1. With the help of a dice and/or random generators available R simulate a Moran model for a haploid population of only six individuals with selection.

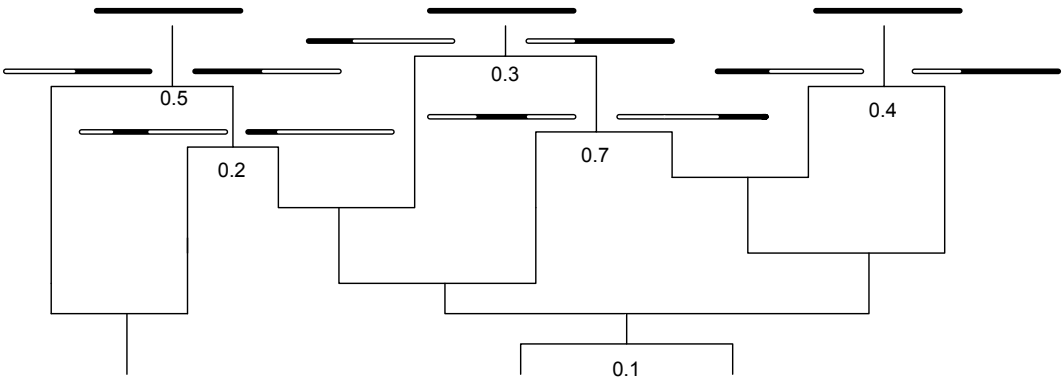
2. With the help of a random generator in R simulate an Ancestral Selection Graph on a piece of paper, assuming a sample size of six with weak selection (and no recombination).

3. Simulate with coala/msms population genetic data containing selective sweeps and assess the sensitivity of Kim and Nielsen's ω and of iHS, SDS and nS_L to detect the sweeps. Also explore under what conditions demographic effects such as bottlenecks can lead to false positives in sweep detection.

4. Calculate Kim and Nielsen's ω for the following sequence alignment, which is already reduced to polymorphisms, and in which dots always refer to the same nucleotide as in the sequence 1.

```
1 ACGTCT
2 ...CGA
3 ......
4 T.CC.A
5 TAC.G.
6 TAC...
7 TA....
8 ...CGA
```

5. In the following (incomplete) ancestral recombination graph of three lineages, solid black lines visualize for ancestral material and the recombination sites are given relative to the total length of the genomic region:



   (a) For the three lineages at the bottom visualize the sections that contain ancestral material.

   (b) The aim of simulating this ancestral recombination graph is to simulate sequence data. To keep the simulation efficient, we want to simulate as little as possible.

      i. Is there one or more lineage(s) that we do not need to simulate further into the past to simulate sequence data according to the population genetic model (combined with one of the usual sequence evolution models)?

      ii. Are there subsections of the genomic regions for which we do not need to simulate recombination events or mutations further back in the past?

6. A diploid locus with alleles A and a evolves according to an MSMS model with two demes. In the current generation ("generation 1") there are 200 diploid individuals in population 1 and 1000 diploid individuals in population 2. The allele-A frequencies are 0.3 in deme 1 and 0.5 in deme 2. The fitness values of the genotypes are $1 + s_1^{aa} = 1$, $1 + s_1^{Aa} = 1.3$ and $1 + s_1^{AA} = 1.4$ in deme 1, and $1 + s_2^{aa} = 1.1$, $1 + s_2^{Aa} = 1.3$ and $1 + s_2^{AA} = 1.2$ in deme 2. In each generation, 20 individuals from island 1 migrate to island 2 and 20 individuals from island 2 migrate to island 1. The probability of a mutation per generation is 1/10,000 for mutations from a to A and 1/20,000 for mutations from A to a.

   (a) Calculate the expected values for the A allele frequencies of A in demes 1 and 2 in the next generation.

   (b) Calculate the probability that the A allele frequencies in the next generation ("generation 2") are 0.32 in deme 1 and 0.52 in deme 2.

   (c) Assume now that it is actually the case that the frequencies of A are 0.32 in deme 1 and 0.52 in deme 2 in generation 2.

      i. Calculate the probability of two gene copies in generation 2 in deme 1, both of allele A, to stem from the same allele-A copy in generation 1 in deme 1.

      ii. Calculate the probability of two gene copies in generation 2, one in deme 1 and of allele A, the other in deme 2 of allele a, stem from the same gene copy in generation 1.

      iii. Calculate the probability of two gene copies in generation 2 in deme 1, both of allele A, to stem from the same gene copy in generation 1.

7. We trace back the ancestral lineages of two loci 1 and 2 on the two copies of chromosome 1 of a diploid individual. The recombiation rate between the two loci is $\rho = 2N_e r = 10$. We distinguish the following states:

   A: There are two ancestral lineages, each containing ancestral material of both loci.

   B: There are three ancestral lineages. One of them contains ancestral material of both loci (and the other two of one locus each).

   C: There are four separate ancestral lineages.

   D: The ancestral lineages of one and only one locus have coalesced.

   E: The ancestral lineages of the two loci have coalesced at different time points.

   F: The ancestral lineages of the two loci have coalesced at the exact same time point.

   (a) Explain how the ancestral recombination graph induces a Markov chain between the states A to F and calculate the transition probabilities.

   (b) Explain how the ancestral recombination graph defines a continuous time Markov jump process on the states A to F and calculate the transition rates.

   (c) Calculate the probability that the ancestral lineages of the two loci coalesce at the exact same time.

   (d) Calculate the expected value of the time until the lineages of both loci have coalesced.

   (e) Characterize as precisely as possible the conditional distribution of the coalescence time of loci 2 for the case that a coalescence of loci 1 at time 0.8 back in the past is given.

   (f) Carry out the previous calculations also for the cases $\rho = 0.1$ and $\rho = 100$.