LINKAGE MAPPING USING OPEN-POLLINATED POPULATIONS

Ben-Hui Liu, Ron Sederoff, and David O'Malley

Forest Biotechnology Group
Department of Forest, Box 8008
North Carolina State University
Raleigh, NC 27695-8008

Genomic mapping in forest trees is making fast progress. Several approaches have been employed for construction linkage maps for trees based on special features of the different species. Haploid megagametophyte tissue in the seeds, which is genetically identical with the maternal gamete, has been used to map individual mother tree for several conifer species. A pseudo-testcross strategy has been used to construct linkage maps using diploid tissues of full-sib population for *Eucalyptus.* Another approach to genomic mapping in trees has been to find appropriate 3-generation outbred pedigrees, where the 4 grandparents are available and the two parents can be crossed.

Here, we propose to construct linkage map for individual mother trees from open-pollinated population. General likelihood equation for this type of mating, which is a function of recombination frequency, allele frequency in pollen pool, and outcrossing rate, is presented in this poster. If mother genotype is known, two approaches can be used to solve the likelihood equation and get recombination estimation. The first is to put the estimated allele frequencies in pollen pool and outcrossing rate into the likelihood function before estimating the recombination frequency using a standard maximum likelihood method. The second is to estimate recombination frequency, allele frequency in the pollen pool, and outcrossing rate simultaneously using EM algorithm. If genotype of the mother tree is unknown, then EM algorithm is used to solve the likelihood equation. The power of detecting linkage, the quality of recombination frequency estimation, and the feasibility are compared with each other and are compared with a check situation where the maternal genotype, allele frequency in the pollen pool, and outcrossing rate are known.

Success of this research has at least the following impacts on genome mapping in forest trees: (1) Wide use of breeding populations will be feasible for genome mapping. In turn, genome map information can be used in breeding directly. (2) The approach proposed here, haploid megagametophyte method, and the pseudo-testcross strategy can be validated by crossing compression. (3) Mapping populations with control pollination, such as F2, backcross, recombined inbred line etc., are actually special cases of open-pollinated populations, with different rates of outcrossing and allele frequencies in the pollen pool. This will change the design of the computer package for linkage map construction. A single unique likelihood function for open-pollinated population with specifications of outcrossing rate and the allele frequencies can be used for any type of mating. This will make the software and the algorithm generalizable.