QTL Mapping in Autotetraploid Species: Theory and Application to Map QTL Affecting Blight Resistance in Potato

C. A. Hackett, J.E. Bradshaw, Z. W. Luo[1], H. E. Stewart, B. Pande, G. Bryan, R. Waugh & J. W. McNicol

Introduction Linkage maps based on molecular markers are an important tool in plant genetics, especially for locating the quantitative trait loci (QTL) controlling important traits such as yield or disease resistance. Methods for linkage analysis and QTL mapping are well established for diploid species, but the extension to polyploid species is complicated by the large number of possible genotypes in polysomic inheritance. One method for locating QTL, in diploids or polyploids, is to compare the mean trait values for individuals with and without a marker allele. This is a useful preliminary analysis, but less informative than an analysis using all markers on a chromosome simultaneously. Here, we outline the statistical methodology underlying this latter analysis, and apply it to examine the genetic control of maturity and the resistance of the foliage to late blight (*Phytophthora infestans*) in a tetraploid potato population comprising two parents (Stirling and 12601ab1) and their full-sib offspring.

Statistical methodology There are three main stages to the statistical analysis. The first stage is to construct a linkage map of the markers for each parent and to deduce which of the four homologous chromosomes has each marker allele i.e. the marker phase. The second stage is to construct a 'graphical genotype' for each offspring, showing which chromosome segments it has inherited from which parent and where recombinations between chromosomes have occurred. The final stage is to relate the trait data to the graphical genotypes to locate QTLs. We discuss each stage in turn.

LINKAGE ANALYSIS Molecular markers are first separated into linkage groups by identifying groups of markers that do not segregate independently in the offspring. For each linkage group, the probability of a recombination event (the recombination frequency) and the likelihood of linkage (the *lod* score) are calculated for every pair of markers in each possible phase.

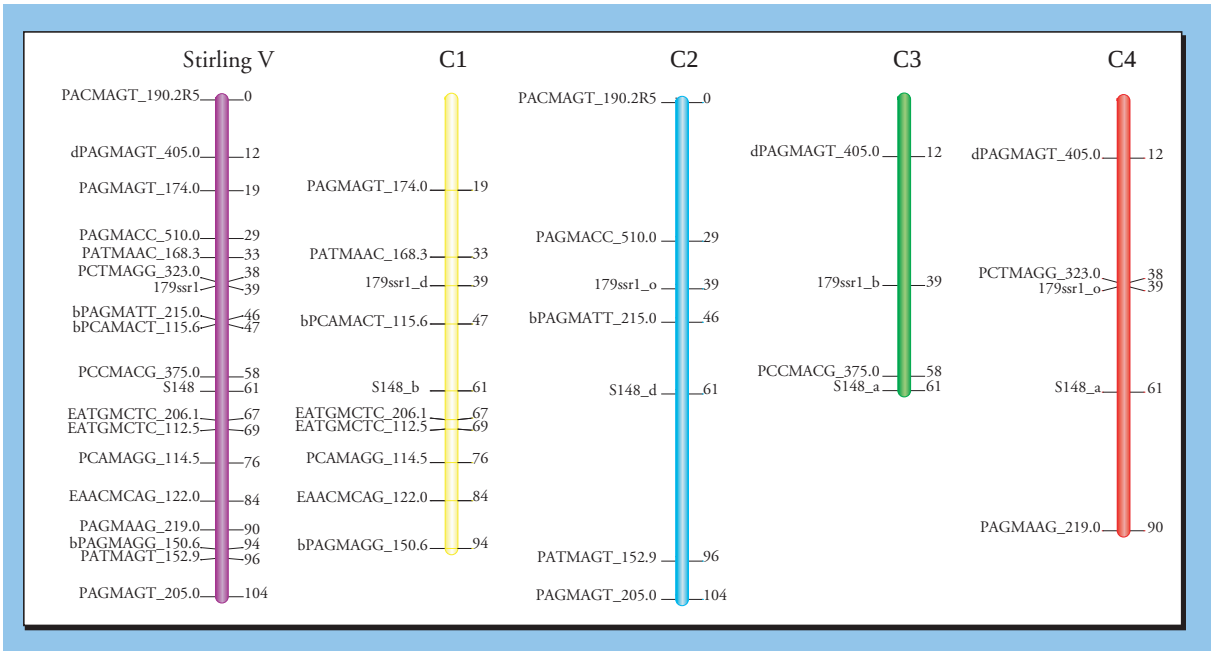

Figure 1 A map of linkage group V for Stirling, showing the overall marker order and the positions of the alleles on the individual chromosomes C1-C4. Marker positions are in centiMorgans (cM).

