

References

- PR05** (2005). Strange bayes indeed: uniform topological prior imply non-uniform clade priors. *Molecular Phylogenetics and Evolution* 34, 203–211.
- Ackerly2000** Ackerly, D. D. (2000, Oct). Taxon sampling, correlated evolution, and independent contrasts. *Evolution* 54(5), 1480–1492.
- Aldous1995** Aldous, D. (1995). *Random Discrete Structures*. New York: Springer.
- aldous96a** Aldous, D. (1996). Probability distributions on cladograms. In D. Aldous and R. Pemantle (Eds.), *Random Discrete Structures*, Volume 75 of *The IMA Volumes in Mathematics and its Applications*, pp. 1–18. Springer-Verlag.
- Aldous2001** Aldous, D. (2001). Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statistical Science* 16(1), 23–34.
- Allman2011a** Allman, E. S., J. H. Degnan, and J. A. Rhodes (2011a). Determining species tree topologies from clade probabilities under the coalescent. *Journal of Theoretical Biology* 289, 96–106.
- Allman2011b** Allman, E. S., J. H. Degnan, and J. A. Rhodes (2011b). Identifying the rooted species tree from the distribution of unrooted gene trees under the coalescent. *Journal of Mathematical Biology* 62, 833–862.
- Andrieu2010** Andrieu, C., A. Doucet, and R. Holenstein (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72(3), 269–342.
- Andrieu2005** Andrieu, C., A. Doucet, and V. Tadic (2005). Online EM for parameter estimation in nonlinear-non Gaussian state-space models. *Proc. IEEE CDC*.
- Andrieu2009** Andrieu, C. and G. O. Roberts (2009). The Pseudo-Marginal Approach for Efficient Monte Carlo Computations. *The Annals of Statistics* 37(2), 697–725.
- Arnason2004** Árnason, E. (2004). Mitochondrial cytochrome *b* variation in the high-fecundity Atlantic cod: trans-Atlantic clines and shallow gene genealogy. *Genetics* 166, 1871–1885.
- Arulampalam2002** Arulampalam, M., S. Maskell, N. Gordon, and T. Clapp (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on* 50(2), 174 –188.
- Backeljau1996** Backeljau, T., L. D. Bruyn, H. D. W. K. Jordaens, S. V. Dongen, and B. Winnepeenninckx (1996). Multiple upgma and neighbor-joining trees and the performance of some computer packages. *Molecular Biology and Evolution* 13(2), 309–313.

- Bafna2004** Bafna, V. and V. Bansal (2004, April). The number of recombination events in a sample history: Conflict graph and lower bounds. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 1(2), 78–90.
- Bagchi1985** Bagchi, A. and A. K. Pal (1985). Asymptotic normality in the generalized Polya–Eggenberger urn model, with an application to computer data structures. *SIAM Journal on Algebraic and Discrete Methods* 6(3), 394–405.
- Bandelt1986** Bandelt, H.-J. and A. Dress (1986, September). Reconstructing the shape of a tree from observed dissimilarity data. *Advances in Applied Mathematics* 7, 309–343.
- Barthelemy1991** Barthélemy, J.-P. and A. Guenoche (1991). *Trees and Proximity Representations*. London: Wiley.
- Beaumont2002** Beaumont, M. A., W. Zhang, and D. J. Balding (2002). Approximate bayesian computation in population genetics. *Genetics* 162, 2025–2035.
- Beckenbach1994** Beckenbach, A. (1994). Mitochondrial haplotype frequencies in oysters: neutral alternatives to selection models. In B. Golding (Ed.), *Non-neutral Evolution*, pp. 188–198. New York: Chapman and Hall.
- Benavoli2007** Benavoli, A., L. Chisci, and A. Farina (2007). Tracking of a ballistic missile with a-priori information. *Aerospace and Electronic Systems, IEEE Transactions on* 43(3), 1000–1016.
- Berti2010** Berti, P., I. Crimaldi, L. Pratelli, and P. Rigo (2010). Central limit theorems for multicolor urns with dominated colors. *Stochastic Processes and their Applications* 120, 1473–1491.
- Blackwell11973** Blackwell, D. and J. B. MacQueen (1973). Ferguson distribution via Pólya Urn schemes. *The Annals of Statistics* 1(2), 353–355.
- Blum2005clade** Blum, M. G. B. and O. Francois (2005a). Minimal clade size and external branch length under the neutral coalescent. *Advances in Applied Probability* 37, 647–662.
- Blum2005index** Blum, M. G. B. and O. Francois (2005b). On statistical tests of phylogenetic tree imbalance: The sackin and other indices revisited. *Mathematical Biosciences* 195, 141–153.
- Blum2006** Blum, M. G. B., O. Francois, and S. Janson (2006). The mean, variance and limiting distribution of two statistics sensitive to phylogenetic tree balance. *The Annals of Applied Probability* 16(4), 2195–2214.
- Bortolussi2006** Bortolussi, N., E. Durand, M. Blum, and O. Francois (2006). aptreeshape: statistical analysis of phylogenetic tree shape. *Bioinformatics* 22(3), 363–364.

- Bortolussi2009** Bortolussi, N., E. Durand, M. Blum, and O. Francois (2009). Analyses of phylogenetic treeshape. <http://cran.r-project.org/web/packages/apTreeshape/apTreeshape.pdf>, Feb 22, 2013.
- Brown1994** Brown, J. K. M. (1994). Probabilities of evolutionary trees. *Systematic Biology* 43(1), 78–91.
- Browning2013** Browning, B. L. and S. R. Browning (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics* 194(2), 459–471.
- Browning2007** Browning, S. R. and B. L. Browning (2007). Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics* 81(5), 1084 – 1097.
- Brumfield2010** Brumfield, R. T. and M. D. Carling (2010). The influence of hybrid zones on species tree inference in manakins. In L. L. Knowles and L. S. Kubatko (Eds.), *Estimating Species Trees, Practical and Theoretical Aspects*, pp. 115–127. Hoboken, NJ: Wiley-Blackwell.
- Bruno2000** Bruno, W. J., N. D. Socci, and A. L. Halpern (2000). Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction. *Molecular Biology and Evolution* 17, 189–197.
- Bryant2004** Bryant, D. (2004). The splits in the neighbourhood of a tree. *Annals of Combinatorics* 8(1), 1–11.
- Bryant2002** Bryant, D. and V. Moulton (2002). NeighborNet: an agglomerative method for the construction of planar phylogenetic networks. *Algorithms in Bioinformatics*, 375–391.
- Bryant2009** Bryant, D. and M. A. Steel (2009). Computing the distribution of a tree metric. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 6(3), 420–426.
- Buckley2006** Buckley, T. R., M. Cordeiro, D. C. Marshall, and C. Simon (2006). Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (*Maoricicada Dugdale*). *Systematic Biology* 55(3), 411–425.
- Buneman1971** Buneman, P. (1971). The Recovery of Trees from Measures of Dissimilarity. In D. Kendall and P. Tautu (Eds.), *Mathematics in the Archeological and Historical Sciences*, pp. 387–395. Edinburgh University Press.
- Buneman1974_DM** Buneman, P. (1974a). A Characterisation of Rigid Circuit Graphs. *Discrete Mathematics* 9, 205–212.
- Buneman1974_JCT** Buneman, P. (1974b). A Note on the Metric Properties of Trees. *Journal of Combinatorial Theory (B)* 17, 48–50.

