Correspondence

Accounting for polysomic inheritance in quantitative trait loci mapping of autopolyploids

A comment on Chen *et al*. (2021): 'A tetrasomic inheritance model and likelihood-based method for mapping quantitative trait loci in autotetraploid species.'

We recently became aware of an inaccurate reference to our work published in your journal which we would like to discuss here in order to rectify the record. The statement in question asserted that '… the method described in Bourke *et al.* (2019) lacks an essential component; that is, genetic linkage analysis between quantitative trait locus (QTLs) and surrounding markers for QTL mapping in an outbred autotetraploid segregating population, as we have presented here' (Chen *et al.*, 2021). We would like to briefly explain why we think this assertion is incorrect.

Quantitative trait loci analysis is a statistical approach to find associations between genotypic and phenotypic data, detecting regions of the genome that contribute to phenotypic variation. The concept of QTL 'interval mapping' was developed which partly decoupled the detection of QTL from marker positions, allowing the effects of putative QTL to be estimated at any position between markers (Lander & Botstein, 1989). Interval mapping approaches employing linear regression rather than maximum likelihood to estimate QTL effects and locations were subsequently introduced, achieving similar results (Haley & Knott, 1992; Martinez & Curnow, 1992).

Much of the focus of subsequent QTL mapping methodologies has been based on the use of posterior or conditional genotype probabilities generated using Hidden Markov Models (HMM; Lander & Green, 1987; Broman *et al.*, 2003; Hackett *et al.*, 2013). In such an approach, marker information is shared across loci in a Markov chain procedure to derive the posterior probabilities of inheritance of parental alleles, for each offspring in a mapping population. Zheng *et al.* (2016) developed a Hidden Markov Model for autotetraploids accounting for tetrasomic inheritance and double reduction in biparental autotetraploid populations and more recently extended this to multiparental populations (Zheng *et al.*, 2016, 2021). Although offspring posterior genotype probabilities are calculated and returned for marker positions only, they are derived from shared inheritance information of all markers along a chromosome, handled in a Markov chain. The interpolation of such probabilities to any position between markers can be achieved from the HMM output by, for example fitting cubic splines between marker positions (Hackett *et al.*, 2013; Bourke *et al.*, 2019). Although somewhat lacking in statistical rigour, this step gives reliable and predictable results, since high marker densities and low recombination levels result in only gradual fluctuations in posterior genotype probabilities across chromosomes. In practice, one may choose an appropriate grid of positions (e.g. at every 1 centiMorgan along a chromosome) to then perform whole-genome QTL scans at these positions (Bourke *et al.*, 2021).

The computational advantage of this approach becomes apparent when generating significance thresholds using permutation tests, particularly when using modern high-density maps, as each chromosome has *c.* 100 positions to test rather than potentially thousands. Testing at thousands of separate positions also has little advantage over a grid search approach due to limited population sizes and their associated limited genetic resolution. 'Fine-mapping' at a smaller grid size is always possible within a QTL peak region, if sub-centiMorgan resolution of QTL positions is expected due to population-level considerations.

It is in this context that we challenge the assertion that genetic linkage between a QTL and surrounding markers is lacking in our HMM-based approach. In our understanding, such linkage is explicitly modelled in the transition matrices of the HMM used to generate posterior genotype probabilities. The claim that the method of Chen *et al.* (2021) is unique in accounting for the full complexity of tetrasomic inheritance is also misleading – the multivalent-aware HMM of Zheng *et al.* (2016, 2021) has a clearly demonstrated the ability to correctly predict offspring derived from multivalent pairing structures carrying either double reduction products or multihomologue mosaics of more than two parental homologous chromosomes. These predictions (i.e. the posterior genotype probabilities) are directly used in the genotype–phenotype association model; thus, all the 'key features of tetrasomic inheritance' are included. The assertion that nonadditive effects are ignored in our approach is only the case in the initial whole-genome scan (implemented in the POLYQTLR package, available via CRAN; Bourke *et al.*, 2021). Nonadditive or epistatic effects can be subsequently modelled, allowing the user to develop a multi-QTL model that allows for nonadditive effects at each position.

