**ALA 5.2 QTL Analysis**

**Prerequisites**

Understanding of:

1. Linkage Disequilibrium
   1. Conceptual basis
   2. Estimation
   3. Sources of linkage disequilibrium
   4. Decay of Linkage disequilibrium
2. Associations between markers and phenotypes.
   1. Genome Wide Association Studies
3. QTL mapping

Completion of ALA4.1, including ability to conduct simple simulation experiments; Introduction to Breeding Management System (BMS)

**Purpose**

Provide understanding of the different factors affecting the power of detecting QTL.

**Background**

Success of QTL mapping depends on various factors, such as population type, population size, heritability, gene/QTL effect, choice of QTL mapping method, among others. Different from QTL analysis of experimental data, where the trait architecture is unknown, simulation experiments allow you to “set the stage” and ask, how often are known QTL detected (power of QTL detection) in multiple simulation runs.

**Tasks**

**This is from ALA4.1 (should have been addressed earlier, not part of ALA5.1)**

The main purpose of this task is to become familiar with data simulation. An important task of this exercise is to establish the data needed for a QTL mapping analysis of grain yield by simulation. Two different types of populations will be used, a F2 and a DH population (N=200 each). For field testing of the F2 population, F2 individuals will be represented by F2:3 individuals. Both populations are tracing back to the same original cross. In other words, the trait architecture is identical in both mapping populations. Both populations will be evaluated in three environments. There are significant G x E interactions. The species studied here has 5 chromosomes of 100 cM length each. Dominant marker costs are low, and the mapping individuals will be mapped with markers in 1 cM spacing. Six QTL control this trait. Two QTL are linked in coupling, and two are linked in repulsion. The remaining two QTL are unlinked and located on different chromosomes than the two linked QTL pairs.

1. Describe input parameters of genetic architecture that you chose based on the constraints above, such as QTL location and effect.
2. Create a simulated DH population evaluated in 3 environments using the above mentioned conditions.
3. Create a simulated F2:3 population evaluated in 3 environments using the above mentioned conditions.
4. Do a proper field data analysis, as learned in Quantitative Methods, and present key phenotypic data for both populations, including means across and for individual environments, variance components (G, E, GxE), and heritability estimates.

**Tasks for ALA5.1, directly connecting to ALA4.1:**

1. In extension of ALA4.1, generate 10 DH and F2:3 QTL mapping populations (N=200) each for repeated QTL experiments.
2. Load data into BMS for QTL mapping.
3. Conduct QTL analyses for these in total 20 populations using simple interval mapping (SIM) and composite interval mapping as implemented in BMS.
4. Summarize power of detecting each of the six QTL, depending on population type, QTL mapping method, linkage of QTL and size of QTL effect. Moreover, report on the position of detected QTL relative to the true QTL position, as well as detection of false positives.
5. Extend the power of QTL detection analysis to another factor of your choice.

**Tentative answers** (can differ, based on context / assumptions made)