



QTL detection in a pedigreed breeding population of diploid potato

George Korontzis · Marcos Malosetti · Chaozhi Zheng · Chris Maliepaard · Han A. Mulder · Pim Lindhout · Roel F. Veerkamp · Fred A. van Eeuwijk

Received: 15 March 2020 / Accepted: 2 August 2020 / Published online: 26 August 2020
© The Author(s) 2020

Abstract Diploid hybrid potato breeding is emerging as an alternative to breeding tetraploid potato clones. The development of diploid breeding varieties involves recent, shallow pedigrees with a limited number of founders. Within this context, alternative QTL detection methodologies should be considered to enable identification of relevant QTLs and characterize the founders of the pedigree. To that end, we are using a dataset of multiple diploid potato F_3 families under selection derived by a cross between an inbred *Solanum chacoense* and an outbred diploid *Solanum*

tuberosum, and identify QTLs for tuber fresh weight. We used three methods for QTL detection: (1) a Genome Wide Association Study model, (2) a linkage approach tailored to the population under study and (3) a more general approach for modelling multiallelic QTLs in complex pedigrees using identity-by-descent (IBD) probabilities. We show that all three approaches enable detection of QTLs in the population under study, but the method that makes better use of IBD information has a more direct and detailed interpretation by linking QTL alleles to the founders.

G. Korontzis (✉) · M. Malosetti ·
C. Zheng · F. A. van Eeuwijk
Biometis, Wageningen University and Research,
Droevendaalsesteeg 1, P.O. Box 16,
6700 AA Wageningen, The Netherlands
e-mail: George.korontzis@wur.nl

F. A. van Eeuwijk
e-mail: fred.vaneeuwijk@wur.nl

C. Maliepaard
Plant Breeding, Wageningen University and Research,
Droevendaalsesteeg 1, P.O. Box 386,
6700 AJ Wageningen, The Netherlands

H. A. Mulder · R. F. Veerkamp
Animal Breeding and Genomics, Wageningen University
and Research, Droevendaalsesteeg 1, P.O. Box 338,
6700 AH Wageningen, The Netherlands

P. Lindhout
Solynta, Dreijenlaan 2, 6703 HA Wageningen, The
Netherlands

Keywords QTL mapping · Diploid potato · Hybrid potato · GWAS · Identity-by-descent · Pedigree

Introduction

Diploid hybrid potato and QTL mapping

Diploid hybrid potato breeding promises a paradigm shift in potato breeding consisting of replacing the tetraploid clones with diploid hybrids propagated by true seed (Lindhout et al. 2011; Jansky et al. 2016). This transition is based on the development of self-compatible diploid potato germplasm and, subsequently, the gradual selection against the genetic load through selfing. The resulting inbred lines enable the exploitation of heterosis through hybrid breeding, the introgression of genes through backcrossing and the

utilization of all breeding methods and genomic tools developed for diploid crops. Besides the consequences regarding breeding and crop management (Lindhout et al. 2018), this shift also influences quantitative trait loci (QTL) mapping methodology.

In potato, QTLs have been mapped using both tetraploid and diploid populations. QTLs in the tetraploid level were detected initially by association mapping (Gebhardt et al. 2004; Simko et al. 2004), because of the complex tetrasomic inheritance that delayed the development of proper methodology and software for linkage mapping (Bradshaw et al. 2008; McCord et al. 2011; Hackett et al. 2013; Massa et al. 2015; Bourke et al. 2015, 2018). Therefore, most linkage studies in potato have been conducted at diploid level, in F_1 populations of biparental crosses involving heterozygous parents (full-sib families) (Gebhardt et al. 1989; Schäfer-Pregl et al. 1998; Manrique-Carpintero et al. 2015). Selfing was hindered by the self-incompatibility and inbreeding depression of diploid potato and therefore the use of selfed material has been limited to the study of segregation distortion and self-fertility (Peterson et al. 2016; Zhang et al. 2019).

A change in QTL mapping methodology driven by diploid hybrid breeding is the use of inbred material in the mapping population either as parents (Endelman and Jansky 2016) and/or as offspring, such as F_2 (Endelman and Jansky 2016; Meijer et al. 2018) or sib-mated F_2 (Braun et al. 2017). In this article we focus on another aspect of QTL mapping in the “diploid hybrid” era, which is the development of diploid potato breeding populations. By the term breeding populations here we refer to populations with: (1) population structure derived by selection and genetic drift and (2) more than two alleles possibly segregating for a QTL. Unlike conventional potato breeding programs, but also unlike many other crops, the relevant pedigrees in diploid hybrid potato breeding are shallow, very detailed and they include a limited number of founders. These conditions can allow for linkage mapping directly in the breeding population ensuring relevance of the identified QTLs.

Within this context, QTL mapping methods developed for inbred-derived biparental populations, single mapping populations and experimental populations are not directly applicable because of their inability to model multiple alleles and/or to allow for the

population structure. We can identify three suitable frameworks:

1. Genome Wide Association Studies (GWAS)
2. Joint analysis of families (we will refer to this as Stratified Linkage Analysis—SLA)
3. QTL mapping in complex pedigrees (we will refer to this as General Linkage Analysis - GLA)

Those three approaches vary in terms of (i) the number of QTL alleles modelled, (ii) the use of Identity-by-Descent (IBD) or Identity-by-State (IBS) information, (iii) the way IBD is inferred, (iv) the way population structure is taken into account and (v) their generality.

Genome-wide association studies

GWAS methodology (also called Linkage Disequilibrium mapping or Association mapping) was developed for large panels of individuals with varying relatedness. It consists in estimating a substitution effect of a biallelic QTL under a correction for population structure. Population structure correction is of crucial importance to avoid false positives. In most cases, this correction is achieved by including a genomic relationship matrix to model genetic covariance of individuals through a random polygenic term, although several other approaches have been proposed (Yu et al. 2006; Sneller et al. 2009; Würschum 2012). Further elaboration of the methodology has aimed at avoiding confounding between the tested polymorphism and the relationship matrix by excluding from the estimation of relationship all markers located in the tested chromosome (Rincent et al. 2014; Yang et al. 2014).

The main advantage of GWAS is its simple application in all types of populations without the need of explicit pedigree information. Its main disadvantage is that it does not make use of IBD information. Modelling the QTL alleles only using IBS information is based on the assumption that all individuals carrying the same marker allele also carry the same QTL allele, irrespective of the origin of the marker allele. Furthermore, the use of IBD information through linkage mapping would enable modelling multiallelic QTLs, performing interval mapping and directly linking the QTL effects to the pedigree founders.

