

Supplementary material

S2. Comparison with QTL detection by iCIM

The results of a QTL detection scan by iCIM (Inclusive Composite Interval Mapping) for plant height (HT) are shown in Figure S1. Similarly to what was observed with conventional CIM the results show an overestimation of detected QTLs. The main reason for that is the confounding effects originating from the correlation between non-linked chromosome regions. In this particular example, it is known that the two major genes related to plant height in this population are located on chromosomes 3H and 5H. However, the output from iCIM shows a very strong signal on chromosome 6H as well, which suggests a major QTL on that chromosome too. Not surprisingly, we have found a high linkage disequilibrium (LD) between markers on the region on chromosome 6H and those nearby the gene on chromosome 3H ($r^2=0.2647$). The large LD between markers in different chromosomes have caused the signal detected by iCIM on chromosome 6H, which is a case of false positive. Although an attractive approach for QTL mapping, iCIM relies on assumptions that are not met in this population (independence of residuals while performing stepwise regression), therefore we should be cautious when interpreting the results obtained in our population. In addition, the two panels in Figure S1 show that the results from iCIM are highly dependent on the settings used for cofactors selection by stepwise regression (P values used as thresholds). The profile on the left (A) show a QTL on chromosomes 1H, 2H, 3H (2 QTLs), 4H (2 QTLs), 5H (3 QTLs), and 6H, however the profile on the right (B) only show a QTL on chromosomes 2H, 3H (2 QTLs), 5H (2 QTLs), and 6H. Note that some of the detections in Fig S1A are lost in Fig S1B, notably the large QTL on chromosome 4H, but also the ones on 1H and 5H. The different results of iCIM after changing the stepwise selection settings, once more points to the instability of results in the case of non-independence between markers in different linkage groups. Changing the threshold for variable inclusion in the stepwise selection of markers from 0.05 to 0.01, had produced a different set of selected markers, which in turn resulted in a different final set of detected QTLs. Note, that under standard segregating populations there should be no major differences because of this change in settings.

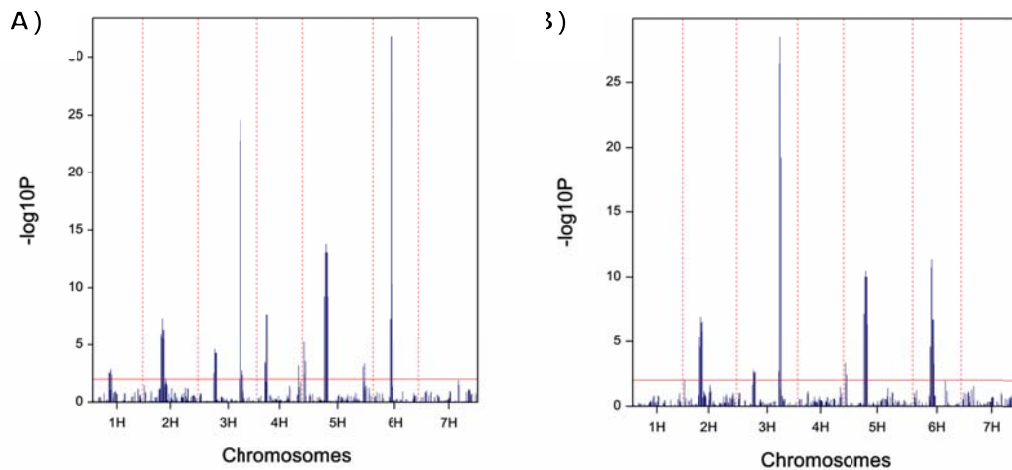


Fig S1. Genome wide results for plant height (HT) by iCIM under two different settings: A) largest P -value for entering variables set to 0.05, and the smallest P -value for removing variables set to 0.10, B) largest P -value for entering variables set to 0.01, and the smallest P -value for removing variables set to 0.10. Results are given as P values on a $-\log_{10}$ scale. Each spike corresponds to a particular SNP, with those exceeding the red horizontal line being significantly associated with the trait.