Chen *et al.* (2021) also stated that their study 'enables QTL mapping analysis to be conducted in autotetraploid species'. Their implementation in an R package released with their publication (called 'TetraQTLAnalysis') requires that the user already knows the rate of double reduction at each locus, a nontrivial input requirement that most users will be unable to provide. In the test dataset with their package, the function QvMethod took over 17 h to detect a single QTL in a mapping population of 300 individuals carrying 20 markers along a single chromosome (using a desktop computer with 32 Gb RAM and an Intel Xeon W-2133 3.6 GHz processor). No significance thresholds were returned with the

New Phytologist

results. The same QTL peak was discovered using POLYQTLR (Bourke *et al.*, 2021) in 5 s, including the steps of posterior genotype probability calculation and threshold setting using a permutation test ($n = 1000$). The results were almost identical in terms of the predicted peak position and QTL configuration (QQoo × oooo at 36.9 cM vs QQoo × oooo at 35.7 cM). We provide a sample script to replicate this result (Supporting Information Notes S1). In modern datasets with hundreds or thousands of markers (and multiple chromosomes), the method of Chen *et al.* (2021) would quickly become computationally infeasible.

Chen *et al.* (2021) have extended the interval mapping framework to autotetraploids while accounting for double reduction. As a theoretical exercise, their work has much merit, but a practical and reproducible application of their method to realistic datasets has yet to be demonstrated. Our method does not omit linkage between markers and putative QTL, rather this linkage is modelled in a multivalent-aware HMM procedure. As demonstrated here, this turns out to be a far more practical approach to accounting for polysomic inheritance in QTL mapping of autopolyploids.

Acknowledgements

The software tool POLYQTLR described in the correspondence was developed with support from the Dutch TKI project 'Novel genetic and genomic tools for polyploid crops' (project no. KV 1605-20, BO-26.03-009-004) and the USDA's National Institute of Food and Agriculture (NIFA) Speciality Crop Research Initiative project 'Tools for genomics-assisted breeding in polyploids: Development of a community resource' (2020-51181-32156/SCRI).

Competing interests

None declared.

Author contributions

PMB drafted the correspondence, and REV and CM edited the correspondence. All authors read and approved the final version.

ORCID

Peter M. Bourke https://orcid.org/0000-0002-0665-6508
Chris Maliepaard https://orcid.org/0000-0002-7319-5270
Roeland E. Voorrips https://orcid.org/0000-0003-4388-9244

Data availability

Data are available in the Supporting Information.

Peter M. Bourke*, Roeland E. Voorrips and Chris Maliepaard

Plant Breeding, Wageningen University & Research, 6708PB, Wageningen, the Netherlands
(*Author for correspondence: email peter.bourke@wur.nl)

References

Bourke PM, Hackett CA, Voorrips RE, Visser RGF, Maliepaard C. 2019. Quantifying the power and precision of QTL analysis in autopolyploids under bivalent and multivalent genetic models. *G3: Genes, Genomes, Genetics* **9**: 2107–2122.

Bourke PM, Voorrips RE, Hackett CA, van Geest G, Willemsen JH, Arens P, Smulders MJ, Visser RG, Maliepaard C. 2021. Detecting quantitative trait loci and exploring chromosomal pairing in autopolyploids using POLYQTLR. *Bioinformatics* **37**: 3822–3829.

Broman KW, Wu H, Sen Ś, Churchill GA. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* **19**: 889–890.

Chen J, Leach L, Yang JX, Zhang FJ, Tao Q, Dang ZY, Chen Y, Luo ZW. 2021. A tetrasomic inheritance model and likelihood-based method for mapping quantitative trait loci in autotetraploid species. *New Phytologist* **230**: 387–398.

Hackett CA, McLean K, Bryan GJ. 2013. Linkage analysis and QTL mapping using SNP dosage data in a tetraploid potato mapping population. *PLoS ONE* **8**: e63939.

Haley CS, Knott SA. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **69**: 315–324.

Lander ES, Botstein D. 1989. Mapping mendelian factors underlying quantitative traits using Rflp linkage maps. *Genetics* **121**: 185–199.

Lander ES, Green P. 1987. Construction of multilocus genetic-linkage maps in humans. *Proceedings of the National Academy of Sciences, USA* **84**: 2363–2367.

Martinez O, Curnow RN. 1992. Estimating the locations and the sizes of the effects of quantitative trait loci using flanking markers. *Theoretical and Applied Genetics* **85**: 480–488.

Zheng C, Voorrips RE, Jansen J, Hackett CA, Ho J, Bink MC. 2016. Probabilistic multilocus haplotype reconstruction in outcrossing tetraploids. *Genetics* **203**: 119–131.

Zheng CZ, Amadeu RR, Munoz PR, Endelman JB. 2021. Haplotype reconstruction in connected tetraploid F1 populations. *Genetics* **219**: iyab106.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Notes S1 R script to replicate the timing and output comparison of TetraQTLAnalysis and POLYQTLR (requires R software to run).

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Key words: autopolyploid, double reduction, interval mapping, polysomic inheritance, quantitative trait loci (QTL) analysis.

Received, 23 February 2023; accepted, 17 November 2023.