- Burgess2008** Burgess, R. and Z. Yang (2008). Estimation of hominoid ancestral population sizes under bayesian coalescent models incorporating mutation rate variation and sequencing errors. *Molecular Biology and Evolution* 25(9), 1979–1994.
- Cardona12008** Cardona, G., F. Rossell, and G. Valiente (2008). Extended Newick: it is time for a standard representation of phylogenetic networks. *BMC Bioinformatics* 9(532-540).
- Chen2001** Chen, F. C. and W. H. Li (2001). Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *American Journal of Human Genetics* 68, 444–456.
- Chen2009macs** Chen, G. K., P. Marjoram, and J. D. Wall (2009). Fast and flexible simulation of dna sequence data. *Genome Res.* 19, 136–142.
- Colless1982** Colless, D. H. (1982). Review of “Phylogenetics: The theory and practice of phylogenetic systematics”. *Systematic Zoology* 31, 100–104.
- Cummings2008** Cummings, M. P., M. C. Neel, and K. L. Shaw (2008). A genealogical approach to quantifying lineage divergence. *Evolution* 62(9), 2411–2422.
- Darlu2002** Darlu, P. and G. Lecointre (2002). When does the incongruence length difference test fail? *Molecular Biology and Evolution* 19, 432–437.
- Daubin2004** Daubin, V. and H. Ochman (2004). Quartet mapping and the extent of lateral transfer in bacterial genomes. *Molecular Biology and Evolution* 1, 86–89.
- deQuieroz2007** de Quieroz, K. (2007). Species concepts and species delimitation. *Systematic Biology* 56, 879–886.
- deQuieroz1990** de Quieroz, K. and J. Gauthier (1990). Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systematic Zoology* 39(4), 307–322.
- Degnan2010** Degnan, J. H. (2010). Probabilities of gene tree topologies with intraspecific sampling given a species tree. In L. L. Knowles and L. S. Kubatko (Eds.), *Estimating Species Trees, Practical and Theoretical Aspects*, pp. 53–77. Hoboken, NJ: Wiley-Blackwell.
- Degnan2009** Degnan, J. H., M. DeGiorgio, D. Bryant, and N. A. Rosenberg (2009). Properties of consensus methods for inferring species trees from gene trees. *Systematic biology* 58(1), 35–54.
- Degnan2006** Degnan, J. H. and N. A. Rosenberg (2006). Discordance of species trees with their most likely gene trees. *PLoS Genetics* 2, 762–768.
- Degnan2005** Degnan, J. H. and L. A. Salter (2005). Gene tree distributions under the coalescent process. *Evolution* 59, 24–37.

- Desper2004** Desper, R. and O. Gascuel (2004). Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. *Molecular Biology and Evolution* 21, 587–598.
- DiltheyMHC** Dilthey, A., C. Cox, Z. Iqbal, M. R. Nelson, and G. McVean (0000). Improved genome inference in the mhc using a population reference graph. *unknown* 0(0), 0.
- Dilthey2013** Dilthey, A., S. Leslie, L. Moutsianas, J. Shen, C. Cox, M. R. Nelson, and G. McVean (2013, 02). Multi-population classical hla type imputation. *PLoS Comput Biol* 9(2), e1002877.
- Donnelly1999** Donnelly, P. and T. Kurtz (1999). Particle representations for measure-valued population models. *The Annals of Probability* 27(1), 166–205.
- Doucet2007** Doucet, A. (2007). Sequential Monte Carlo methods lectures at Machine Learning Summer School (MLSS), Tübingen 2007. <http://videlectures.net/mlss07-doucet.smcm/>.
- Doucet2001** Doucet, A., N. de Freitas, and N. Gordon (2001). An introduction to sequential Monte Carlo methods. In A. Doucet, N. de Freitas, and N. Gordon (Eds.), *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag.
- Doucet2000** Doucet, A., S. Godsill, and C. Andrieu (2000). On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing* 10, 197–208.
- Doucet2008** Doucet, A. and A. M. Johansen (2008). A tutorial on particle filtering and smoothing: Fifteen years later. www.cs.ubc.ca/~arnaud/doucet_johansen_tutorialPF.pdf.
- Doyle1997** Doyle, J. J. (1997). Trees within trees: Genes and species, molecules and morphology. *Systematic Biology* 46(3), 537–553.
- Dress86** Dress, A., A. von Haeseler, and M. Krueger (1986). Reconstructing phylogenetic trees using variants of the “four-point condition”. *Studien zur Klassifikation* 17, 299–305.
- Drummond2007beast** Drummond, A. and A. Rambaut (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7(1), 214.
- Drummond2005** Drummond, A. J., A. Rambaut, B. Shapiro, and O. G. Pybus (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution* 22(5), 1185–1192.
- Drummond2012Beast** Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* 29(8), 1969–1973.

- Durbin2014** Durbin, R. (2014). Efficient haplotype matching and storage using the positional BurrowsWheeler transform (PBWT). *Bioinformatics* 30(9), 1266–1272.
- Dutheil2009** Dutheil, J. Y., G. Ganapathy, A. Hobolth, T. Mailund, M. K. Uyenoyama, and M. H. Schierup (2009). Ancestral population genomics: the coalescent hidden Markov model approach. *Genetics* 183(1), 259–274.
- Edwards2007** Edwards, S., L. Liu, and D. K. Pearl (2007). High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences of the United States of America* 104(14), 5936–5941.
- Eldon2011** Eldon, B. (2011). Estimation of parameters in large offspring number models and ratios of coalescence times. *Theoretical Population Biology* 80, 16–28.
- Eldon2015** Eldon, B., M. Birkner, J. Blath, and F. Freund (2015). Can the site-frequency spectrum distinguish exponential population growth from multiple-merger coalescents? *Genetics* 199(3), 841–856.
- Eldon2011slides** Eldon, B. and J. H. Degnan (2011). Concordance between species trees and gene genealogies with asynchronous multiple mergers. Slides, www.newton.ac.uk/programmes/PLG/seminars/062016301.pdf.
- Eldon2012** Eldon, B. and J. H. Degnan (2012). Multiple merger gene genealogies in two species: monophyly, paraphyly, and polyphyly for two examples of Lambda coalescents. *Theoretical Population Biology* 82, 117–130.
- Eldon2006** Eldon, B. and J. Wakeley (2006). Coalescent processes when the distribution of offspring number among individuals is highly skewed. *Genetics* 172, 2621–2633.
- Eldon2008** Eldon, B. and J. Wakeley (2008). Linkage disequilibrium under skewed offspring distribution among individuals in a population. *Genetics* 178, 1517–1532.
- Eriksson2009** Eriksson, A., B. Mahjani, and B. Mehlig (2009). Sequential markov coalescent algorithms for population models with demographic structure. *Theoretical Population Biology* 76(2), 84 – 91.
- Ewing2006** Ewing, G. and A. Rodrigo (May 2006). Coalescent-based estimation of population parameters when the number of demes changes over time. *Molecular Biology and Evolution* 23(5), 988–996.
- Excoffier2011** Excoffier, L. and M. Foll (2011). fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics* 27(9), 1332–1334.
- Excoffier2000** Excoffier, L., J. Novembre, and S. Schneider (2000). Computer note. simcoal: a general coalescent program for the simulation of molecular data in interconnected populations with arbitrary demography. *Journal of Heredity* 91(6), 506–509.

Felsensteinlist	Felsenstein, J. Phylogeny programs. http://evolution.genetics.washington.edu/phylip/software.html .
Felsenstein1978misleading	Felsenstein, J. (1978a). Cases in which parsimony or compatibility methods will be positively misleading. <i>Systematic Biology</i> 27, 401–410.
Felsentein1978	Felsenstein, J. (1978b). The number of evolutionary trees. <i>Systematic Zoology</i> 27, 27–33.
Felsenstein1981	Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. <i>Journal of Molecular Evolution</i> 17, 368–376.
Felsenstein1993	Felsenstein, J. (1993). <i>PHYLIP (PHYLogeny Inference Package) version 3.6a2</i> . Department of Genetics, University of Washington, Seattle: Distributed by the author.
Felsentein2004	Felsenstein, J. (2004). <i>Inferring Phylogenies</i> . Sunderland, MA: Sinauer Associates.
Ferretti2013	Ferretti, L., F. Disanto, and T. Wiehe (2013). The effect of single recombination events on colaescent tree height and shape. <i>PLOS ONE</i> 8(4), e60123. doi:10.1371/journal.pone.0060123.
Fisher1930	Fisher, R. A. (1930). <i>The genetical theory of natural selection</i> . Oxford: Clarendon Press.
Fitch1971	Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specific tree topology. <i>Systematic Zoology</i> 20(4), 406–416.
Galtier2004	Galtier, N. and A. Jean-marie (2004). Markov-modulated markov chains and the covarion process of molecular evolution. <i>Journal of Computational Biology</i> 11(4), 727–733.
Zhang2009	Gaoyu, Z., L. Qiongfei, L. Qing, and Z. Zhizhao (2009). High frequency financial time series forecasting via particle filtering. In <i>Information Management, Innovation Management and Industrial Engineering, 2009 International Conference on</i> , Volume 4, pp. 62–65.
Gascuel1997	Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. <i>Molecular Biology and Evolution</i> 14(7), 685–695.
Gerard2011	Gerard, D., H. L. Gibbs, and L. Kubatko (2011). Estimating hybridization in the presence of coalescence using phylogenetic intraspecific sampling. <i>BMC Evolutionary Biology</i> 11.
Gernhard2008	Gernhard, T., K. Hartmann, and M. A. Steel (2008). Stochastic properties of generalised yule models, with biodiversity applications. <i>Mathematical Biology</i> 57, 713–735.

