

Suggested Papers for Class Presentation

Diffusion process and importance sampling

- De Iorio, M. and Griffiths, R. C. (2004) Importance sampling on coalescent histories. I. *Adv. Appl. Probab.* 36(2): 417–433.
- De Iorio, M. and Griffiths, R. C. (2004) Importance sampling on coalescent histories. II: Subdivided population models *Adv. Appl. Probab.* 36(2): 434–454.
- De Iorio, M., Griffiths, R. C., Lebois, R. and Rousset, F. (2005). Stepwise mutation likelihood computation by sequential importance sampling in subdivided population models. *Theor. Popul. Biol.* 68: 41–53.

Frequency spectrum

- Sawyer, S. A. and Hartl, D. L. (1992) Population Genetics of Polymorphism and Divergence. *Genetics* 132: 1161–1176.
- Bustamante, C. D., Wakeley, J., Sawyer, S., and Hartl, D. L. (2001) Directional Selection and the Site-Frequency Spectrum. *Genetics* 159: 1779–1788.
- Griffiths, R. C. (2003). The frequency spectrum of a mutation, and its age, in a general diffusion model. *Theor. Popul. Biol.* 64: 241–251.

Estimating recombination rates

- Fearnhead, P. and Donnelly, P. (2001). Estimating recombination rates from population genetic data. *Genetics* 159: 1299–1318.
- Hudson, R. R. (2001). Two-locus sampling distributions and their application. *Genetics* 159: 1805–1817.
- Li, N. and Stephens, M. (2003). Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* 165: 2213–2233.
- Smith, N. G. C. and Fearnhead, P. (2005) A novel method with improved power to detect recombination hotspots from polymorphism data reveals multiple hotspots in human genes. *Am. J. Hum. Genet* 77: 781–794.
- Smith, N. G. C. and Fearnhead, P. (2005) A Comparison of Three Estimators of the Population-scaled Recombination Rate: Accuracy and Robustness. *Genetics* 171: 2051–2062.
- Mcvean, G., Myers, S., Hunt, S., Deloukas, P., Bentley, D., and Donnelly, P. (2004). The fine-scale structure of recombination rate variation in the human genome. *Science* 304: 581–584.
- Myers, S., Bottolo, L., Freeman, C., Mcvean, G., and Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science* 310: 321–324.

MCMC methods for the coalescent

- Kuhner, M. K., Yamato, J., and Felsenstein, J. (1995) Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. *Genetics* 140: 1421–1430.
- Kuhner, M. K., Yamato, J., and Felsenstein, J. (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* 149: 429–434.
- Kuhner, M. K., Beerli, P., Yamato, J. and Felsenstein, J. (2000) Usefulness of single nucleotide polymorphism data for estimating population parameters. *Genetics* 156: 439–447.
- Kuhner, M. K., Yamato, J. and Felsenstein, J. (2000) Maximum likelihood estimation of recombination rates from population data. *Genetics* 156: 1393–1401.
- Beerli, P. and Felsenstein, J. (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *PNAS* 98: 4563–4568.

Gene trees vs. species trees

- Rosenberg, N. A. (2002) The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology* 61: 225–247.
- Degnan, J. H. and Salter L. A. (2005) Gene tree distributions under the coalescent process. *Evolution* 59: 24–37.
- Degnan, J. H. and Rosenberg, N. A. (2006) Discordance of species trees with their most likely gene trees. *PLoS Genetics* 2: 762–768.

Population structure

- Pritchard, J.K., Stephens, M. and Donnelly, P. J. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K.K. et al., (2002) Genetic structure of human populations. *Science* 298: 2981–2985.
- Price, A.L, Patterson, N.J, Plenge, R.M., Weinblatt, M.E., Shadick, N.A., and Reich, D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics* 38: 904–909.
- Huelsenbeck, J.P. and Andolfatto, P. (2007) Inference of Population Structure Under a Dirichlet Process Model *Genetics* 175: 1787–1802.
- Sankararaman, S., Sridhar, S., Kimmel, G., and Halperin, E. (2008) Estimating local ancestry in admixed populations. *Am. J. Hum. Genet.* 82: 290–303.

Coalescent theory with selection

- Kaplan, N. L., T. Darden and R. R. Hudson. (1988) The coalescent process in models with selection. *Genetics* 120: 819–829.
- Kaplan, N. L, Hudson, R. R. and Langley, C. H. (1989) The “Hitchhiking Effect” Revisited. *Genetics* 123: 887–899.
- Coop, G. and Griffiths, R. C. (2004). Ancestral inference on gene trees under selection. *Theor. Popul. Biol.* 66: 219–232.

Multiple collisions

- Pitman, J. (1999). Coalescents with multiple collisions. *Annals of Probability* 27(4): 1870–1902.
- Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability* 36(4): 1116–1125.
- Schweinsberg, J. (2000). Coalescents with simultaneous multiple collisions. *Electronic Journal of Probability* 5, paper 12.

Selective sweep I

- Durrett, R. and Schweinsberg, J. (2004) Approximating selective sweeps. *Theor. Popul. Biol.* 66: 129–138.
- Schweinsberg, J and Durrett, R. (2005) Random partitions approximating the coalescence of lineages during a selective sweep. *Ann. Appl. Probab.* 15: 1591–1651.
- Durrett, R and Schweinsberg, J. (2005) A coalescent model for the effect of advantageous mutations on the genealogy of a population. *Stochastic Process. Appl.* 115: 1628–1657

Selective sweep II

- Etheridge, A., Pfaffelhuber, P., and Wakolbinger, A. (2006) An approximate sampling formula under genetic hitchhiking, *Ann. Appl. Probab.* 16(2): 685–729.
- Pfaffelhuber, P. and Studeny, A. (2007) Approximating genealogies for partially linked neutral loci under a selective sweep, *J. Math. Biol.* 55: 299–330.