Stratified linkage analysis

By the term Stratified Linkage Analysis we refer to a wide group of approaches for joint analysis of families that use IBD information and estimate multiple effects per QTL by exploiting the population stratification. Typical examples of populations where these approaches are applicable are Nested Association Mapping (NAM) populations (Yu et al. 2008; Buckler et al. 2009; Giraud et al. 2014; Garin et al. 2017) and populations derived from a diallel cross (Rebai and Goffinet 1993; Jannink and Jansen 2001; Blanc et al. 2006). Although in such inbred-derived multiparental populations more than two QTL alleles might be present, IBD estimation is rather straightforward because only two founder alleles segregate in each family. Multiple QTL effects corresponding to the founder alleles can be estimated by inferring the parental origin and in the case of NAM populations this is equivalent to fitting within family an indicator expressing the contrast between common and peripheral parent.

Population correction usually consists in including a family specific mean and a heterogeneous genetic residual variance. However, in the presence of selection, genetic covariance correction is reported to be appropriate for linkage mapping populations as it is for association panels (Malosetti et al. 2011).

SLA approaches have limited generality as they are developed for specific types of populations and the explicit modelling of population structure demands large families. With increased pedigree depth and complexity, two problems arise: (i) IBD estimation is not straightforward and demands more elaborate methods and (ii) the number of parameters to be estimated (for explicit modelling of population stratification and multiple QTL alleles) increases.

General linkage analysis

By General Linkage Analysis we refer to methods for linkage analysis in more complex populations. The main element of those methods is the more elaborate estimation of IBD probabilities in complex pedigrees that are deeper, with more founders (potentially outbred), more families and fewer individuals per family (Thompson 2000; Almasry and Blangero 1998; Abecasis et al. 2002; Bink et al. 2002; Zheng et al. 2014, 2015, 2018). Because of the complexity of IBD

estimation in these methods, they are sometimes called IBD-based approaches (Xie et al. 1998; Crepieux et al. 2004), although all linkage mapping methods use IBD information (unlike association mapping methods such as GWAS that use only IBS). Estimation of IBD probabilities results in an n by f matrix, where n is the number of individuals and f is the number of founder alleles. As in every linkage method, IBD can allow for interval mapping. Because of the possibly high number of founders these approaches usually fit a random multiallelic QTL effect to decrease the number of parameters to be estimated. For this reason this approach is usually called random model (Xu and Atchley 1995) or variance component method (George et al. 2000). However, if the number of founder alleles is small, IBD probabilities can also be fitted as fixed (Wei and Xu 2016), as long as the IBD probabilities to the founders are available and not only the pairwise IBD matrix of the mapping population. Correction for population structure is usually achieved using a genomic relationship matrix, as in GWAS. These methods were initially developed for animal genetics (Fernando and Grossman 1989) but have also been applied, as an alternative to GWAS, to complex plant multiparental populations such as MAGIC (Verbyla et al. 2014) and pedigreed plant breeding populations (Parrisseeux and Bernardo 2004; Crepieux et al. 2004; van Eeuwijk et al. 2010).

The advantage compared to GWAS is the use of IBD information that allows modelling multiallelic QTLs and interval mapping. Compared to SLA, the superiority of GLA lies in the generality of the approach that can allow for routine application in pedigreed breeding populations and increase the relevance of the identified QTLs. Furthermore, the explicit estimation of founder QTL allele effects, enables tracing the origin of the QTL alleles and characterizing the founders. The main disadvantage of the method is the need for accurate pedigree information.

GLA is well suited to the recent and rather simple pedigree of diploid potato, where pedigree information is available and characterizing the founders is essential for the further development of the diploid potato germplasm.

Objectives

In this study, we develop methodology for QTL detection in pedigreed breeding populations, focusing on shallow pedigrees, derived by few possibly outbred founders. To that end, we will use a population described by Meijer et al. (Meijer et al. 2018). This population is a set of F₃ families with a defined pedigree structure derived from a cross between an inbred (*Solanum chacoense*) and an outbred parent (*Solanum tuberosum*). It is a diploid potato QTL mapping population with exceptionally high level of inbreeding. We here use it as a proxy of a complex plant breeding pedigree on the grounds that it has more than two alleles segregating, it is under selection and it represents an important part of the breeding germplasm. The structure and complexity of the segregating population implies that no standard method for QTL mapping is immediately applicable in this population. The objective here was to compare the following three different approaches to map QTLs in this population: (1) a GWAS approach ignoring IBD information, (2) a simplified IBD-based approach tailored for this population (SLA), and (3) a general IBD-based approach (GLA) (Table 1).

Materials and methods

Population and data

The mapping population consists of 840 individuals originating from a cross between an inbred *S. chacoense* genotype (we will refer to it as the “Sc” parent), and a highly heterozygous diploid *S. tuberosum* genotype (the “St” parent) (Meijer et al. 2018). The mapping population is the third generation of this cross, as shown in Fig. 1. Specifically, two F₁ individuals from the “Sc” × “St” cross were selected and each one selfed to give rise to two main families (called “Fam1” and “Fam2”). From the F₁ plant that originated Fam1, six F₂ individuals were selected and selfed to produce six F₃ subfamilies (F₁ plant 1 in Fig. 1). The F₁ plant that produced the Fam2 was also selfed but only three F₂ plants were selected to give three F₃ subfamilies (F₁ plant 2 in Fig. 1). Each of the nine F₃ subfamilies consisted of approximately 100 individuals.

Note that since “Sc” is highly homozygous this parent contributes only one allele to the population, which comes from a *S. chacoense* background (we will refer to it as the “Sc” allele). On the other hand the parental genotype “St” contributes (when heterozygous at that locus) two alleles from a *S. tuberosum* background (“St_a” and “St_b” alleles). Therefore, at a particular locus, up to three alleles can be segregating in the entire population. Also, and because of the crossing scheme, within each of the F₃ subfamilies at most two alleles can segregate, the “Sc” and one of either “St_a” or “St_b”. A final important remark is that the population is under selection forces, natural because of inbreeding depression and artificial as result of the regular breeding process (only best individuals were advanced). Consequently, the population departs from the usual assumptions of a standard segregating mapping population.

All 840 individuals were evaluated in a greenhouse experiment for a number of traits as described by Meijer et al. (2018). Here, we focus on total tuber fresh weight per plant (TFW). TFW is not expected to be an accurate estimator of yield per hectare, but it is a quantitative trait and as such suitable for a comparison of QTL models. To enable interpretation of the results, effects will be reported as a percentage of the population mean. TFW showed continuous variation and is expected to have a polygenic inheritance. In addition, all the individuals and the two founders were genotyped, resulting in 120 SNP markers being polymorphic within all or some of the F₃ subfamilies as described by Meijer et al. (2018). Those SNPs were mapped on a linkage map with 12 groups and a total length of 817.7 cM.