- Gordon1993** Gordon, N., D. Salmond, and A. F. M. Smith (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *Radar and Signal Processing, IEE Proceedings F* 140(2), 107–113.
- Gravel05072011** Gravel, S., B. M. Henn, R. N. Gutenkunst, A. R. Indap, G. T. Marth, A. G. Clark, F. Yu, R. A. Gibbs, T. . G. Project, and C. D. Bustamante (2011). Demographic history and rare allele sharing among human populations. *Proceedings of the National Academy of Sciences*.
- Green2010** Green, R. E., J. Krause, A. W. Briggs, T. Maricic, U. Stenzel, M. Kircher, N. Patterson, H. Li, W. Zhai, M. H.-Y. Fritz, N. F. Hansen, E. Y. Durand, A.-S. Malaspinas, J. D. Jensen, T. Marques-Bonet, C. Alkan, K. Prfer, M. Meyer, H. A. Burbano, J. M. Good, R. Schultz, A. Aximu-Petri, A. Butthof, B. Hber, B. Hffner, M. Siegemund, A. Weihmann, C. Nusbaum, E. S. Lander, C. Russ, N. Novod, J. Affourtit, M. Egholm, C. Verna, P. Rudan, D. Brajkovic, . Kucan, I. Guic, V. B. Doronichev, L. V. Golovanova, C. Lalueza-Fox, M. de la Rasilla, J. Fortea, A. Rosas, R. W. Schmitz, P. L. F. Johnson, E. E. Eichler, D. Falush, E. Birney, J. C. Mullikin, M. Slatkin, R. Nielsen, J. Kelso, M. Lachmann, D. Reich, and S. Pbo (2010). A draft sequence of the neandertal genome. *Science* 328(5979), 710–722.
- Griffiths1981** Griffiths, R. C. (1981). Neutral two-locus multiple allele models with recombination. *Theoretical Population Biology* 19(2), 169 – 186.
- Griffiths1991** Griffiths, R. C. (1991). The two-locus ancestral graph. *Lecture Notes-Monograph Series* 18, 100–117.
- Griffiths1997ARG** Griffiths, R. C. and P. Marjoram (1997). An ancestral recombination graph. In *Progress in Population Genetics and Human Evolution (Minneapolis, MN, 1994)*, Volume 87 of *IMA Vol. Math. Appl.*, pp. 257–270. New York: Springer.
- Griffiths2003** Griffiths, R. C. and S. Tavaré (2003). The genealogy of a neutral mutation. In P. J. Green, N. L. Hjort, and S. Richardson (Eds.), *Highly Structured Stochastic Systems*, Volume 27 of *Oxford Statistical Science Series*, pp. 393–413. Oxford University Press.
- Gronau2011** Gronau, I., M. J. Hubisz, B. Gulko, C. G. Danko, and A. Siepel (2011a). *Nature Genetics* 43, 1031 – 1034.
- Ilan2011** Gronau, I., M. J. Hubisz, B. Gulko, C. G. Danko, and A. Siepel (2011b). Bayesian inference of ancient human demography from individual genome sequences. *Nature Genetics* 43, 1031–1034.
- Guigo1996** Guigo, R., I. Muchnik, and T. F. Smith (1996). Reconstruction of ancient molecular phylogeny. *Molecular Phylogenetics and Evolution* 6(2), 189 – 213.
- Guindon2010** Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology* 59(3), 307–321.

- Guindon2003** Guindon, S. and O. Gascuel (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* 52(5), 696–704.
- Gusev2009** Gusev, A., J. K. Lowe, M. Stoffel, M. J. Daly, D. Altshuler, J. L. Breslow, J. M. Friedman, and I. Pe’er (2009). Whole population, genome-wide mapping of hidden relatedness. *Genome Research* 19(2), 318–326.
- Gutenkunst2009dadi** Gutenkunst, R. N., R. D. Hernandez, S. H. Williamson, and C. D. Bustamante (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet* 5(10), e1000695.
- Harding1971** Harding, E. F. (1971). The probabilities of rooted tree-shapes generated by random bifurcation. *Advances in Applied Probability* 3(1), 44–77.
- Harris2013** Harris, K. and R. Nielsen (2013). Inferring demographic history from a spectrum of shared haplotype lengths. *PLoS Genet* 9(6), e1003521.
- Hasegawa1991** Hasegawa, M., H. Kishino, and N. Saitou (1991). On the maximum likelihood method in molecular phylogenetics. *Journal of Molecular Evolution* 32, 443–445.
- Heard1992** Heard, S. B. (1992). Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. *Evolution* 46(6), 1818–1826.
- Hedgecock1994** Hedgecock, D. (1994). Does variance in reproductive success limit effective population sizes of marine organisms? In A. Beaumont (Ed.), *Genetics and Evolution of Aquatic Organisms*, pp. 1222–1344. London: Chapman and Hall.
- Hedgecock2011** Hedgecock, D. and A. I. Pudovkin (2011). Sweepstakes reproductive success in highly fecund marine fish and shellfish: a review and commentary. *Bull Mar Sci* 87, 971–1002.
- Hedgecock1982** Hedgecock, D., M. Tracey, and K. Nelson (1982). Genetics. In L. Abele (Ed.), *The Biology of Crustacea*, Volume 2, pp. 297–403. New York: Academic Press.
- Hein2005** Hein, J., M. H. Schierup, and C. Wiuf (2005). *Gene Genealogies, Variation and Evolution : A Primer in Coalescent Theory*. Oxford University Press, Oxford, UK.
- Heled2010** Heled, J. and A. Drummond (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27(3), 570–580.
- Hillis1992** Hillis, D. M. and J. P. Huelsenbeck (1992). Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* 83, 189–195.
- Ho2011** Ho, S. Y. and B. Shapiro (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources* 11(3), 423–434.

- Hoeffding1963** Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* 58(301), 13–30.
- Holder2003** Holder, M. and P. O. Lweis (2003). Phylogeny estimation: traditional and bayesian approaches. *Nature reviews Genetics* 4, 275–284.
- Holland2003** Holland, B. and V. Moulton (2003). Consensus networks: A method for visualising incompatibilities in collections of trees. In G. Benson and R. Page (Eds.), *Algorithms in Bioinformatics*, Volume 2812 of *Lecture Notes in Computer Science*, pp. 165–176. Springer Berlin Heidelberg.
- Holland2008** Holland, B. R., S. Benthin, P. J. Lockhart, V. Moulton, and K. T. Huber (2008). Using supernetworks to distinguish hybridization from lineage-sorting. *BMC Evolutionary Biology* 8, 202–213.
- Holland2004** Holland, B. R., K. T. Huber, V. Moulton, and P. J. Lockhart (2004). Using consensus networks to visualize contradictory evidence for species phylogeny. *Molecular Biology and Evolution* 21(7), 1459–1461.
- Holman2005** Holman, E. W. (2005). Nodes in phylogenetic trees: The relation between imbalance and number of descendent species. *Systematic Biology* 54(6), 895–899.
- Huber2006** Huber, K. and V. Moulton (2006). Phylogenetic networks from multi-labelled trees. *Journal of Mathematical Biology* 52(5), 613–632.
- Hudson1983TPB** Hudson, R. R. (1983a). Properties of a neutral allele model with intragenic recombination. *Theoretical Population Biology* 23, 183–201.
- Hudson1983Evo1** Hudson, R. R. (1983b). Testing the constant-rate3 neutral allele model with protein sequence data. *Evolution* 37, 203–217.
- Hudson1990** Hudson, R. R. (1990). Gene genealogies and the coalescent process. *Oxford Surveys Evolution Biology* 7, 1–44.
- Hudson2002ms** Hudson, R. R. (2002). Generating samples under a Wright–Fisher neutral model. *Bioinformatics* 18, 337–338.
- Hudson2002** Hudson, R. R. and J. A. Coyne (2002). Mathematical consequences of the genealogical species concept. *Evolution* 56(8), 1557–1565.
- Hudson1985** Hudson, R. R. and N. L. Kaplan (1985). Statistical properties of the number of recombination events in the hisotry of a sample of DNA sequences. *Genetics* 111, 147–164.
- Huelsensbeck2001 mrbays** Huelsenbeck, J. P. and F. Ronquist (2001). MrBayes: Bayesian inference of phylogeny. *Bioinformatics* 17, 754–755.
- Huson2010** Huson, D., R. Rupp, and C. Scornavacca (2010). *Phylogenetic Networks: Concepts, Algorithms and Applications*. Cambridge, UK: Cambridge University Press.

