-dimensional data.*Annual Review of Statistics and its Application*, 1(1), 279.

15:00-15:30 Tea Time

Wednesday, August 10

Boyer Hall, Room 159

10:30-11:30 **Research Talk: Jo Hardin**

Assumptions in Normalizing RNASeq Data

[1] Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, *10*(1), 57-63.
[2] 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), 1.
[3] Dillies, M. A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., … & Guernec, G. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis.*Briefings in Bioinformatics*, *14*(6), 671-683.
[4] Lovén, J., Orlando, D. A., Sigova, A. A., Lin, C. Y., Rahl, P. B., Burge, C. B., … & Young, R. A. (2012). Revisiting global gene expression analysis. *Cell*, *151*(3), 476-482.

15:00-16:00 **Research Talk: Jonathan Flint**

Quantitative Trait Locus Detection

Note: This talk will take place in Boelter Hall, Room 4760.

16:00-17:00 External (non-CGSI) Event

UCLA Bioinformatics Lecture: Transcriptome analysis by hybrid sequencing
Kin Fai Au, Ph.D., Assistant Professor of Internal Medicine and Biostatistics, University of Iowa
Kin Fai Au Seminar Information

Note: This talk will take place in Boelter Hall, Room 4760.

Thursday, August 11

Boelter Hall, Room 4760

10:30-11:30	**Tutorial: David Koslicki** Bacterial Community Reconstruction via Compressed Sensing [1] Candès, E. J., & Wakin, M. B. (2008). An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2), 21-30. [2] Koslicki, D., Foucart, S., & Rosen, G. (2013). Quikr: a method for rapid reconstruction of bacterial communities via compressive sensing. *Bioinformatics*, btt336. [3] Koslicki, D., Chatterjee, S., Shahrivar, D., Walker, A. W., Francis, S. C., Fraser, L. J., … & Corander, J. (2015). ARK: Aggregation of Reads by K-Means for Estimation of Bacterial Community Composition. *PloS one*, 10(10), e0140644. [4] Chatterjee, S., Koslicki, D., Dong, S., Innocenti, N., Cheng, L., Lan, Y., … & Corander, J. (2014). SEK: sparsity exploiting k-mer-based estimation of bacterial community composition. *Bioinformatics*, 30(17), 2423-2431.
11:30-12:30	**Tutorial: Jason Ernst** Functional Genomics Time-series Analysis [1] Ernst, J., Nau, G. J., & Bar-Joseph, Z. (2005). Clustering Short Time Series Gene Expression Data. *Bioinformatics*, 21(suppl 1), i159-i168. [2] Ernst, J., & Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinformatics*, 7(1), 1. [3] Ernst, J., Vainas, O., Harbison, C. T., Simon, I., & Bar-Joseph, Z. (2007). Reconstructing dynamic regulatory maps. *Molecular Systems Biology*, 3(1), 74. [4] Bar-Joseph, Z., Gitter, A., & Simon, I. (2012). Studying and modeling dynamic biological processes using time-series gene expression data. *Nature Reviews Genetics*, 13(8), 552-564.
15:00-15:30	Tea Time
	Friday, August 12 **Boyer Hall, Room 130**
10:00-11:00	**Research Talk: Joel Mefford** Linear Mixed Models – And Their Equivalents
15:00-15:30	Tea Time