Statistical models for QTL detection

The Sc allele is chosen as reference level since the Sc parent is inbred and the Sc allele present in both families enabling comparison of the three approaches. In the GWAS model (Table 1), the putative QTL allele effect is modelled on the basis of IBS as the substitution of an Sc SNP allele by a different one. In SLA, marker data are simplified as explained below to resemble those of a biparental inbred-derived population and express the contrast between the two founders ignoring heterozygosity of the St founder. The informative marker data are recoded into AA, AB or BB where A denotes the Sc allele and B the St allele

Table 1 Statistical models for single QTL scan with the three approaches

Terms	Description
GWAS	$y_{ijk} = \mu + x_{ijk}\alpha + u_{ijk} + e_{ijk}$ (1)
y_{ijk}	Response of individual k belonging to family i and subfamily j
μ	Mean
x_{ijk}	Marker state (number of Sc SNP alleles)
α	Fixed SNP allele substitution effect
u_{ijk}	$u_{ijk} \sim N(0, G\sigma_u^2)$ random polygenic term
G	Genomic relationship matrix
e_{ijk}	$e_{ijk} \sim N(0, \sigma_e^2)$ residual term
SLA	$y_{ijk} = \mu + \phi_i + s_j + x'_{ijk}\alpha'_i + e'_{ijk}$ (2)
ϕ_i	Fixed effect of family i
s_j	$s_j \sim N(0, \sigma_s^2)$ random effect of subfamily j
x'_{ijk}	$x'_{ijk} = Pr(QQ M_i) - Pr(qq M_i)$ interval mapping genetic predictor for inbred biparental populations based on informative markers
α'_i	Effect of substitution of an inbred founder allele by one of the outbred founder alleles within family i
e'_{ijk}	$e'_{ijk} \sim N(0, \sigma_{e_j}^2)$ genetic residual term with heterogeneous variance for subfamilies
GLA	$y_{ijk} = \mu + (x''_{ijkf})^T \alpha_f + u_{ijk} + e_{ijk}$ (3)
x''_{ijkf}	Vector of three IBD probabilities (indexed by f) for individual k of family i and subfamily j
α_f	Vector of the three founder allele effects

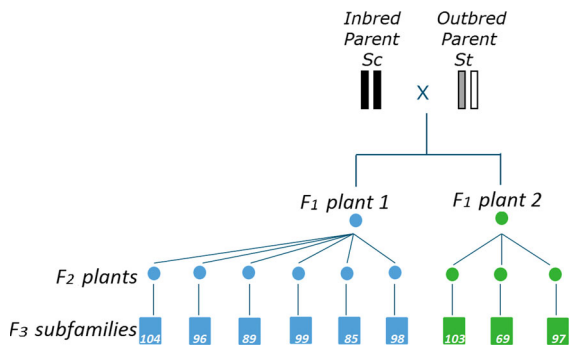


Fig. 1 Pedigree structure. Three founder origins denoted with black (“Sc”), gray (“St_a”) and white (“St_b”). Two families derived by two different F₁ plants (Fam1 on the left, Fam2 on the right). Number of F₃ individuals of each subfamily noted inside the box

(St_a or St_b). Markers for which the St parent is homozygous are informative for all F₃ individuals. Markers for which the St parent is heterozygous are informative for the F₃ individuals that originate from an heterozygous F₁ plant. Subsequently, the standard procedure of R qtl for biparental inbred-derived F₃ populations was followed to estimate genotype

probabilities for every marker position and every 5cM (Broman et al. 2003). A genotypic predictor expressing the substitution of an Sc allele by an St allele was calculated as $Pr(BB | M_{ijk}) - Pr(AA | M_{ijk})$ where M_{ijk} is the recoded marker data. Given that only one St allele can be present in each family, this genotypic predictor was fitted within family, implicitly modelling a multiallelic QTL in our SLA approach. In GLA, the marker data of founders and offspring were combined with the pedigree of the population to estimate the probabilities of a randomly drawn allele from each of the 840 F₃ individuals being identical by descent with each of the three founder alleles using the RABBIT package (Zheng et al. 2015). The resulting 840 by 3 matrix allows for modelling explicitly three QTL alleles across families. The estimation was done for every marker position and for a grid of 5cM. Although the marker data were initially filtered for low minor SNP allele frequency, IBD probabilities might indicate that one of the three founder alleles is absent from the population. This can happen either because of transmitting the same St allele to both F₁ plants or due to strong selection in

subsequent generations against any of the three alleles. Therefore, we estimated frequencies of each founder allele by averaging the IBD probabilities over the population and filtered again with a threshold of 0.05 setting IBD probabilities of the low frequency founder to 0 and rescaling the other two to sum to 1. To enable comparisons with the other two methods, the Sc allele was chosen as reference level and probabilities were multiplied by two to estimate the effect of substitution of an Sc allele.

The differences between those three ways of modelling the putative QTL are explained following the example of Fig. 2. For each individual in Fig. 2 we see the marker genotype along with the number of copies of each of the three different founder alleles (“Sc”|“St_a”|“St_b”). If the two F₁ plants have inherited different St alleles, the segregating marker state does not correspond to the origin of the allele and plants 1.3.1, 1.3.2 and 1.3.3 are grouped together although they might be carrying different QTL alleles. SLA and GLA allow this distinction by using flanking marker information but they differ in the fact that SLA estimates QTL effects within family assuming that different St alleles are transmitted. If we consider a similar example (not shown here) with the difference that both F₁ plants have inherited the second St allele (St_b) and have genotype GA, SLA will still use the same genotypic predictor as in the previous example and report two effects but GLA will not estimate an effect for one of the two St alleles that will have (close to) zero probability for all individuals. In that case, GLA would properly use information across families by estimating one effect using all individuals.

Regarding the correction for population structure, in GWAS and GLA the genetic covariance approach was followed, including in the model a polygenic term with genetic covariance modelled by a genomic relationship matrix as described by VanRaden (first method) (VanRaden 2008) and implemented in the synbreed R package (Wimmer et al. 2012),

$$G = \frac{ZZ'}{2 \sum^M p_m(1 - p_m)} \tag{4}$$

where $Z = X - P$, X is the marker matrix coded as (0, 1, 2), P the non-reference allele frequencies multiplied by two, and p_m the allele frequency at marker m . Rincent et al. (2014) showed that the drawback of a single genome-wide relationship matrix is that

