























fect phylogenies provide sensible groupings of samples at multiple resolutions. TreeQA explores the space of all possible groupings implied by the perfect phylogenies in a carefully designed order so that intermediate computations can be maximally reused. Our experimental results on both simulated and real data show that TreeQA can efficiently conduct quantitative GWA analysis and is more effective than the previous methods.

### References

1. R. G. Miller. *Simultaneous Statistical Inference*. Springer Verlag New York, 1981.
2. R. R. Hudson and N. L. Kaplan. Statistical properties of the number of recombination events in the history of a sample of dna sequences. *Genetics*, 111(1):147C164, 1985.
3. R. Agarwala, D. Fernandez-Baca, and G. Slutzki. Fast algorithms for inferring evolutionary trees. *Journal of Computational Biology*, 2(3):397–408, 1995.
4. H. Toivonen, P. Onkamo, K. Vasko, and et al. Data mining applied to linkage disequilibrium mapping. *Am J Hum Genet*, 67:133–145, 2000.
5. J. Akey, L. Jin, and M. Xiong. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? *Eur J. Hum Genet.*, 9(4):291–300, 2001.
6. F. Larrriba, S. Lessarda, and N. J. Schork. Gene mapping via the ancestral recombination graph. *Theoretical Population Biology*, 62(2):215–229, 2002.
7. A. P. Morris, J. C. Whittaker, and D. J. Balding. Fine-scale mapping of disease loci via shattered coalescent modeling of genealogies. *Am J Hum Genet*, 70(3), 2002.
8. P. Onkamo, V. Ollikainen, P. Sevon, and et al. Association analysis for quantitative traits by data mining: Qhpm. *Ann. Hum. Genet.*, 66:419–429, 2002.
9. X. Wang and B. Paigen. Quantitative trait loci and candidate genes regulating hdl cholesterol. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 22:1390, 2002.
10. J. Li and T. Jiang. Haplotype-based linkage disequilibrium mapping via direct data mining. *Bioinformatics*, 21(24):4384–4393, 2005.
11. T. Mailund, M. H. Schierup, and et al. Coasim: A flexible environment for simulating genetic data under coalescent model. *BMC Bioinformatics*, 6(252), 2005.
12. E. Waldron, J. Whittaker, and D. Balding. Fine mapping of disease genes via haplotype clustering. *Genetic Epidemiology*, 30:170–179, 2005.
13. S. Zöllner and J. K. Pritchard. Coalescent-based association mapping and fine mapping of complex trait loci. *Genetics*, 169(2):1071–92, 2005.
14. T. Mailund, S. Besenbacher, and M. H. Schierup. Whole genome association mapping by incompatibilities and local perfect phylogenies. *BMC Bioinformatics*, 7:454, 2006.
15. P. McClurg, M. T. Pletcher, T. Wiltshire, and A. I. Su. Comparative analysis of haplotype association mapping algorithms. *BMC Bioinformatics*, 7:61, 2006.
16. M. J. Minichiello, and R. Durbin. Mapping Trait Loci by Use of Inferred Ancestral Recombination Graphs. *Am J Hum Genet*, 79(5):910–922, 2006.
17. I. Pe'er and et al. Evaluating and improving power in whole-genome association studies using fixed marker sets. *Nature Genetics*, 38:663–667, 2006.
18. P. Sevon, H. Toivonen, and V. Ollikainen. Treedt: Tree pattern mining for gene mapping. *IEEE Transactions on Computational Biology and Bioinformatics*, 3(2), 2006.
19. H. Yang, T. A. Bell, G. A. Churchill, and F. P.-M. de Villena. On the subspecific origin of the laboratory mouse. *Nature Genetics*, 39, 2007.
20. J. P. Szatkiewicz, G. L. Beane, Y. Ding, L. Hutchins, F. P.-M. de Villena, and G. A. Churchill. An imputed genotype resource for the laboratory mouse. *Mammalian Genome*, 19(3), 2008.
21. T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. Introduction to algorithms.