[1] Current address: School of Biosciences, The University of Birmingham, Birmingham B15 2TT

locus	Phenotype	Position	Graphical Genotype	
PACMAGT_190.2R5	A	0	O	A
dPAGMAGT_405.0	O	12.8	O	O
PAGMAGT_174.0	A	19.0	A	O
PAGMACC-510.0	A	28.8	O	A
PATMAAC_168.3	A	33.0	A	O
PCTMAGG_323.0	O	37.5	O	O
179ssrl	BD	38.7	D	O
bPAGMATT_215.0	A	45.8	O	A
bPCAMACT_115.6	A	46.9	A	O
PCCMACG_375.0	O	57.9	O	O
S148	ABD	60.5	A	D
EATGMCTC_206.1	O	66.9	O	O
EATGMCTC_112.5	O	69.2	O	O
PCAMAGG_114.5	O	75.8	O	O
EAACMCAG_122.0	O	83.6	O	O
PAGMAAG_219.0	A	90.4	A	O
bPAGMAGG_150.6	O	93.5	O	O
PATMAGT_152.9	A	96.0	O	A
PAGMAGT_205.0	A	103.9	O	A

Figure 2 A graphical genotype for the inheritance of chromosomes from linkage group V of Stirling. The B allele of 179ssr1 and the B allele of S148 are from the 12601ab1 parent.

If both markers are single dose markers, with an allele present on one chromosome of one of the parents, there are two possible phases: the alleles are both on the same chromosome (coupling phase) or they are on different chromosomes (repulsion phase). For markers such as simple sequence repeat markers (SSRs), where several alleles can be identified, there can be up to 24 possible phases for each parent. The EM algorithm enables recombination frequencies to be calculated efficiently for every phase. The recombination frequency for the phase with the highest likelihood is used to order the markers. There are a large number of possible marker orders: 10 markers can be ordered in 1,814,400 ways! However a computer search algorithm, simulated annealing, compares orders to find the optimal one. Figure 1 shows a linkage map for chromosome V of Stirling, based on these methods.

GRAPHICAL GENOTYPING For QTL interval mapping it is necessary to infer the genotype for each offspring at possible QTL locations between the mapped markers. To do this, we create a 'graphical genotype' for each offspring, to show how chromosome segments have been inherited from each parent. Consider, for example, the S148 SSR locus on the map of Stirling, which has four different alleles in this population. The parental genotypes are BDAA for Stirling and CBBB for 12601ab1, so an offspring with

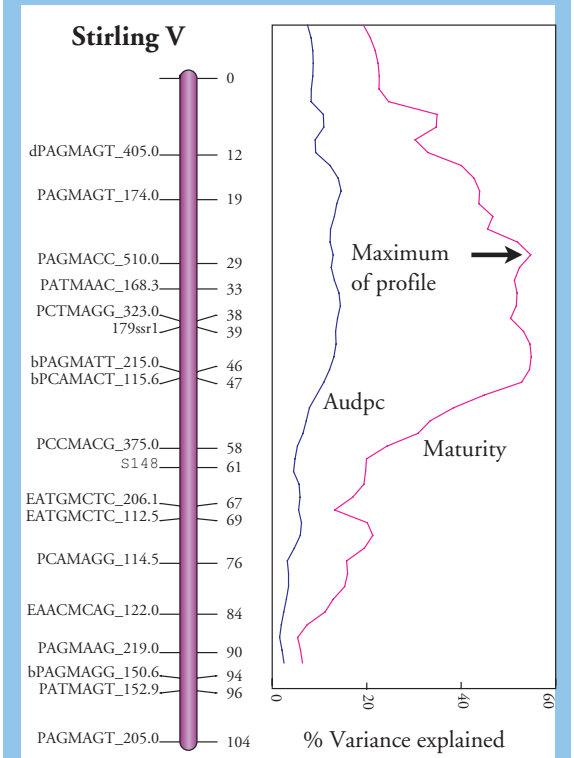

Figure 3 Interval mapping of maturity and resistance of the foliage to late blight (AUDPC) for Stirling linkage group V.