- Huson2007Splits** Huson, D. H. and D. Bryant (2007). Contents user manual for SplitsTree4 V4.8.
- Huson2004** Huson, D. H., T. Dezulian, T. Klopper, and M. A. Steel (2004). Phylogenetic super-networks from partial trees. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 1, 151–158.
- Huson2005** Huson, D. H., T. Klopper, P. J. Lockhart, and M. A. Steel (2005). Reconstruction of reticulate networks from gene trees. *Research in Computational Molecular Biology 3500*, 233–249.
- Huson2007Dendro** Huson, D. H., D. C. Richter, C. Rausch, T. Dezulian, M. Franz, and R. Rupp (2007). Dendroscope: An interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8(1), 460.
- Huynh2005** Huynh, T. N. D., J. Jansson, N. B. Nguyen, and W.-K. Sung (2005). Constructing a smallest refining galled phylogenetic network. *Annual International Conference on Research in Computational Molecular Biology 2005, LNBI 3500*, 265–280.
- Degnan2012** James H. Degnan, Noah A. Rosenberg, T. S. (2012). The probability distribution of ranked gene trees on a species tree. *Mathematical Biosciences* (235), 45–55.
- Jiggins2008** Jiggins, C. D., C. Salazar, M. Linares, and J. Mavarez (2008). Hybrid trait speciation and heliconius butterflies. *Philosophical Transactions of the Royal Society of London. Series B* 363, 3047–3054.
- Jack2015** John D. O’Brien, Zamin Iqbal, L. A.-E. (2015). An integrative statistical model for inferring strain admixture within clinical plasmodium falciparum isolates.
- Joly2009** Joly, S., P. A. McLenachan, and P. J. Lockhart (2009). A statistical approach for distinguishing hybridization and incomplete lineage sorting. *The American Naturalist* 174, 54–70.
- Jones2011** Jones, G. R. (2011). Tree models for macroevolution and phylogenetic analysis. *Systematic Biology* 60(6), 735–746.
- Jordan1869** Jordan, C. (1869). Sur les assemblages des lignes. *Journal für die Reine und Angewandte Mathematik* 70, 185–190.
- JC69** Jukes, T. H. and C. R. Cantor (1969). Evolution of protein molecules. In H. N. Munro (Ed.), *Mammalian protein metabolism*, pp. 21–123. New York: Academic Press.
- Kalinowski2009** Kalinowski, S. (2009). How well do evolutionary trees describe genetic relationships among populations? *Heredity* 102, 506–513.
- Kang1975** Kang, A. N. C. and D. A. Ault (1975). Some properties of a centroid of a free tree. *Information Processing Letters* 4(1), 18–20.
- Karr1993** Karr, A. F. (1993). *Probability*. New York: Springer-Verlag.

- Kerr2013** Kerr, A. M. (2013). *Holothuria (semperothuria) roseomaculata* n. sp. (aspidochirotrida: Holothuriidae), a coral-reef inhabiting sea cucumber from the western pacific ocean. *Zootaxa* 3641(4), 384–394.
- Kim1993** Kim, J. (1993, Sep). Improving the accuracy of phylogenetic estimation by combining different methods. *Systematic Biology* 42(3), 331–340.
- Kim1993nj** Kim, J., F. J. Rohlf, and R. R. Sokal (1993, Apr). The accuracy of phylogenetic estimation using the neighbor-joining method. *Evolution* 47(2), 471–486.
- Kimura1971** Kimura, M. (1971). Theoretical foundation of population genetics at the molecular level. *Theoretical Population Biology* 2(2), 174 – 208.
- Kingman1982** Kingman, J. F. C. (1982). On the genealogy of large populations. *Journal of Applied Probability* 19, 27–43.
- Kirkpatrick1993** Kirkpatrick, M. and M. Slatkin (1993). Searching for evolutionary patterns in the shape of a phylogenetic tree. *Evolution* 47(4), 1171–1181.
- Kotz2000** Kotz, S., N. Balakrishnan, and N. L. Johnson (2000). *Continuous Multivariate Distributions, Volume 1, Models and Applications* (2 ed.). New York: Wiley.
- Kubatko2009** Kubatko, L. S. (2009). Identifying hybridization events in the presence of coalescence via model selection. *Systematic Biology* 58, 478–488.
- Kubatko2009stem** Kubatko, L. S., B. C. Carstens, and L. L. Knowles (2009). STEM: species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics* 25(7), 971–973.
- Kubatko2007** Kubatko, L. S. and J. Degnan (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology* 56, 17–24.
- Lapointe2010** Lapointe, F.-J., P. Lopez, Y. Boucher, J. Koenig, and E. Baptiste (2010). Clanistics: a multi-level perspective for harvesting unrooted gene trees. *Trends in Microbiology* 18, 341–347.
- Lawson2012** Lawson, D. J., G. Hellenthal, S. Myers, and D. Falush (2012, 01). Inference of population structure using dense haplotype data. *PLoS Genet* 8(1), e1002453.
- Leigh2008** Leigh, J. W., E. Susko, M. Baumgartner, and A. J. Roger (2008). Testing congruence in phylogenomic analysis. *Systematic Biology* 57(1), 104–115.
- Li2006** Li, H., A. Coghlan, J. Ruan, L. J. Coin, J.-K. Heriche, L. Osmotherly, R. Li, T. Liu, Z. Zhang, L. Bolund, G. K.-S. Wong, W. Zheng, P. Dehal, J. Wang, and R. Durbin (2006). Treefam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Research* 34, 572–580.
- Li2011** Li, H. and R. Durbin (2011). Inference of human population history from individual whole-genome sequences. *Nature* 475, 493–496.

- Li2003** Li, N. and M. Stephens (2003, December). Modeling Linkage Disequilibrium and Identifying Recombination Hotspots Using Single-Nucleotide Polymorphism Data. *Genetics* 165(4), 2213–2233.
- Li2000** Li, S., D. K. Pearl, and H. Doss (2000). Phylogenetic tree construction using Markov Chain Monte Carlo. *Journal of the American Statistical Association* 95(450), 493–508.
- Liu1998** Liu, J. S., R. Chen, and W. H. Wong (1998). Rejection control and sequential importance sampling. *Journal of the American Statistical Association* 93(443), 1022–1031.
- LiuPhD** Liu, L. (2006). *Reconstructing posterior distributions of a species phylogeny using estimated gene tree distributions*. Ph. D. thesis, Ohio State University.
- Liu2008** Liu, L. (2008). BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics* 24, 2542–2543.
- Liu2007** Liu, L. and D. K. Pearl (2007). Species trees from gene trees: Reconstructing bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* 56(3), 504–514.
- Liu2008DNA** Liu, L., D. K. Pearl, R. T. Brumfield, and S. V. Edwards (2008). Estimating species trees using multiple allele DNA sequence data. *Evolution* 62(8), 2080–2091.
- Liu2010Phybase** Liu, L. and L. Yu (2010). Phybase: an R package for species tree analysis. *Bioinformatics* 26(7), 962–963.
- Liu2009review** Liu, L., L. Yu, L. Kubatko, D. K. Pearl, and S. V. Edwards (2009). Coalescent methods for estimating phylogenetic trees. *Molecular Phylogenetics and Evolution* 53(1), 320–328.
- Liu2010MT** Liu, L., L. Yu, and D. K. Pearl (2010). Maximum tree: a consistent estimator of the species tree. *Journal of Mathematical Biology* 60, 95–106.
- Liu2009Sys** Liu, L., L. Yu, D. K. Pearl, and S. V. Edwards (2009). Estimating species phylogenies using coalescence times among sequences. *Systematic Biology* 58(5), 468–477.
- Locke2011** Locke, D. P., L. W. Hillier, W. C. Warren, K. C. Worley, L. V. Nazareth, D. M. Muzny, S.-P. Yang, Z. Wang, A. T. Chinwalla, P. Minx, M. Mitreva, L. Cook, K. D. Delehaunty, C. Fronick, H. Schmidt, L. A. Fulton, R. S. Fulton, J. O. Nelson, V. Magrini, C. Pohl, T. A. Graves, C. Markovic, A. Cree, H. H. Dinh, J. Hume, C. L. Kovar, G. R. Fowler, G. Lunter, S. Meader, A. Heger, C. P. Ponting, T. Marques-Bonet, C. Alkan, L. Chen, Z. Cheng, J. M. Kidd, E. E. Eichler, S. White, S. Searle, A. J. Vilella, Y. Chen, P. Flicek, J. Ma, B. Raney, B. Suh, R. Burhans, J. Herrero, D. Haussler, R. Faria, O. Fernando, F. Darr, D. Farr, E. Gazave, M. Oliva, A. Navarro, R. Roberto, O. Capozzi, N. Archidiacono, G. D. Valle, S. Purgato, M. Rocchi, M. K. Konkel, J. A.