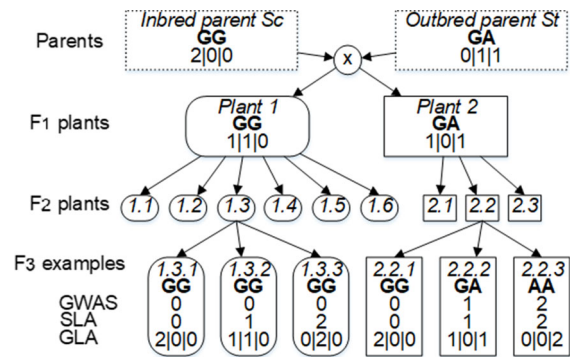


Fig. 2 Example of genotype coding of putative QTL. For illustration only 6 F₃ individuals are shown here. Identifier for individual (in italic) and example SNP marker genotype (here a G/A SNP) (in bold). The three numbers (e.g. 0|1|1) correspond to the actual number of copies of each of the three founder alleles (“Sc”|“St_a”|“St_b”). At the bottom, the three different ways of modelling the putative QTL. In GWAS, as number of copies of the non-reference SNP allele (inbred Sc allele as reference). In SLA, as number of copies of allele of St origin. In GLA, as number of copies of each founder allele

markers used in the computation of matrix G interfere with the signal at the tested position, thus reducing detection power. Therefore, as an alternative, a set of 12 chromosome-specific G^c matrices were obtained by excluding the SNPs of the corresponding chromosome c . So, while performing a GWAS or GLA scan, the relationship matrix used was the one that did not include the markers on the chromosome being scanned (this implies that matrix G had to be updated when moving from one chromosome to the next). In all cases, semi-positive definiteness of the G matrix was checked, and if necessary an approximation of the matrix was used, replacing non-positive eigenvalues by a small positive constant (Piepho et al. 2008). In SLA, correction for population structure was focused on the known stratification, by fitting a fixed family effect for the two families and a random subpopulation effect for the nine subfamilies. In addition, the genetic residual variance was allowed to vary between subfamilies.

All three models were fitted in ASReml-R (Butler et al. 2009) and a Wald test was used to assess association between genotype variation and trait variation. To map the QTLs, the corresponding P values on a $-\log_{10}$ scale were plotted against map positions. A Bonferroni-type multiple testing correction threshold was defined following Li and Ji (2005) to control the experiment-wise error rate at 0.05. When

two significant positions were found on the same chromosome, they were considered to be linked to different QTLs when they were spaced by at least 30 cM. For each of the three approaches, after an initial search by the single-QTL models of Table 1, all significant QTLs were combined in a multi-QTL model, and a backward elimination process was implemented. In GWAS and GLA, the set of chromosome specific genomic relationship matrices was replaced by a G matrix estimated by all chromosomes without QTLs. For the backward elimination, a conditional Wald test was used, dropping QTLs from the full model whenever the contribution of the particular QTL was found not significant at a significance level of 0.05. The remaining significant QTLs were retained in a final multi-QTL model that was used to produce estimates of allele effects.

Results

In both approaches that include a polygenic term (GWAS and GLA), we observed an overall increase in significance when using a chromosome-specific instead of a standard genomic relationship matrix. This result was expected, as including also markers physically linked with the putative QTL when estimating the genomic relationship matrix results in the genetic background interfering with the specific position that is being tested. The scatter plots in Fig. 3 where most of the values are above the 1:1 line, show that identification of QTLs on a $-\log_{10} p$ value threshold of 3, would be impossible with a standard G matrix.

The overall trend of the QTL mapping profiles was consistent across methods indicating presence of QTLs in chromosomes 1, 2 and 3 (Fig. 4). However, inconsistencies were also observed, reflecting differences between the three approaches.

GWAS approach

The results of the GWAS approach with a chromosome-specific relationship matrix (Fig. 4) show peaks above the threshold on the end of chromosome 1, beginning and middle of chromosome 2, beginning and middle of chromosome 3 and beginning of chromosome 5. Because the two peaks on chromosome 2 were spaced by less than 30 cM, only the

largest one (the one at the beginning of the chromosome) was selected for the next modelling step. All five candidate positions were found significant when fitted together in a multi-QTL model, and their positions and estimated effects are shown in Table 2. We observe that for all but one of the significant SNPs, the allele originating from the *S. tuberosum* background increased TFW (positive sign). Only for the QTL on chromosome 3, the superior allele was the one from *S. chacoense*. The largest effect was of the QTL on chromosome 5, with a point estimate of 22.45% of the population TFW mean. However, it was also the one with the higher standard error (clearly higher than for the estimates of other QTL effects) and its effect might be overestimated because of its low minor allele frequency (0.051).

SLA approach

The SLA profile in Fig. 4 shows similarities and differences with the results from the GWAS approach. Significant signal was found on chromosomes 2 and 3 as with the GWAS, but not on chromosome 1 and 5 (although on chromosome 1 a peak close to the threshold was observed). In comparison with the GWAS result, an extra peak was found on chromosome 12. The QTL on chromosome 12 should be interpreted with caution, given that it is driven by a single marker in the end of the chromosome. The

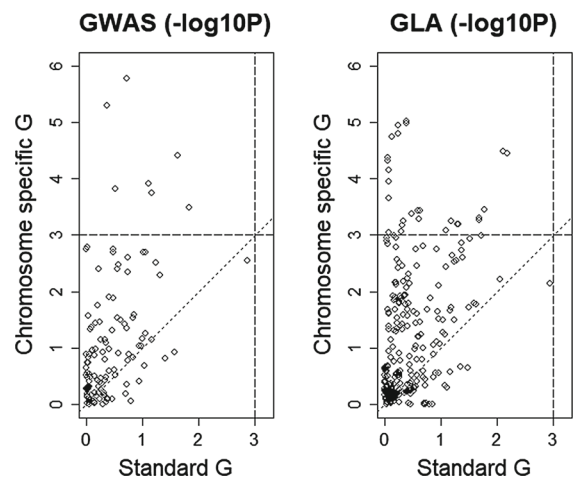


Fig. 3 Comparison of $-\log_{10} p$ values between the two ways of modelling the polygenic term (standard G matrix on the x axis and chromosome-specific G matrix on the y axis) for the GWAS (left) and the GLA approach (right)

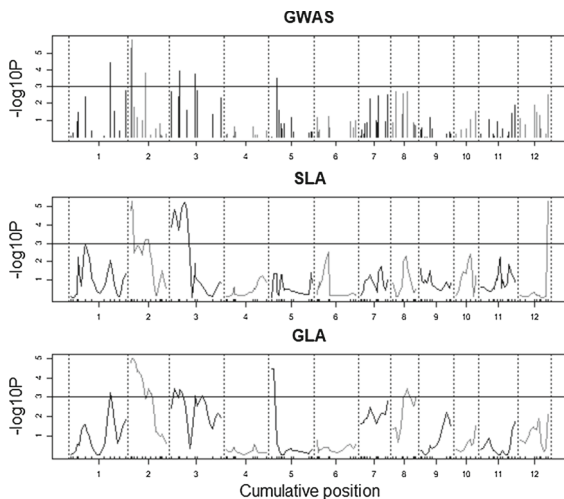


Fig. 4 QTL detection profiles for tuber fresh weight showing the $-\log_{10} p$ values along the 12 chromosomes from single-QTL scans with the three approaches. Li and Ji threshold of 3.0 depicted

decreased significance around the QTL in the beginning of chromosome 5, that could be related to the maturity gene *StCDF1*, can be attributed to the loss of information when the state of markers is not enough to infer IBD because of heterozygosity in the St founder and homozygosity in the F_1 plant. While the two peaks on chromosome 2 were more than 30 cM apart and so, considered as two different candidate QTLs, the two peaks on chromosome 3 were at a shorter distance so only the strongest position was selected as candidate QTL.

