

OEB 275R: Phylogenetics and Phylogeography in the Era of Genomics

Course Instructor: Scott Edwards (MCZ Labs 306, sedwards@fas.harvard.edu)

Teaching Fellow: Tauana Cunha (MCZ Labs 510, tauanacunha@g.harvard.edu) Office Hours: Tue 5-6pm

Academic Integrity Policy

Discussion and the exchange of ideas are essential to academic work, and this practice will apply to assignments in this course. For all activities in this course, including homework and the final project, you may freely consult with class colleagues, borrow and exchange ideas. You may find it useful to discuss your chosen topic with your peers, particularly if you are working on the same topic as a classmate. However, you should ensure that any written work you submit for evaluation is the result of your own research and writing and that it reflects your own approach to the topic. Of course, you must write any written material yourself and you must also adhere to standard citation practices in biological science and properly cite any books, articles, websites, lectures, personal communications or other exchanges of information that help you produce your work. If you received any help with your writing (feedback on drafts, ideas etc), you must also acknowledge this assistance. This is standard scientific practice and putting it into practice now will help you become a better scientist.

Syllabus

Date	Details
30-Aug	L01 - Scott: Intro to phylogeography/phylogenetics
6-Sep	L02 - Scott: Phylogenetics in the context of molecular evolution
13-Sep	L03 - Scott: Coalescent theory: a bridge between population genetics and phylogenetics
20-Sep	L04 - Axel Meyer: Gene and genome duplications
27-Sep	L05/S - Allison Shultz: Phylogeography and population genetics using Rad-seq/Stacks
4-Oct	L06/S - Ziheng Yang: Species delimitation and phylogeny using Bayesian Phylogenetics and Phylogeography
11-Oct	L07 - Craig Moritz: Phylogeography and biodiversity discovery
18-Oct	L08/S - Peter Beerli: Bayes factors and population model selection/Migrate-n
25-Oct	L09/S - Tim Sackton: Genome analysis: population genetics and comparative genomics
1-Nov	L10/S - Molly Schumer: Ancestry inference in hybrids
8-Nov	L11/S - Nicholas Patterson: Population genetics and population history using f-statistics
15-Nov	L12/S - Luay Nakhleh: The multispecies network coalescent
22-Nov	THANKSGIVING RECESS
29-Nov	L13 - Summary, project lab session and prospectus

Phylogenetics

- Shen, X., Hittinger, C. T., & Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Publishing Group*, 1(April), 1–10. <https://doi.org/10.1038/s41559-017-0126>
- Xi, Z., Liu, L., & Davis, C. C. (2016). The impact of missing data on species tree estimation. *Molecular Biology and Evolution*, 33(3), 838–860. <https://doi.org/10.1093/molbev/msv266>
- Edwards, S. V., Xi, Z., Janke, A., Faircloth, B. C., McCormack, J. E., Glenn, T. C., ... Davis, C. C. (2016). Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Molecular Phylogenetics and Evolution*, 94, 447–462. <https://doi.org/10.1016/j.ympev.2015.10.027>
- Xi, Z., Liu, L., & Davis, C. C. (2015). Genes with minimal phylogenetic information are problematic for coalescent analyses when gene tree estimation is biased. *Molecular Phylogenetics and Evolution*, 92, 63–71. <https://doi.org/10.1016/j.ympev.2015.06.009>
- Pyron, R. A. (2015). Post-molecular systematics and the future of phylogenetics. *Trends in Ecology & Evolution*, 30(7), 384–389. <https://doi.org/10.1016/j.tree.2015.04.016>
- Liu, L., Xi, Z., Wu, S., Davis, C. C., & Edwards, S. V. (2015). Estimating phylogenetic trees from genome-scale data. *Annals of the New York Academy of Sciences*, 1360(1), 36–53. <https://doi.org/10.1111/nyas.12747>
- Gatesy, J., & Springer, M. S. (2014). Phylogenetic Analysis at Deep Timescales: Unreliable Gene Trees, Bypassed Hidden Support, and the Coalescence/Concatalaescence Conundrum. *Molecular Phylogenetics and Evolution*, 80, 231–266. <https://doi.org/10.1016/j.ympev.2014.08.013>
- Jiang, W., Chen, S.-Y., Wang, H., Li, D.-Z., & Wiens, J. J. (2014). Should genes with missing data be excluded from phylogenetic analyses? *Molecular Phylogenetics and Evolution*, 80, 308–318. <https://doi.org/10.1016/j.ympev.2014.08.006>
- Liu, L., Xi, Z., & Davis, C. C. (2014). Coalescent Methods Are Robust to the Simultaneous Effects of Long Branches and Incomplete Lineage Sorting. *Molecular Biology and Evolution*, 32(3), 791–805. <https://doi.org/10.1093/molbev/msu331>
- McCormack, J. E., Hird, S. M., Zellmer, A. J., Carstens, B. C., & Brumfield, R. T. (2013). Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, 66(2), 526–38. <https://doi.org/10.1016/j.ympev.2011.12.007>
- Yang, Z., & Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nature Reviews Genetics*, 13(5), 303–314. <https://doi.org/10.1038/nrg3186>
- Edwards, S. V. (2009). Is a new and general theory of molecular systematics emerging? *Evolution*, 63(1), 1–19. <https://doi.org/10.1111/j.1558-5646.2008.00549.x>
- Degnan, J. H., & Rosenberg, N. a. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, 24(6), 332–340. <https://doi.org/10.1016/j.tree.2009.01.009>