Walker, B. Ullmer, M. A. Batzer, A. F. A. Smit, R. Hubley, C. Casola, D. R. Schrider, M. W. Hahn, V. Quesada, X. S. Puente, G. R. Ordoez, C. Lpez-Otn, T. Vinar, B. Brejova, A. Ratan, R. S. H. and Webb Miller, C. Kosiol, H. A. Lawson, V. Taliwal, A. L. Martins, A. Siepel, A. RoyChoudhury, X. Ma, J. Degenhardt, C. D. Bustamante, R. N. Gutenkunst, T. Mailund, J. Y. Dutheil, A. Hobolth, M. H. Schierup, O. A. Ryder, Y. Yoshinaga, P. J. de Jong, G. M. Weinstock, J. Rogers, E. R. Mardis, R. A. Gibbs, and R. K. Wilson (2011). Comparative and demographic analysis of orang-utan genomes. *Nature* 469, 529–533.

Lohmueller2009 Lohmueller, K. E., C. D. Bustamante, and A. G. Clark (2009). Methods for human demographic inference using haplotype patterns from genomewide single-nucleotide polymorphism data. *Genetics* 182(1), 217–231.

Lopes2009 Lopes, J. S., D. Balding, and M. A. Beaumont (2009). Popabc: a program to infer historical demographic parameters. *Bioinformatics* 25(20), 2747–2749.

Lunter2012SMBE Lunter, G. (2012). A new Markovian approximation to the coalescent with recombination. Society for Molecular Biology & Evolution. http://imgpublic.mci-group.com/ie/PCO/AllAbstracts_FINAL.pdf. Aug 14, 2013.

Lunter2011 Lunter, G. and M. Goodson (2011). Stampy: A statistical algorithm for sensitive and fast mapping of illumina sequence reads. *Genome Research* 21(6), 936–939.

Ma2008 Ma, J., A. Ratan, B. J. Raney, B. B. Suh, W. Miller, and D. Haussler (2008). The infinite sites model of genome evolution. *Proceedings of the National Academy of Sciences of the United States of America* 105(38), 14254–14261.

Maddison1997nexus Maddison, D. R., D. L. Swofford, and W. P. Maddison (1997). NEXUS: An extensible file format for systematic information. *Systematic Biology* 46(4), 590–621.

Maddison1997 Maddison, W. P. (1997). Gene trees in species trees. *Systematic Biology* 46(3), 523–536.

Maddison2006 Maddison, W. P. and L. L. Knowles (2006). Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* 55(1), 21–30.

Mahmoud2008 Mahmoud, H. (2008). *Pólya Urn Models*. Boca Raton: Chapman and Hall/CRC.

Mailund2011 Mailund, T., J. Y. Dutheil, A. Hobolth, G. Lunter, and M. H. Schierup (2011). Estimating Divergence Time and Ancestral Effective Population Size of Bornean and Sumatran Orangutan Subspecies Using a Coalescent Hidden Markov Model. *PLoS Genet* 7(3), e1001319.

Mailund2012 Mailund, T., A. E. Halager, M. Westergaard, J. Y. Dutheil, K. Munch, L. N. Andersen, G. Lunter, K. Prüfer, A. Scally, A. Hobolth, and M. H. Schierup. A new isolation with migration model along complete genomes infers very

different divergence processes among closely related great ape species. *PLOS Genetics* 8(12), doi:10.1371/journal.pgen.1003125.

Mallet2007 Mallet, J., M. Beltrán, W. Neukirchen, and M. Linares (2007). Natural hybridization in heliconiine butterflies: the species boundary as a continuum. *BMC Evolutionary Biology* 7, 28.

Menske2012 Manske, M., O. Miotto, S. Campino, S. Auburn, J. Almagro-Garcia, G. Maslen, J. O’Brien, A. Djimde, O. Doumbo, I. Zongo, J. Ouedraogo, P. Michon, I. Mueller, P. Siba, A. Nzila, S. Borrmann, S. Kiara, K. Marsh, H. Jiang, X. Su, C. Amaratunga, R. Fairhurst, D. Socheat, F. Nosten, M. Imwong, N. White, M. Sanders, E. Anastasi, D. Alcock, E. Drury, S. Oyola, M. Quail, D. Turner, V. Ruano-Rubio, D. Jyothi, L. Amenga-Etego, C. Hubbart, A. Jeffreys, K. Rowlands, C. Sutherland, C. Roper, V. Mangano, D. Modiano, J. Tan, M. Ferdig, A. Amambua-Ngwa, D. Conway, S. Takala-Harrison, C. Plowe, J. Rayner, K. Rockett, T. Clark, C. Newbold, M. Berriman, B. MacInnis, and D. Kwiatkowski (2012, 7). Analysis of plasmodium falciparum diversity in natural infections by deep sequencing. *Nature* 487(7407), 375–379.

Margush1981 Margush, T. and F. R. McMorris (1981). Consensus n-trees. *Bulletin of Mathematical Biology* 43, 239–244.

Marjoram2006Review Marjoram, P. and S. Tavaré (2006). Modern computational approaches for analysing molecular genetic variation data. *Nature Reviews Genetics* 7(10), 759–770.

Marjoram2006 Marjoram, P. and J. Wall (2006). Fast “coalescent” simulation. *BMC Genetics* 7(1), 16.

Marshall12011 Marshall, D. C., K. B. R. Hill, J. R. Cooley, and C. Simon (2011). Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: Lessons from New Zealand cicadas (Genus Kikihia). *Systematic Biology* 60, 482–502.

Marshall12008 Marshall, D. C., K. Slon, J. R. Cooley, K. B. Hill, and C. Simon (2008). Steady plio-pleistocene diversification and a 2-million-year sympatry threshold in a New Zealand cicada radiation. *Molecular Phylogenetics and Evolution* 48, 1054–1066.

Mathew2013 Mathew, L. A., P. R. Staab, L. E. Rose, and D. Metzler (2013). Why to account for finite sites in population genetic studies and how to do this with jaatha 2.0. *Ecology and Evolution*, n/a–n/a.

Mathieson2014 Mathieson, I. and G. McVean (2014, 08). Demography and the age of rare variants. *PLoS Genet* 10(8), e1004528.

Matsumoto1998 Matsumoto, M. and T. Nishimura (1998). Mersenne Twister: A 623-dimensionally equidistributed uniform Pseudo-Random number generator. *ACM Transactions on Modeling and Computer Simulation* 8(1), 3–30.