We started the final model selection with five candidate QTLs (including the peak on chromosome 1, which was just below the threshold). When fitted together the second peak of chromosome 2 (position 35 cM) was found not significant and was removed from the model. Table 2 shows the QTL effects estimated from the final multi-QTL model. Two effects are estimated for each QTL, corresponding to the two families (Fam1 and Fam2). For the QTL on chromosome 1 the allele substitution effect was significant in Fam1 but not in Fam2. We can also observe that the effect had a negative sign, which points to the *S. chacoense* allele (“Sc”) as the one increasing TFW. For the QTL on chromosome 2, both alleles showed a significant effect and were consistent in pointing to the “St” allele as the one increasing TFW (positive sign). Note that the effect of this QTL is consistent with the QTL previously identified by the

Table 2 Locations, allele effects and standard errors of the identified QTLs with the three approaches. Effects are estimated from the final multi-QTL model and expressed as percentage of the population mean. GWAS effects refer to substitution of an Sc SNP allele by a different SNP allele, SLA effects refer to substitution of an allele of Sc origin by an allele of St origin within the two families and GLA effects refer to substitution of a Sc allele by an St_a or an St_b allele

GWAS

Chr	Pos	Effect	<i>p</i> value
chr_01	76	12.07 (2.71)	0.000008
chr_02	3.2	11.28 (2.47)	0.000005
chr_03	16	− 8.52 (2.69)	0.001566
chr_03	46	8.26 (2.60)	0.001499
chr_05	10	22.43 (6.15)	0.000267

SLA

Chr	Pos	Effect Fam1	Effect Fam2	<i>p</i> value
chr_1	27	− 8.95 (2.47)	− 0.96 (5.67)	0.037198
chr_2	3.2	10.21 (2.34)	13.91 (5.63)	0.001818
chr_3	25	− 11.61 (2.85)	23.37 (7.77)	0.001645
chr_12	54	1.31 (4.47)	34.22 (6.52)	0.000998

GLA

Chr	Pos	Effect St_a	Effect St_b	<i>p</i> value
chr_01	76	11.57 (3.20)	12.40 (4.41)	0.009490
chr_02	5	14.74 (6.04)	10.15 (2.83)	0.008267
chr_03	5	− 12.10 (2.86)	−	0.000024
chr_03	46	6.56 (2.55)	−	0.009999
chr_05	5	− 2.98 (2.70)	28.00 (6.95)	0.008553
chr_08	27	11.65 (4.47)	10.60 (3.05)	0.029837

GWAS model, which also pointed to the “St” allele as the superior one. A more contrasting case was observed for the QTL on chromosome 3, where both QTL alleles were significant but the sign opposite. So, within Fam1 the “Sc” allele increased TFW, while the opposite was true for the allele segregating in Fam2 (the “St” increased TFW). Because the reference allele is always the same (the “Sc” allele), differences in the estimated effects within families both in magnitude and sign suggests that two different *S. tuberosum* alleles may have been inherited in the population, one of them within Fam1, and the other

one within Fam2. We can also conclude that the best allele at this particular QTL is the *S. tuberosum* allele that segregates within Fam2, followed by the “Sc” allele, and finally the second *S. tuberosum* allele that happened to segregate within Fam1. We should highlight this type of extra information available by switching from a GWAS approach to the one here, where not only it is possible to rank up to three alleles but also trace in which part of the pedigree a particular allele can be found.

GLA approach

The last profile plot in Fig. 4 shows the results of QTL detection by the GLA approach. Seven peaks were found above the significance threshold. In comparison with the GWAS and SLA approaches, the GLA approach seems to suggest a larger set of candidate QTLs (7 vs. 5). The GLA model detected five candidates similar to the ones GWAS did, plus two additional ones. After fitting all seven candidate positions, one position (chromosome 2 at 35 cM) was dropped out from the model, resulting in a final multi-QTL model with six QTLs.

The allele substitution effects as estimated from the final multi-QTL model are given in Table 2. One first difference to point out is that with the GLA model an effect can be estimated for each of the two *S. tuberosum* alleles (“St_a” and “St_b”). This is an improvement over the SLA model as it gives an explicit estimate for each of the two alleles. For each QTL, two allele effects were estimated, except when one of the two *S. tuberosum* alleles was not present in the population (or present at a low frequency). A first general observation is that for most of the QTLs, the favourable alleles (i.e. the one increasing TFW) were from the *S. tuberosum* background, since most of the significant effects had a positive sign, which translates as an improvement (increase) over the reference *S. chacoense* allele. We also observed that for two of the QTLs only one of the two *S. tuberosum* alleles was segregating, which is a possible indication of either selection, the effect of the strong bottleneck effect in this population (started from two F₁ plants), or both.

Some extra insight about the QTLs was possible when applying the GLA model in comparison with the SLA model. For example, with the SLA model, a QTL on chromosome 2 was detected, and two effects estimated, one within Fam1 and another one within

Fam2 (10.21 and 13.91% of the population TFW mean respectively). However, from that analysis we could not conclude whether the estimates were for two different *S. tuberosum* alleles or just one allele estimated within the two different families. The results from the GLA analysis shed some more light on it, as the IBD information showed that two different *S. tuberosum* haplotypes were inherited in the two families in that region. This does not prove the presence of three different QTL alleles, but it supports that hypothesis by excluding the case of the same *S. tuberosum* segment being transmitted. Another example of additional information about QTLs was observed on chromosome 3. With the SLA model, one QTL at 25 cM on chromosome 3 was found, and two effects were obtained: – 11.61% within Fam1 and 23.37% within Fam2. With those results, the suggestion was that two different alleles must be segregating at this particular QTL. However, the results from the GLA analysis suggest a different conclusion involving two QTLs with only one *S. tuberosum* allele in each QTL. The GLA analysis suggests one QTL at 5 cM with an effect of – 12.10%, and a second one at 46.4 cM with an effect of 6.56%. The difference between the two analyses, is that in the GLA analysis IBD information reveals that it is more likely that only one *S. tuberosum* allele is segregating within these regions on chromosome 3. Therefore, the hypothesis of two closely linked but different QTLs is more likely than the hypothesis of a single QTL. The result from the SLA model is akin to the well-known effect of a ghost QTL detected somewhere in-between two close-by QTLs. The difference between the two family-specific effects observed in SLA must have been the result of limited segregation in Fam2. The QTLs detected on chromosomes 1, 2, 5 and 8 seemed to segregate with three alleles each. Again, the possibility to rank the alleles and establish the possible genetic background source is an added value of the GLA approach. For example, we can identify within the population those individuals that have a higher probability of carrying the right version of the “St” allele for selection purposes.

