1. Simulate datasets of sequence data sampled from two populations that stem from a recent joint ancestral popultion and compute the JSFS of the data. Explore (e.g. by averaging over many simulations) how the expectation values of the entries of the JSFS depend on parameters like the time since the split, rates of geneflow between the population, population size ratios and population growth rates.

2. Implement ABC-PMC in R for a demographic model like in the above exercise and explore the performance of this approach. (For the actual ABC steps you can use the abc command from the abc package.)

3. Two populations emerged from a recent split of an ancestral population. There may still be a little amount of gene-flow between the populations.

   (a) Explore for different sizes of datasets (start with 10 independent gene loci) how accurate ABC, IM/IMa, Lamarc and Jaatha can estimate the model parameters θ for the ancestral and the two descendant populations, time of populations split, migration rates and population growth rates.

   (b) How can we deal with recombination within loci? Does recombination within loci increase or decrease the accuracy of parameter estimations?

4. Assume you simulate an ancestral recombination graph (ARG) for a n whole chromosomes sampled from a neutral, constant-size population (with positive recombination rate). From all the trees that this ARG assigns to single nucleotide positions, you choose four:

   Tree A is the tree at position 1000.

   Tree B is the next tree after Tree A, that is, you move on from position 1000 until there is a recombination event somewhere on a branch of tree A. Tree B is then the tree of the nucleotide position right after the recombination event.

   Tree C is the 100th tree in the list of all trees appearing from left two right, that is, the tree after the 100th recombination event (where only recombination events that effect the current tree are counted).

   Tree D is the next tree after Tree C, that is, you move on until there is a recombination event somewhere on a branch of tree C. Tree D is then the tree of the nucleotide position right after the recombination event.

   Which of the following statements (a) to (k) are true? Substantiate your answers either logically or with computer simulations. (Start with $n = 2$ and $n = 3$). If you rely on computer simulations, try to find logical explanations for your observations.

   (a) The probability distribution of Tree A is that of a standard Kingman coalescent.

   (b) The probability distribution of Tree B is that of a standard Kingman coalescent.

   (c) The probability distribution of Tree C is that of a standard Kingman coalescent.

   (d) Tree B has the same probability distribution as tree A.

   (e) The expected total branch lenght of A is smaller than that of B.

(f) The expected total branch lenght of B is smaller than that of A.

(g) Tree C has the same probability distribution as tree A.

(h) The expected total branch lenght of A is smaller than that of C.

(i) Tree D has the same probability distribution as tree C.

(j) The expected total branch lenght of C is smaller than that of D.

(k) The expected total branch lenght of D is smaller than that of C.