- Mau1999** Mau, B., M. A. Newton, and B. Larget (1999). Bayesian phylogenetic inference via markov chain monte carlo methods. *Biometrics* 55(1), 1–12.
- McKenziethesis** McKenzie, A. (2000). *Stochastic Speciation Models for Evolutionary Trees*. Ph. D. thesis, University of Canterbury, Christchurch.
- McKenzie2000** McKenzie, A. and M. A. Steel (2000). Distributions of cherries for two models of trees. *Mathematical Biosciences* 164, 81–92.
- McVean2005** McVean, G. A. and N. J. Cardin (2005). Approximating the coalescent with recombination. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360(1459), 1387–1393.
- mcvean_approximating_2005** McVean, G. A. T. and N. J. Cardin (2005, July). Approximating the coalescent with recombination. *Phil. Trans. R. Soc. B* 360(1459), 1387–1393.
- Meng2009** Meng, C. and L. S. Kubatko (2009). Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: a model. *Theoretical Population Biology* 75, 35–45.
- Mihaylova2004** Mihaylova, L. and R. Boel (2004). A particle filter for freeway traffic estimation. In *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, Volume 2, pp. 2106–2111 Vol.2.
- Miller2012** Miller, W., S. C. Schuster, A. J. Welch, A. Ratan, O. C. Bedoya-Reina, F. Zhao, H. L. Kim, R. C. Burhans, D. I. Drautz, N. E. Wittekindt, L. P. Tomsho, E. Ibarra-Laclette, L. Herrera-Estrella, E. Peacock, S. Farley, G. K. Sage, K. Rode, M. Obbard, R. Montiel, L. Bachmann, . Inglfsson, J. Aars, T. Mailund, . Wiig, S. L. Talbot, and C. Lindqvist (2012). Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proceedings of the National Academy of Sciences*.
- Mitchell1978** Mitchell, S. L. (1978). Another characterization of the centroid of a tree. *Discrete Mathematics* 24, 277–280.
- mooers97a** Mooers, A. O. and S. B. Heard (1997). Evolutionary process from phylogenetic tree shape. *Quarterly Review of Biology* 72, 31–54.
- Morgan2009** Morgan-Richards, M., R. D. Smissen, L. D. Shepherd, G. P. Wallis, J. J. Hayward, C. hang Chan, G. K. Chambers, and H. M. Chapman (2009, March). A review of genetic analyses of hybridisation in New Zealand. *Journal of the Royal Society of New Zealand* 39(1), 15–34.
- Mossel2008** Mossel, E. and S. Roch (2010). Incomplete lineage sorting: consistent phylogeny estimation from multiple loci. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 7, 166–171.
- Naduvilezhath2011** NADUVILEZHATH, L., L. E. ROSE, and D. METZLER (2011). Jaatha: a fast composite-likelihood approach to estimate demographic parameters. *Molecular Ecology* 20(13), 2709–2723.

- Nei1987** Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Nichols2001** Nichols, R. (2001, July). Gene trees and species trees are not the same. *Trends in Ecology and Evolution* 16(7).
- Nieselt2001** Nieselt-Struwe, K. and A. von Haeseler (2001). Quartet-mapping, a generalization of the likelihood-mapping procedure. *Molecular Biology and Evolution* 7(18), 1204–1219.
- nordborg98a** Nordborg, M. (1998). On the probability of neanderthal ancestry. *American Journal of Human Genetics* 63, 1237–1240.
- nordborg01a** Nordborg, M. (2001). Coalescent theory. In D. J. Balding, M. Bishop, and C. Cannings (Eds.), *Handbook of Statistical Genetics*, Chapter 7, pp. 179–212. Chichester, UK: Wiley.
- Olsen1990** Olsen, G. (1990). Gary Olsen’s interpretation of the “Newick’s 8:45” tree format standard. <http://evolution.genetics.washington.edu/phylip/newick.doc.html>. Feb 21, 2013.
- Omeara2010** O’Meara, B. C. (2010). New heuristic methods for joint species delimitation and species tree inference. *Systematic Biology* 59(1), 59–73.
- Omeara2006** O’Meara, B. C. and P. C. W. Ccile An, Michael J. Sanderson (2006). Testing for different rates of continuous trait evolution using likelihood. *Evolution* 60(5), 922–933.
- Pamilo1988** Pamilo, P. and M. Nei (1988). Relationships between gene trees and species trees. *Molecular Biology and Evolution* 5, 568–583.
- Paradis2004** Paradis, E., J. Claude, and K. Strimmer (2004). Ape: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290.
- Pearson2016** Pearson, R. D., R. Amato, S. Auburn, O. Miotto, J. Almagro-Garcia, C. Amaratunga, S. Suon, S. Mao, R. Noviyanti, H. Trimarsanto, J. Marfurt, N. M. Anstey, T. William, M. F. Boni, C. Dolecek, H. T. Tran, N. J. White, P. Michon, P. Siba, L. Tavul, G. Harrison, A. Barry, I. Mueller, M. U. Ferreira, N. Karunaweera, M. Randrianarivelojosia, Q. Gao, C. Hubbart, L. Hart, B. Jeffery, E. Drury, D. Mead, M. Kekre, S. Campino, M. Manske, V. J. Cornelius, B. MacInnis, K. A. Rockett, A. Miles, J. C. Rayner, R. M. Fairhurst, F. Nosten, R. N. Price, and D. P. Kwiatkowski (2016, June). Genomic analysis of local variation and recent evolution in *Plasmodium vivax*. *Nat Genet* 48, 959–964.
- Pinelis2003** Pinelis, I. (2003). Evolutionary models of phylogenetic trees. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270(1522), 1425–1431.
- Pitman1999** Pitman, J. (1999). Coalescents with multiple collisions. *The Annals of Probability* 27(4), 1870–1902.

- Poe1998** Poe, S. (1998, March). Sensitivity of phylogeny estimation to taxonomic sampling. *Systematic Biology* 47(1), 18–31.
- Poe1999** Poe, S. and D. L. Swofford (1999, March). Taxon sampling revisited. *Nature* 398, 299–300.
- Pybus2000** Pybus, O. G., A. Rambaut, and P. H. Harvey (2000). An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics* 155(3), 1429–1437.
- Rmanual** R Core Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.
- Rannala1996** Rannala, B. and Z. Yang (1996). Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution* 43, 304–311.
- Rannala2003** Rannala, B. and Z. Yang (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164, 1645–1656.
- Ranwez2001** Ranwez, V. and O. Gascuel (2001). Quartet-based phylogenetic inference: Improvements and limits. *Molecular Biology and Evolution* 18(6), 1103–1116.
- Reich2010** Reich, D., R. E. Green, M. Kircher, J. Krause, N. Patterson, E. Y. Durand, B. Viola, A. W. Briggs, U. Stenzel, P. L. F. Johnson, T. Maricic, J. M. Good, T. Marques-Bonet, C. Alkan, Q. Fu, S. Mallick, H. Li, M. Meyer, E. E. Eichler, M. Stoneking, M. Richards, S. Talamo, M. V. Shunkov, A. P. Derevianko, J.-J. Hublin, J. Kelso, M. Slatkin, and S. Pbo (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060.
- Reyer2008** Reyher, H.-U. (2008). Mating with the wrong species can be right. *Trends in Ecology and Evolution* 23(6), 289–292.
- Robinson1981** Robinson, D. R. and L. R. Foulds (1981). Comparison of phylogenetic trees. *Mathematical Biosciences* 53, 131–147.
- rogers96a** Rogers, J. S. (1996). Central moments and probability distributions of three measures of phylogenetic tree imbalance. *Systematic Biology* 45, 99–110.
- Rokas2005** Rokas, A. and S. B. Carroll (2005). More genes or more taxa? the relative contribution of gene number and taxon number to phylogenetic accuracy. *Molecular Biology and Evolution* 22(5), 1337–1344.
- Rokas2003** Rokas, A., B. L. Williams, N. King, and S. B. Carroll (2003). Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425, 798–804.
- Rosen1978** Rosen, D. E. (1978). Vicariant patterns and historical explanation in biogeography. *Systematic Zoology* 27(2), 159–188.