Discussion

The combination of population structure, possible segregation of multiple QTL alleles and availability of

pedigree information differentiates this population from commonly used mapping populations in potato breeding. However, we were able to identify QTLs with three approaches. Applying GWAS models is a straightforward solution but has the limitation of only modelling biallelic QTL effects and only at marker positions (no interval mapping). Our SLA approach is a better alternative, that exploits the known population structure, in a tailored-made approach to model multiple alleles indirectly by nesting effects within a particular level of the population structure. While a possible option, it can only be applied in simple pedigrees like the one shown here or inbred-derived NAM populations. However, as we showed with our example, the SLA approach can miss useful information since it needs to assume that different alleles segregate in each family. By comparison of the estimated allele effects one can indirectly hypothesize similar or different alleles depending on how different the estimated effects are in the different parts of the pedigree. In contrast, the GLA approach proposed here directly uses IBD estimation from markers and pedigree data to design the incidence matrix that is used to test for QTLs and to estimate QTL effects. In other words, the ancestry of the alleles that flow along the pedigree is better reflected in the GLA QTL detection model.

False marker-trait associations caused by confounding effects with the genetic background is a well understood phenomenon in GWAS studies. Examples of linkage mapping using breeding populations under selection have also described the detection of spurious QTLs because of the genetic background effect (Kennedy et al. 1992; Malosetti et al. 2011). In such cases, the use of a genomic relationship matrix offers flexibility to modelling genetic background effects that result from population stratification and cryptic relatedness. Another advantage of including a polygenic term in our model is that it decreases the need for composite interval mapping and multi-QTL models, in the sense that each putative QTL is tested while correcting for the genetic background. However, it is reasonable that major QTLs will be more properly modelled as fixed cofactors than through a random polygenic term.

We observed that using a chromosome specific G matrix that excludes markers at the particular chromosome helps to increase the signal given by putative QTLs, in agreement with previous studies (Rincen

et al. 2014; Yang et al. 2014). The method is straightforward to apply when performing a QTL search based on a single QTL model like we did here. Some further elaboration will be needed if the QTL model includes QTLs in most of the chromosomes. In such cases, the estimation of the genetic relationship matrix should at least exclude the markers in the vicinity of the QTLs that are part of the model. In those cases, a decision has to be made regarding the size of the window around the QTL within which to exclude the markers, for example using global or local LD decay information. An alternative approach is the one described by Wei et al., consisting in adding the polygenic counterpart (BLUP) of the tested marker to the response, resulting in a marker-specific response vector (Wei and Xu 2016).

We suggest the use of IBD probabilities to explicitly model QTL alleles in reference to the founders. Software for linkage analysis in complex pedigrees is suitable as long as it allows the incorporation of all aspects of the relevant pedigree, including selfing and it reports the probabilities of each allele in the offspring being IBD to each founder allele. An IBD-approach with random QTL effects can be performed using only the pairwise IBD matrix of the offspring but best linear unbiased prediction of QTL effects will not be possible. Similar approaches can be used for non-pedigreed populations, by using conserved haplotype segments as IBD evidence. In that case, the interpretation of founder allele effects is not relevant but modelling of multiple QTL alleles can still give further insight. In more complex pedigreed populations, fixed QTL effects will lead to an increased number of parameters to be estimated. Therefore, it would be more appropriate for QTLs in complex pedigrees to be modelled as random effects. Other extensions of the methodology include the clustering of founder haplotypes locally into a smaller set of ancestral haplotypes (Meuwissen and Goddard 2001; ter Braak et al. 2010; Leroux et al. 2014) and the combined use of different QTL models (Garin et al. 2017).

We should note a difference between IBD-based methods and IBS methods. IBD probabilities include in the analysis uncertainty that is not only related to marker calling but also to the transmission of genes in the pedigree and the genotype between markers. Consequently, QTL alleles are modelled by a vector of continuous values and the interpretation of allele

frequency changes as it does not always refer to the number of alleles but to the average probability of an allele. Therefore, in regions of greater uncertainty, low to moderate allele frequency can be observed in cases of alleles that are in fact absent from the population, leading to spurious results. In a less extreme case, where the allele is actually present but the uncertainty drives IBD probabilities towards the expectation from the pedigree, significance of QTL presence might not be affected, but the allele effect will be overestimated. Further research in the use of IBD probabilities should focus on assessing the accuracy of allele effects estimation.

When IBD information is used for estimation of multiple founder allele effects, each founder allele effect should be modelled by a vector that shows allelic variation and allows its separation from other founder alleles. It is advised to explore the IBD probabilities and if necessary exclude or treat positions that are not suitable for this analysis. Such cases include: (i) IBDs that show presence of the allele in only a few individuals, (ii) IBDs that are close to zero for all individuals because one of the founder alleles is selected against and is not present in the mapping population or (iii) IBD probabilities of two founder alleles that are approximately equal to each other for all individuals because of extended IBS between founder haplotypes (e.g. homozygosity of an outbred founder) causing colinearity.

It is important to develop proper criteria for evaluation of marker quality and information content of IBD probabilities in breeding populations. This is because the minor allele frequency is not an adequate measure of polymorphism in populations with heterozygosity or uncertainty of origin and the genotypic information coefficient (van Ooijen 2009; Bourke et al. 2019) is not directly applicable to populations under selection.

It should be noted that the rationale behind the design of SLA was to use population stratification for modelling both the QTL and the population structure. Similarly, GLA was aiming at offering a general approach to those two model elements. Consequently, here we compared those two approaches, but further research is needed to study each of those two elements separately. Although applying SLA with a polygenic term, instead of the explicit modelling of population stratification, affects the QTL profile (results not

shown here), it does not affect the validity of the discussion on the interpretation of QTL effects.

The aim of this study was to demonstrate the application of those different approaches that are applicable to breeding pedigrees and discuss the interpretation of the estimated effects. However, we should stress the importance of validation of reported QTL before application and the benefits of densely genotyped and accurately phenotyped experimental populations for deeper understanding of genetic architecture.

In diploid potato, further insight should be gained by extending QTL detection to include dominance effects, given the prevalence of inbreeding depression. Furthermore, dense marker sets can increase the mapping resolution and allow the study of residual heterozygosity, unravelling further the genetic architecture of yield as reported by Marand et al. (2019).

