Universität Bielefeld
Technische Fakultät

AG Genominformatik
Prof. Dr. Jens Stoye

Algorithms in Genome Research
Winter 2013/2014

Exercises

Number 11, Discussion: 2014 January 31

1. Remember and explain a few of the terms (hopefully) mentioned in class:
   - genotype, haplotype
   - segregating site
   - genealogy
   - infinite sites model
   - four-gametes test
2. Give general formulas for the following questions. If this is difficult, enumerate the solutions for small examples.
   (a) For k markers of two alleles each, how many different haplotype vectors are possible?
   (b) If the haplotypes come in blocks of 10 sites each, how does this decrease the number of different haplotype vectors?
   (c) For m founder sequences and l recombination hot spots, how many haplotype vectors are possible (under the assumption that recombinations only occur at hot spots)?
   (d) For k individuals, what is the maximum number of different segregating sites such that the four-gametes test does not fail?
3. Given the following instance of the haplotype inference problem (HIP): 212022, 102212, 101110, 211020, 122110.
   - Apply Clark's algorithm.
   - Solve the pure parsimony variant of the HIP.
   - Solve the perfect phylogeny haplotyping problem (PHP).