- Rosenberg2002** Rosenberg, N. A. (2002). The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology* 61, 225–247.
- Rosenberg2003** Rosenberg, N. A. (2003). The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly and polyphyly in a coalescent model. *Evolution* 57(7), 1465–1477.
- Rosenberg2006** Rosenberg, N. A. (2006). The mean and variance of the numbers of r-pronged nodes and r-caterpillars in yule-generated genealogical trees,. *Annals of Combinatorics* 10, 129–146.
- Rosenberg2007:counting** Rosenberg, N. A. (2007a). Counting coalescent histories. *Journal of Computational Biology*, 360–377.
- Rosenberg2007:monophyly** Rosenberg, N. A. (2007b). Statistical tests for taxonomic distinctiveness from observations of monophyly. *Evolution* 61(2), 317–323.
- Ruan2008** Ruan, J., H. Li, Z. Chen, A. Coghlan, L. J. M. Coin, Y. Guo, J.-K. Heriche, Y. Hu, K. Kristiansen, R. Li, T. Liu, A. Moses, J. Qin, S. Vang, A. J. Vilella, A. Ureta-Vidal, L. Bolund, J. Wang, and R. Durbin (2008). Treefam: 2008 update. *Nucleic Acids Research* 36, 735–740.
- Sackin1972** Sackin, M. J. (1972). “Good” and “bad” phenograms. *Systematic Zoology* 21(2), 225–226.
- Sagitov1999** Sagitov, S. (1999). The general coalescent with asynchronous mergers of ancestral lines. *Journal of Applied Probability* 36, 1116–1125.
- Saitou1987** Saitou, N. and M. Nei (1987). The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4(4), 406–425.
- Sanderson1996** Sanderson, M. J. (1996). How many taxa must be sampled to identify the root node of a large clade ? *Systematic Biology* 45(2), 168–173.
- Sanderson2011** Sanderson, M. J., M. M. McMahon, and M. A. Steel (2011, July). Terraces on phylogenetic tree space. *Science* 333, 448–450.
- Sang2000** Sang, T. and Y. Zhong (2000). Testing hybridization hypotheses based on incongruent gene trees. *Systematic Biology* 49, 422–434.
- Sankoff1975a** Sankoff, D. D. (1975). Minimal mutation trees of sequences. *SIAM Journal of Applied Mathematics* 28, 35–42.
- Sankoff1975b** Sankoff, D. D. and P. Rousseau (1975). Locating the vertices of a Steiner tree in arbitrary space. *Mathematical Programming* 9, 240–246.
- Sargsyan2008** Sargsyan, O. and J. Wakeley (2008). A coalescent process with simultaneous multiple mergers for approximating the gene genealogies of many marine organisms. *Theoretical Population Biology* 74, 104–114.

- Saunders1984** Saunders, I. W., S. Tavaré, and G. A. Watterson (1984). On the genealogy of nested subsamples from a haploid population. *Advances in Applied Probability* 16, 471–491.
- Schaffner01112005** Schaffner, S. F., C. Foo, S. Gabriel, D. Reich, M. J. Daly, and D. Altshuler (2005). Calibrating a coalescent simulation of human genome sequence variation. *Genome Research* 15(11), 1576–1583.
- Schliep2011** Schliep, K. and E. Paradis (2011). phangorn: phylogenetic analysis in r. *Bioinformatics* 27(4), 592–593.
- Schmidt2002** Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* 18(3), 502–504.
- Schoniger1995** Schoniger, M. and A. von Haeseler (1995, Dec). Performance of the maximum likelihood, neighbor joining, and maximum parsimony methods when sequences sites are not independent. *Systematic Biology* 44(4), 533–547.
- Schweinsberg2003** Schweinsberg, J. (2003). Coalescent processes obtained from supercritical Galton–Watson processes. *Stochastic Processes and their Applications* 106, 107–139.
- Semple2003** Semple, C. and M. A. Steel (2003). *Phylogenetics*. Oxford, UK: Oxford University Press.
- Sergio2012** Sergio, L. and J. Hey (2012). Demographic inference using spectral methods on snp data, with an analysis of the human out-of-africa expansion. *Genetics* 192(2).
- Shapiro2004** Shapiro, B., A. J. Drummond, A. Rambaut, M. C. Wilson, P. E. Matheus, A. V. Sher, O. G. Pybus, M. T. P. Gilbert, I. Barnes, J. Binladen, E. Willerslev, A. J. Hansen, G. F. Baryshnikov, J. A. Burns, S. Davydov, J. C. Driver, D. G. Froese, C. R. Harington, G. Keddie, P. Kosintsev, M. L. Kunz, L. D. Martin, R. O. Stephenson, J. Storer, R. Tedford, S. Zimov, and A. Cooper (2004). Rise and Fall of the Beringian Steppe Bison. *Science* 306(5701), 1561–1565.
- Sharkey2001** Sharkey, M. J. and J. W. Leathers (2001). Majority does not rule: The trouble with majority-rule consensus trees. *Cladistics* 17, 282–284.
- Shaw1998** Shaw, K. L. (1998). Species and the diversity of natural groups. In D. J. Howard and S. H. Berlocher (Eds.), *Endless Forms: Species and Speciation.*, pp. 44–56. Oxford, UK: Oxford University Press.
- Sheehan2013** Sheehan, S., K. Harris, and Y. S. Song (2013). Estimating variable effective population sizes from multiple genomes: A sequentially markov conditional sampling distribution approach. *Genetics* 194(3), 647–662.
- Shlyakhter2014** Shlyakhter, I., P. C. Sabeti, and S. F. Schaffner (2014). Cosi2: an efficient simulator of exact and approximate coalescent with selection. *Bioinformatics* 30(23), 3427–3429.

- Simon2006** Simon, C., T. R. Buckley, F. Frati, J. B. Stewart, and A. T. Beckenbach (2006). Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* 37, 545–579.
- Simon1994** Simon, C., F. Frati, A. Beckenbach, B. Crespi, H. Liu, and P. Flook (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* 87(6), 651–701.
- Skouras1998** Skouras, K. (1998). Absolute continuity of markov chains. *Journal of Statistical Planning and Inference* 75, 1–8.
- Slowinski1990** Slowinski, J. B. (1990). Probabilities of n-trees under two models: a demonstration that asymmetrical interior nodes are not improbable. *Systematic Zoology* 39(1), 89–94.
- Smith2007** Smith, C. I. (2007). Historical biogeography: The new synthesis. *Current Biology* 17, 598–600.
- Smythe1996** Smythe, R. T. (1996). Central limit theorems for urn models. *Stochastic Processes and their Applications* 65, 115–137.
- Sokal1958** Sokal, R. and C. Michener (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin* 38(22), 1409–1438.
- Song2005** Song, Y. S. and J. Hein (2005, March). Constructing minimal ancestral recombination graphs. *Journal of Computational Biology* 12(2), 147–169.
- Staab2015scrm** Staab, P. R., S. Zhu, D. Metzler, and G. Lunter (2015). scrm: efficiently simulating long sequences using the approximated coalescent with recombination. *Bioinformatics* 31(10), 1680–1682.
- PS06** Steel, M. and K. M. Pickett (2006). On the impossibility of uniform priors on clades. *Molecular Phylogenetics and Evolution* 39, 585–586.
- Steel1993parsimony** Steel, M. A. (1993). Distributions on bicoloured binary trees arising from the principle of parsimony. *Discrete Applied Mathematics* 43, 245–261.
- Steel2012** Steel, M. A. (2012). Root location in random trees: A polarity property of all sampling consistent phylogenetic models except one. *Molecular Phylogenetics and Evolution* 65(1), 345 – 348.
- Steel2001** Steel, M. A. and A. McKenzie (2001). Properties of phylogenetic trees generated by yule-type speciation models. *Mathematical Biosciences* 170, 91–112.
- Steel1993** Steel, M. A. and D. Penny (1993). Distributions of tree comparison metrics – some new results. *Systematic Biology* 42(2), 126–141.
- Steel2008** Steel, M. A. and A. Rodrigo (2008, Apr). Maximum likelihood supertrees. *Systematic Biology* 57(2), 243–250.

Sage2011 Stein, W. et al. (2011). *Sage Mathematics Software (Version 4.7)*. The Sage Development Team. <http://www.sagemath.org>.

Strimmer1997 Strimmer, K., N. Goldman, and A. von Haeseler (1997). Bayesian probabilities and quartet puzzling. *Molecular Biology and Evolution* 2(14), 210–211.

Strimmer1996nj Strimmer, K. and A. von Haeseler (1996a, Dec). Accuracy of neighbor joining for n-taxon trees. *Systematic Biology* 45(4), 516–523.