The ability to map QTLs that segregate in a breeding population is crucial for the success of a breeding program as it can guide breeding decisions on germplasm enhancement, population improvement and choice of hybrid parents. QTL information can be used in the context of Marker-Assisted Selection for simple traits (e.g. for the introgression of resistance genes) or genomic prediction for traits that are regulated by a combination of major and minor genes (Bernardo 2014). In the case under study, the use of proper methodology can boost the improvement of the currently emerging diploid potato germplasm and consequently the development of diploid hybrid potato as described by Lindhout et al. (2011).

Acknowledgements This work is part of the research programme PredAPloid with project number 14520, which is financed by the Dutch Research Council (NWO). We would like to thank Solynta for making the data available to us.

Author Contributions GK performed the analysis and wrote the manuscript. MM developed SLA. CZ computed the IBD probabilities for GLA. MM, CM, HAM, CZ, RFV and FAVe supervised the study. All authors reviewed manuscript and discussed results.

Funding This work is part of the research programme PredAPloid with project number 14520, which is financed by the Dutch Research Council (NWO).

Availability of data and material Not applicable.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval Not applicable.

Code availability Not applicable.

Consent for publication Not applicable

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin-rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30(1):97–101
- Almasy L, Blangero J (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 62(5):1198–1211
- Bernardo R (2014) Genomewide selection when major genes are known. *Crop Sci* 54(1):68
- Bink M, Uimari P, Sillanpää M, Janss L, Jansen R (2002) Multiple QTL mapping in related plant populations via a pedigree-analysis approach. *Theor Appl Genet* 104(5):751–762
- Blanc G, Charcosset A, Mangin B, Gallais A, Moreau L (2006) Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. *Theor Appl Genet* 113(2):206–224
- 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. *Genetics* 9(7):2107–2122
- Bourke PM, Voorrips RE, Visser RG, Maliepaard C (2015) The double-reduction landscape in tetraploid potato as revealed by a high-density linkage map. *Genetics* 201(3):853–863
- Bourke PM, Voorrips RE, Visser RG, Maliepaard C (2018) Tools for genetic studies in experimental populations of polyploids. *Front Plant Sci* 9:513
- Bradshaw JE, Hackett CA, Pande B, Waugh R, Bryan GJ (2008) QTL mapping of yield, agronomic and quality traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). *Theor Appl Genet* 116(2):193–211
- Braun SR, Endelman JB, Haynes KG, Jansky SH (2017) Quantitative trait loci for resistance to common scab and cold-induced sweetening in diploid potato. *Plant Genome* 10(3):1–9
- Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19(7):889–890
- Buckler ES, Holland JB, Bradbury PJ, Acharya CB, Brown PJ, Browne C, Ersoz E, Flint-Garcia S, Garcia A, Glaubitz JC, Goodman MM, Harjes C, Guill K, Kroon DE, Larsson S, Lepak NK, Li H, Mitchell SE, Pressoir G, Peiffer JA, Rosas MO, Rocheford TR, Romay MC, Romero S, Salvo S, Villeda HS, Sofia da Silva H, Sun Q, Tian F, Upadaya N, Ware D, Yates H, Yu J, Zhang Z, Kresovich S, McMullen MD (2009) The genetic architecture of maize flowering time. *Science* 325(5941):714–718
- Butler D, Cullis B, Gilmour A, Gogel B (2009) ASReml-R reference manual (Version 3). Brisbane, The State of Queensland, Department of Primary Industries and Fisheries
- Crepeux S, Lebreton C, Servin B, Charmet G (2004) Quantitative trait loci (QTL) detection in multicross inbred designs. *Genetics* 168(3):1737–1749
- Endelman JB, Jansky SH (2016) Genetic mapping with an inbred line-derived F₂ population in potato. *Theor Appl Genet* 129(5):935–943
- Fernando R, Grossman M (1989) Marker assisted selection using best linear unbiased prediction. *Genetics Selection Evolution* 21(4):467
- Garin V, Wimmer V, Mezouk S, Malosetti M, van Eeuwijk F (2017) How do the type of QTL effect and the form of the residual term influence QTL detection in multi-parent populations? A case study in the maize EU-NAM population. *Theor Appl Genet* 130(8):1753–1764
- Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schüler K (2004) Assessing genetic potential in germplasm collections of crop plants by marker-trait association: A case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Mol Breed* 13(1):93–102
- Gebhardt C, Ritter E, Debener T, Schachtschabel U, Walkemeier B, Uhrig H, Salamini F (1989) RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theor Appl Genet* 78(1):65–75
- George AW, Visscher PM, Haley CS (2000) Mapping quantitative trait loci in complex pedigrees: a two-step variance component approach. *Genetics* 156(4):2081–2092
- Giraud H, Lehermeier C, Bauer E, Falque M, Segura V, Bauland C, Camisan C, Camp L, Meyer N, Ranc N, Schipprack W, Flament P, Melchinger AE, Menz M, Moreno-González J, Ouzunova M, Charcosset A, Schön CC, Moreau L (2014) Linkage disequilibrium with linkage analysis of multiline crosses reveals different multiallelic QTL for hybrid performance in the flint and dent heterotic groups of maize. *Genetics* 198(4):1717–1734
- 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(5):e63939
- Jannink JL, Jansen R (2001) Mapping epistatic quantitative trait loci with one-dimensional genome searches. *Genetics* 157(1):445–454