the A, B and D alleles at this locus must have received the A and D alleles from Stirling, and two copies of the B allele from 12601ab1 i.e. at this locus it has received from Stirling segments from chromosome 2, and either chromosome 3 or chromosome 4. We identify possible segments for each locus in turn, and then use a branch and bound algorithm to identify the chromosome configurations that give the observed phenotypes with as few recombinations as possible. One such graphical genotype for chromosomes from linkage group V inherited from the Stirling parent is shown in Figure 2. This individual's phenotypic data is consistent with it having inherited chromosome 2 without recombination, and a recombined chromosome with part from Stirling chromosome 1 and part from Stirling chromosome 4. The recombination is located between the 179ssr1 and the S148 loci.

QTL MAPPING The graphical genotype enables us to infer the genotype at a possible QTL at any position along the chromosome, assuming that there are no double recombinations between markers. For the individual shown in Figure 2, we deduce that it

receives alleles from chromosomes 1 and 2, i.e. a QTL genotype $Q_{12,}$ with probability 1 at positions from the start up to locus 179ssr1, a genotype of Q_{24} with probability 1 at positions from S148 to the end, and a mixture of these two genotypes at positions between 179ssr1 and S148, with probabilities depending on the position. For interval mapping, we consider the possibility of a QTL at a grid of positions along the chromosome, using steps of 1-2 cM. At each position, a mixture of normal distributions is used to relate the trait value of each individual to its inferred QTL genotype(s) at that position, using the QTL genotype probabilities. The mixture model can be fitted by weighted regression in an iterative manner, modifying the initial weights from the graphical genotype to include information on the trait. We obtain a profile of the percentage of the trait variance accounted for by a QTL for the trait at each location.

Maturity and resistance of foliage to late blight

Figure 3 shows a QTL profile for resistance of the foliage to late blight (measured as the area under the disease progress curve, AUDPC) and maturity for Stirling linkage group V. There are significant QTLs for both traits, with the QTL for maturity accounting for 55% of the variance, and that for blight resistance accounting for 14%. The most likely position is between markers PATMAAC_168.3 and PCT-MAGG_323.0. The QTL analysis also shows that the allele on chromosome 1 has a significantly different effect on the trait from those on chromosomes 2, 3 and 4. There is a strong linear relationship between blight resistance and maturity. When this effect is removed by a linear regression analysis, no further variance in blight resistance is explained by a QTL on linkage group V. These results indicate that we have a QTL for maturity on chromosome V, with offspring carrying the allele from chromosome 1 maturing earlier than those without this allele and that this QTL has an indirect effect on blight resistance, through the change in maturity, with early maturity associated with increased susceptibility.

Another QTL for blight resistance is found on Stirling linkage group IV, and this accounts for 24% of the trait variance. However there is no evidence for a QTL affecting maturity on linkage group IV: the resistance mechanism is different from that on chromosome V. For this QTL, offspring receiving alleles from both chromosomes 1 and 4 have the highest AUDPC scores, while offspring receiving alleles from both chromosomes 2 and 3 have the lowest scores. This indicates a double-dose locus for blight resistance, with offspring with either the allele from chromosome 2 or from chromosome 3 having some resistance, and those with alleles from both chromosomes 2 and 3 being most resistant. These results are directly relevant to potato breeding and cultivar production at the tetraploid level. For example, it should be relatively easy to identify progeny clones that combine the blight resistance alleles from linkage group IV with early maturity from chromosome V, an important goal in potato breeding.

Acknowledgements

This work was financially supported by grants from the UK Biotechnology and Biological Research Council, and by the Scottish Executive Environment and Rural Affairs Department.