Strimmer1996 Strimmer, K. and A. von Haeseler (1996b). Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Molecular Biology and Evolution* 13(7), 964–969.

Tajima1983 Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* 105, 437–460.

Takacs1991 Takács, L. (1991). A bernoulli excursion and its various applications. *Advances in Applied Probability* 23(3), 557–585.

Takahata1985 Takahata, N. and M. Nei (1985). Gene genealogy and variance of interpopulation nucleotide differences. *Genetics* 110, 325–344.

Tallmon2004 Tallmon, D. A., G. Luikart, and M. A. Beaumont (2004). Comparative evaluation of a new effective population size estimator based on approximate Bayesian computation. *Genetics* 167(2), 977–988.

Tavare1984 Tavaré, S. (1984). Line-of-descent and genealogical processes and their applications in population genetics models. *Theoretical Population Biology* 26(2), 119–164.

Tellier14 Tellier, A. and C. Lemaire (2014). Coalescence 2.0: a multiple branching of recent theoretical developments and their applications. *Mol Ecol* 23, 2637–2652.

Than2009 Than, C. and L. Nakhleh (2009). Species tree inference by minimizing deep coalescences. *PLoS Computational Biology* 5(9), 12.

Than2008 Than, C., D. Ruths, and L. Nakhleh (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics* 9, 322.

valiente2002algorithms Valiente, G. (2002). *Algorithms on Trees and Graphs*. Berlin, Germany: Springer.

Iersel2008 van Iersel, L., J. Keijsper, S. Kelk, L. Stougie, F. Hagen, and T. Boekhout (2008). Constructing level-2 phylogenetic networks from triplets. *Annual International Conference on Research in Computational Molecular Biology 2008. LNCS 4955*, 450–462.

- Verdu2013** Verdu, P., N. S. Becker, A. Froment, M. Georges, V. Grugni, L. Quintana-Murci, J.-M. Hombert, L. Van der Veen, S. Le Bomin, S. Bahuchet, E. Heyer, and F. Austerlitz (2013). Sociocultural behavior, sex-biased admixture, and effective population sizes in central African Pygmies and Non-Pygmies. *Molecular Biology and Evolution* 30(4), 918–937.
- Vinh2011** Vinh, L. S., A. Fuehrer, and A. von Haeseler (2011). Random tree-puzzle leads to the Yule–Harding distribution. *Molecular Biology and Evolution* 28(2), 873–877.
- Wakeley2008** Wakeley, J. (2008). *Coalescent theory: An Introduction*. Greenwood Village, CO: Roberts and Co.
- Wakeley1997** Wakeley, J. and J. Hey (1997). Estimating ancestral population parameters. *Genetics* 145, 847–855.
- Watterson1975** Watterson, G. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* 7(2), 256 – 276.
- Watterson1984** Watterson, G. A. (1984). Lines of descent and the coalescent. *Theoretical Population Biology* (26), 77–92.
- Wegmann2010** Wegmann, D. and L. Excoffier (2010). Bayesian inference of the demographic history of chimpanzees. *Molecular Biology and Evolution* 27(6), 1425–1435.
- Wheeler2008** Wheeler, W. C. and K. M. Pickett (2008). Topology-bayes versus clade-bayes in phylogenetic analysis. *Molecular Biology and Evolution* 25(2), 447–453.
- Wiens2000** Wiens, J. J. and M. R. Servedio (2000, Apr). Species delimitation in systematics: inferring diagnostic differences between species. *The Royal society Biological Science* 267(1444), 631–636.
- Wilkinson2007** Wilkinson, M., J. O. McInerney, R. P. Hirt, P. G. Foster, and T. M. Embley (2007). Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends in Ecology and Evolution* 22(3), 114–115.
- Wiuf1997** Wiuf, C. and J. Hein (1997). On the number of ancestors to a DNA sequence. *Genetics* 147, 1459–1468.
- Wiuf1999** Wiuf, C. and J. Hein (1999). Recombination as a point process along sequences. *Theoretical Population Biology* 55(3), 248–259.
- Wright1931** Wright, S. (1931). Evolution in mendelian populations. *Genetics* 16, 97–159.
- Wu2012** Wu, Y. (2012). Coalescent-based species tree inference from gene tree topologies under incomplete lineage sorting by maximum likelihood. *Evolution* 66(3), 763–775.
- Xifara** Xifara, D. K., I. Mathieson, and G. McVean (0000). The detection, structure and uses of extended haplotype identity in population genetic data. *unknown* 0(0), 0.

- Yang2006** Yang, Z. (2006). *Computational Molecular Evolution*. Oxford: Oxford University Press.
- Yang2010a** Yang, Z. (2010). A likelihood ratio test of speciation with gene flow using genomic sequence data. *Genome Biology and Evolution* 2, 200–211.
- Yang2010b** Yang, Z. and B. Rannala (2010). Bayesian species delimitation using multi-locus sequence data. *Proceedings of the National Academy of Sciences of U.S.A* 107(20), 9264–9269.
- Yu2012** Yu, Y., J. H. Degnan, and L. Nakhleh (2012). The probability of a gene tree topology within a phylogenetic network with applications to hybridization detection. *PLoS Genet.* 8, e1002660.
- Yu2011** Yu, Y., C. Than, J. H. Degnan, and L. Nakhleh (2011). Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Systematic Biology* 60(2), 138–149.
- Yule1925** Yule, G. U. (1925). A mathematical theory of evolution. based on the conclusions of Dr. J.C. Willis, F.R.S. In *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, Volume 213, pp. 21–87. The Royal Society.
- Zhang2011** Zhang, C., D. Zhang, T. Zhu, and Z. Yang (2011). Evaluation of a bayesian coalescent method of species delimitation. *Systematic Biology* 60, 747–761.
- Zhu2013phd** Zhu, S. (2013). *Stochastic tree models and probabilistic modelling of gene trees of given species networks*. Ph. D. thesis, Department of Mathematics, University of Canterbury, Christchurch, New Zealand.
- Zhu2010poster** Zhu, S. and J. H. Degnan (2010). Effect of taxon sampling on constructing species tree from gene trees. Poster, http://www.math.canterbury.ac.nz/~s.zhu/my_files/taxon.pdf.
- Zhu2013sim** Zhu, S., J. H. Degnan, and B. Eldon. Hybrid-lambda: simulation of multiple merger and kingman gene genealogies in species networks and species trees. arxiv.org/abs/1303.0673.
- Zhu2015HybridLambda** Zhu, S., J. H. Degnan, S. J. Goldstien, and B. Eldon (2015). Hybrid-lambda: simulation of multiple merger and kingman gene genealogies in species networks and species trees. *BMC Bioinformatics* 16(292). doi:10.1186/s12859-015-0721-y.
- Zhu2011clade** Zhu, S., J. H. Degnan, and M. A. Steel (2011a). Clades, clans and reciprocal monophyly under neutral evolutionary models. *Theoretical Population Biology* 79, 220–227.
- Zhu2011poster** Zhu, S., J. H. Degnan, and M. A. Steel (2011b). Probabilistic modeling of gene trees given species networks. Poster, <http://www.newton.ac.uk/programmes/PLG/Zhu.pdf>. Feb 21, 2013.

- Zhu2013tp** Zhu, S. and M. A. Steel (2013). Does random tree puzzle produce Yule–Harding trees in the many-taxon limit? *Mathematical Biosciences* 243(1), 109–116.
- Zhu2014pda** Zhu, S., C. Than, and T. Wu (2014). Clades and clans: a comparison study of two evolutionary models. *Journal of Mathematical Biology*, 1–26.
- Zhu2012MBE** Zhu, T. and Z. Yang (2012). Maximum likelihood implementation of an isolation-with-migration model with three species for testing speciation with gene flow. *Molecular Biology and Evolution*.
- Zilversmit2013** Zilversmit, M., E. Chase, D. Chen, P. Awadalla, K. Day, and G. McVean (2013). Hypervariable antigen genes in malaria have ancient roots. *BMC Evolutionary Biology* 13(1), 110.
- Zwickl2002** Zwickl, D. J. and D. M. Hills (2002, Aug). Increased taxon sampling greatly reduces phylogenetic error. *Systematic Biology* 51(4), 588–598.