- Jansky SH, Charkowski AO, Douches DS, Gusmini G, Richael C, Bethke PC, Spooner DM, Novy RG, De Jong H, De Jong WS, Bamberg JB, Thompson AL, Bizimungu B, Holm DG, Brown CR, Haynes KG, Sathuvalli VR, Veilleux RE, Miller JC, Bradeen JM, Jiang J (2016) Reinventing potato as a diploid inbred line-based crop. *Crop Sci* 56(4):1412
- Kennedy BW, Quinton M, van Arendonk J (1992) Estimation of effects of single genes on quantitative traits. *J Anim Sci* 70(7):2000–2012
- Leroux D, Rahmani A, Jasson S, Ventelon M, Louis F, Moreau L, Mangin B (2014) Clusthaplo: a plug-in for MCQTL to enhance QTL detection using ancestral alleles in multi-cross design. *Theor Appl Genet* 127(4):921–933
- Li J, Ji L (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* 95(3):221–227
- Lindhout P, de Vries M, ter Maat M, Ying S, Viquez-Zamora M, van Heusden S (2018) Hybrid potato breeding for improved varieties. In: Wang-Pruski G (ed) Achieving sustainable cultivation of potatoes, vol 1. Breeding improved varieties. Burleigh Dodds Science Publishing, Cambridge, pp 99–122
- Lindhout P, Meijer D, Schotte T, Hutten RCB, Visser RGF, van Eck HJ (2011) Towards F₁ hybrid seed potato breeding. *Potato Res* 54(4):301–312
- Malosetti M, van Eeuwijk FA, Boer MP, Casas AM, Elía M, Moralejo M, Bhat PR, Ramsay L, Molina-Cano JL (2011) Gene and QTL detection in a three-way barley cross under selection by a mixed model with kinship information using SNPs. *Theor Appl Genet* 122(8):1605–1616
- Manrique-Carpintero NC, Coombs JJ, Cui Y, Veilleux RE, Buell CR, Douches D (2015) Genetic map and QTL analysis of agronomic traits in a diploid potato population using single nucleotide polymorphism markers. *Crop Sci* 55(6):2566
- Marand AP, Jansky SH, Gage JL, Hamernik AJ, De Leon N, Jiang J (2019) Residual heterozygosity and epistatic interactions underlie the complex genetic architecture of yield in diploid potato. *Genetics* 212(1):317–332
- Massa AN, Manrique-Carpintero NC, Coombs JJ, Zarka DG, Boone AE, Kirk WW, Hackett CA, Bryan GJ, Douches DS (2015) Genetic linkage mapping of economically important traits in cultivated tetraploid potato (*Solanum tuberosum* L.). *G3 Genes Genom Genet* 5(11):2357–2364
- McCord PH, Sosinski BR, Haynes KG, Clough ME, Yencho GC (2011) Linkage mapping and QTL analysis of agronomic traits in tetraploid potato (*Solanum tuberosum* subsp. *tuberosum*). *Crop Sci* 51(2):771
- Meijer D, Viquez-Zamora M, van Eck HJ, Hutten RCB, Su Y, Rothengatter R, Visser RGF, Lindhout WH, van Heusden AW (2018) QTL mapping in diploid potato by using selfed progenies of the cross *S. tuberosum* × *S. chacoense*. *Euphytica* 214(7):121
- Meuwissen TH, Goddard ME (2001) Prediction of identity by descent probabilities from marker-haplotypes. *Genet Select Evol* 33(6):605
- Parisseaux B, Bernardo R (2004) In silico mapping of quantitative trait loci in maize. *Theor Appl Genet* 109(3):508–514
- Peterson BA, Holt SH, Laimbeer FPE, Doullis AG, Coombs J, Douches DS, Hardigan MA, Buell CR, Veilleux RE (2016) Self-fertility in a cultivated diploid potato population examined with the infinium 8303 potato single-nucleotide polymorphism array. *Plant Genome* 9(3):1–13
- Piepho HP, Möhring J, Melchinger AE, Büchse A (2008) BLUP for phenotypic selection in plant breeding and variety testing. *Euphytica* 161(1–2):209–228
- Rebai A, Goffinet B (1993) Power of tests for QTL detection using replicated progenies derived from a diallel cross. *Theor Appl Genet* 86(8):1014–1022
- Rincint R, Moreau L, Monod H, Kuhn E, Melchinger AE, Malvar RA, Moreno-Gonzalez J, Nicolas S, Madur D, Combes V, Dumas F, Altmann T, Brunel D, Ouzunova M, Flament P, Dubreuil P, Charcosset A, Mary-Huard T (2014) Recovering power in association mapping panels with variable levels of linkage disequilibrium. *Genetics* 197(1):375–387
- Schäfer-Pregl R, Ritter E, Concilio L, Hesselbach J, Lovatti L, Walkemeier B, Thelen H, Salamini F, Gebhardt C (1998) Analysis of quantitative trait loci (QTLs) and quantitative trait alleles (QTAs) for potato tuber yield and starch content. *Theor Appl Genet* 97(5–6):834–846
- Simko I, Costanzo S, Haynes KG, Christ BJ, Jones RW (2004) Linkage disequilibrium mapping of a *Vorticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. *Theor Appl Genet* 108(2):217–224
- Sneller CH, Mather DE, Crepieux S (2009) Analytical approaches and population types for finding and utilizing QTL in complex plant populations. *Crop Sci* 49(2):363
- ter Braak CJF, Boer MP, Totir LR, Winkler CR, Smith OS, Bink MCAM (2010) Identity-by-descent matrix decomposition using latent ancestral allele models. *Genetics* 185(3):1045–1057
- Thompson EA (2000) Statistical inference from genetic data on pedigrees. In: NSF-CBMS Regional conference series in probability and statistics, Vol. 6. Institute of Mathematical Statistics
- van Eeuwijk FA, Boer M, Totir LR, Bink M, Wright D, Winkler CR, Podlich D, Boldman K, Baumgarten A, Smalley M, Arbelbide M, ter Braak CJF, Cooper M (2010) Mixed model approaches for the identification of QTLs within a maize hybrid breeding program. *Theor Appl Genet* 120(2):429–440
- van Ooijen J (2009) MapQTL 6, Software for the mapping of quantitative trait loci in experimental populations of diploid species. Kyazma BV., Wageningen
- VanRaden PM (2008) Efficient methods to compute genomic predictions. *J Dairy Sci* 91(11):4414–4423
- Verbyla AP, George AW, Cavanagh CR, Verbyla KL (2014) Whole-genome QTL analysis for MAGIC. *Theor Appl Genet* 127(8):1753–1770
- Wei J, Xu S (2016) A random-model approach to QTL mapping in multiparent advanced generation intercross (MAGIC) populations. *Genetics* 202(2):471–486
- Wimmer V, Albrecht T, Auinger HJ, Schön CC (2012) Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* 28(15):2086–2087
- Würschum T (2012) Mapping QTL for agronomic traits in breeding populations. *Theor Appl Genet* 125(2):201–210
- Xie C, Gessler DDG, Xu S (1998) Combining different line crosses for mapping quantitative trait loci using the

- identical by descent-based variance component method. *Genetics* 149(2):1139–1146
- Xu X, Atchley WR (1995) A random model approach to interval mapping of quantitative trait loci. *Genetics* 141(3):1189–1197
- Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL (2014) Advantages and pitfalls in the application of mixed-model association methods. *Nat Genet* 46(2):100–106
- Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. *Genetics* 178(1):539–551
- Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, Doebley JF, McMullen MD, Gaut BS, Nielsen DM, Holland JB, Kresovich S, Buckler ES (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38(2):203–208
- Zhang C, Wang P, Tang D, Yang Z, Lu F, Qi J, Tawari NR, Shang Y, Li C, Huang S (2019) The genetic basis of inbreeding depression in potato. *Nat Genet* 51(3):374
- Zheng C, Boer MP, van Eeuwijk FA (2014) A general modeling framework for genome ancestral origins in multiparental populations. *Genetics* 198(1):87–101
- Zheng C, Boer MP, van Eeuwijk FA (2015) Reconstruction of genome ancestry blocks in multiparental populations. *Genetics* 200(4):1073–1087
- Zheng C, Boer MP (2018) Recursive algorithms for modeling genomic ancestral origins in a fixed pedigree. *G3 Genes Genom Genet* 8(10):3231–3245

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.