- Hillis, D. M., Pollock, D. D., McGuire, J. A., & Zwickl, D. J. (2003). Is Sparse Taxon Sampling a Problem for Phylogenetic Inference? *Systematic Biology*, 52(1), 124–126.
- Rokas, A., Williams, B. L., King, N., & Carroll, S. B. (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, 425(6960), 798–804.
<https://doi.org/10.1038/nature02053>
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R., & Bollback, J. P. (2001). Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, 294, 2310–2314.
<https://doi.org/10.1126/science.1065889>
- Maddison, W. P., Biology, S., & Sep, N. (1997). Gene trees in species trees. *Systematic Biology*, 46(3), 523–536.
- Sullivan, J., & Swofford, D. L. (1997). Are guinea pigs rodents? The importance of adequate models in molecular phylogenetics. *Journal of Mammalian Evolution*, 4(2), 77–86.
<https://doi.org/10.1023/A:1027314112438>
- Felsenstein, J. (1985). Phylogenies and the Comparative Method. *The American Naturalist*, 125(1), 1–15.
- Felsenstein, J. (1981). Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17(6), 368–376.
<https://doi.org/10.1007/BF01734359>

Molecular Evolution/Detecting Natural Selection

- Field, Y., Boyle E. A., Telis N., Gao Z., Gaulton K. J., Golan D., Yengo L., Rocheleau G., Froguel P., McCarthy M. I., and Pritchard J. K. 2016. Detection of human adaptation during the past 2000 years. *Science* 354:760-764.
- Vitti, J. J., Grossman S. R., and Sabeti P. C. 2013. Detecting natural selection in genomic data. *Annu Rev Genet* 47:97-120.
- Pritchard, J. K., Pickrell J. K., and Coop G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr Biol* 20:R208-215.
- Pritchard, J. K., and Di Rienzo A. 2010. Adaptation - not by sweeps alone. *Nat Rev Genet* 11:665-667.
- Grossman, S. R., Shlyakhter I., Karlsson E. K., Byrne E. H., Morales S., Frieden G., Hostetter E., Angelino E., Garber M., Zuk O., Lander E. S., Schaffner S. F., and Sabeti P. C. 2010. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* 327:883-886.
- Voight, B. F., Kudaravalli S., Wen X., and Pritchard J. K. 2006. A map of recent positive selection in the human genome. *PLoS Biol* 4:e72.
- Yang, Z., Wong W. S., and Nielsen R. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. *Molecular biology and evolution* 22:1107-1118.
- Nielsen, R. 2005. Molecular signatures of natural selection. *Annu. Rev. Genet.* 39:197-218.
- Garrigan, D., and Hedrick P. W. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution* 57:1707-1722.
- Sabeti, P. C., Reich D. E., Higgins J. M., Levine H. Z., Richter D. J., Schaffner S. F., Gabriel S. B., Platko J. V., Patterson N. J., McDonald G. J., Ackerman H. C., Campbell S. J., Altshuler D., Cooper R., Kwiatkowski D., Ward R., and Lander E. S. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419:832-837.
- Reich, D. E., Cargill M., Bolk S., Ireland J., Sabeti P. C., Richter D. J., Lavery T., Kouyoumjian R., Farhadian S. F., Ward R., and Lander E. S. 2001. Linkage disequilibrium in the human genome. *Nature* 411:199-204.
- Goldman, N., and Yang Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Molecular Biology and Evolution* 11:725-736.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-589.
- Takahata, N. 1987. On the Overdispersed Molecular Clock. *Genetics* 116:169-179.
- Hudson, R., Kreitman M., and Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153-159.

Phylogeography/Empirical Population Genetics

- Rannala, B., and Yang Z. 2017. Efficient Bayesian species tree inference under the multispecies coalescent. *Syst Biol*.
- Nielsen, R., Akey J. M., Jakobsson M., Pritchard J. K., Tishkoff S., and Willerslev E. 2017. Tracing the peopling of the world through genomics. *Nature* 541:302-310.
- Xu, B., and Yang Z. 2016. Challenges in Species Tree Estimation Under the Multispecies Coalescent Model. *Genetics* 204:1353-1368.
- Espindola, A., Ruffley M., Smith M. L., Carstens B. C., Tank D. C., and Sullivan J. 2016. Identifying cryptic diversity with predictive phylogeography. *Proceedings of the Royal Society B-Biological Sciences* 283.
- Edwards, S. V., Schultz A., and Campbell-Staton S. 2015. Next-generation sequencing and the expanding domain of phylogeography. *Folia Zoologica* 64:187-206.
- Carstens, B. C., Pelletier T. A., Reid N. M., and Satler J. D. 2013. How to fail at species delimitation. *Molecular ecology* 22:4369-4383.
- Pickrell, J. K., and Pritchard J. K. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* 8:e1002967.
- Patterson, N., Moorjani P., Luo Y., Mallick S., Rohland N., Zhan Y., Genschoreck T., Webster T., and Reich D. 2012. Ancient admixture in human history. *Genetics* 192:1065-1093.
- Helmkamp, L. J., Jewett E. M., and Rosenberg N. A. 2012. Improvements to a class of distance matrix methods for inferring species trees from gene trees. *Journal of computational biology : a journal of computational molecular cell biology* 19:632-649.
- Hickerson, M. J., Carstens B. C., Cavender-Bares J., Crandall K. A., Graham C. H., Johnson J. B., Rissler L., Victoriano P. F., and Yoder A. D. 2010. Phylogeography's past, present, and future: 10 years after Avise, 2000. *Mol Phylogenet Evol* 54:291-301.
- Excoffier, L., Beaumont M. A., Nielsen R., Robert C., Hey J., Gaggiotti O., Knowles L., Estoup A., Panchal M., Corander J., Hickerson M., Sisson S. A., Fagundes N., Chikhi L., Beerli P., Vitalis R., Cornuet J. M., Huelsenbeck J., Foll M., Yang Z. H., Rousset F., and Balding D. 2010. In defence of model-based inference in phylogeography REPLY. *Molecular Ecology* 19:436-446.
- Beerli, P., and Palczewski M. 2010. Unified framework to evaluate panmixia and migration direction among multiple sampling locations. *Genetics* 185:313-326.
- Nielsen, R., and Beaumont M. A. 2009. Statistical inferences in phylogeography. *Molecular Ecology* 18:1034-1047.
- Degnan, J. H., and Rosenberg N. A. 2009. Gene tree discordance, phylogenetic inference, and the multispecies coalescent. *Trends in Ecology & Evolution* 24:332-340.
- Carstens, B. C., Stoute H. N., and Reid N. M. 2009. An information-theoretical approach to phylogeography. *Molecular Ecology* 18:4270-4282.
- Patterson, N., Price A. L., and Reich D. 2006. Population structure and eigenanalysis. *PLoS Genetics* Volume 2:e190.

- Beaumont, M. A., and Rannala B. 2004. The Bayesian revolution in genetics. *Nature Reviews Genetics* 5:251-261.
- Rannala, B., and Yang Z. 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164:1645-1656.
- Hudson, R. R., and Turelli M. 2003. Stochasticity overrules the "three-times rule": Genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57:182-190.
- Brumfield, R., Nickerson D. A., Beerli P., and Edwards S. V. 2003. The utility of single nucleotide polymorphisms in inferences of population history. *Trends in Ecology and Evolution* 18:249-256.
- Yang, Z. 2002. Likelihood and Bayes estimation of ancestral population sizes in Hominoids using data from multiple loci. *Genetics* 162:1811-1823.
- Irwin, D. E. 2002. Phylogeographic breaks without geographic barriers to gene flow. *Evolution* 56:2383-2394.
- Reich, D. E., Cargill M., Bolk S., Ireland J., Sabeti P. C., Richter D. J., Lavery T., Kouyoumjian R., Farhadian S. F., Ward R., and Lander E. S. 2001. Linkage disequilibrium in the human genome. *Nature* 411:199-204.
- Reich, D. E., Cargill M., Bolk S., Ireland J., Sabeti P. C., Richter D. J., Lavery T., Kouyoumjian R., Farhadian S. F., Ward R., and Lander E. S. 2001. Linkage disequilibrium in the human genome. *Nature* 411:199-204.
- Felsenstein, J. 2001. Accuracy of coalescent likelihood estimates: do we need more sites, more sequences, or more loci? *Molecular Biology and Evolution* 23:691-700.
- Edwards, S. V., and Beerli P. 2000. Perspective: Gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. *Evolution* 54:1839-1854.
- Beerli, P., and Felsenstein J. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics* 152:763-773.
- Nei, M., and Takahata M. 1993. Effective population size, genetic diversity and coalescence time in subdivided populations. *Journal of Molecular Evolution* 37:240-244.
- Hudson, R. R. 1992. Gene trees, species trees and the segregation of ancestral alleles. *Genetics* 131:509-512.
- Hudson, R. R. 1990. Gene genealogies and the coalescent process. *Oxford Surv. Evol. Biol.* 7:1-44.
- Takahata, N. 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* 122:957-966.
- Avise, J. C., Arnold J., Ball R. M., Bermingham E., Lamb T., Neigel J. E., Reeb C. A., and Saunders N. C. 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual review of ecology and systematics*:489-522.
- Neigel, J. E., and Avise J. C. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation, Pages 515-534 in S. Karlin, and E. Nevo, eds. *Evolutionary Processes and Theory*. New York, Academic Press, Inc.