Theor Appl Genet (2014) 127:731–748
DOI 10.1007/s00122-013-2254-y

ORIGINAL PAPER

Identification of agronomically important QTL in tetraploid potato cultivars using a marker–trait association analysis

Björn B. D'hoop · Paul L. C. Keizer · M. João Paulo · Richard G. F. Visser · Fred A. van Eeuwijk · Herman J. van Eck

Received: 29 July 2013 / Accepted: 13 December 2013 / Published online: 10 January 2014

Abstract
Key message **Nineteen tuber quality traits in potato were phenotyped in 205 cultivars and 299 breeder clones. Association analysis using 3364 AFLP loci and 653 SSR-alleles identified QTL for these traits.**
Abstract Two association mapping panels were analysed for marker–trait associations to identify quantitative trait loci (QTL). The first panel comprised 205 historical and contemporary tetraploid potato cultivars that were phenotyped in field trials at two locations with two replicates (the academic panel). The second panel consisted of 299 potato cultivars and included recent breeds obtained from five Dutch potato breeding companies and reference cultivars (the industrial panel). Phenotypic data for the second panel were collected during subsequent clonal selection generations at the individual breeding companies. QTL were identified for 19 agro-morphological and quality traits. Two association mapping models were used: a baseline model without, and a more advanced model with correction for population structure and genetic relatedness. Correction for population structure and genetic relatedness was performed with a kinship matrix estimated from marker information. The detected QTL partly not only confirmed previous studies, e.g. for tuber shape and frying colour, but also new QTL were found like for after baking darkening and enzymatic browning. Pleiotropic effects could be discerned for several QTL.

Communicated by C. Gebhardt.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-013-2254-y) contains supplementary material, which is available to authorized users.

B. B. D'hoop · M. J. Paulo · R. G. F. Visser · H. J. van Eck (✉)
Plant Breeding, Wageningen University and Research Centre, P.O. Box 386, 6700 AJ Wageningen, The Netherlands
e-mail: herman.vaneck@wur.nl

B. B. D'hoop · R. G. F. Visser · H. J. van Eck
Graduate School of Experimental Plant Sciences, Wageningen, The Netherlands

B. B. D'hoop · M. J. Paulo · R. G. F. Visser · F. A. van Eeuwijk · H. J. van Eck
Centre for BioSystems Genomics (CBSG), Wageningen, The Netherlands

P. L. C. Keizer · M. J. Paulo · F. A. van Eeuwijk
Biometris, Wageningen University and Research Centre, Wageningen, The Netherlands

Introduction

Potato (*Solanum tuberosum* L.) is the most important non-cereal staple crop worldwide and is increasingly being cultivated in Asia and developing countries, which traditionally rely on rice, pulses or other root and tuber crops like cassava and sweet potato. Unlike many crops that display an annual genetic gain of approximately one percent yield increase per year, the genetic gain in potato is almost absent. Its current breeding system, characterized by tetraploidy ($2n = 4x = 48$) and intolerance to inbreeding (Bradshaw et al. 2008), does not allow fixation of superior allele combinations in homozygous condition. In effect, potato breeders may capture beneficial allele combinations within a cultivar by chance, but individual beneficial alleles cannot be fixed in tetraploid outbreeding germplasm unless markers are used in "precision breeding" strategies (Li et al. 2008). An initiative to overcome these limitations and to develop true seed F1 hybrids was launched recently (Lindhout et al. 2011).

Recent technological developments in molecular genetics have led to numerous cost-effective methods for the generation of large marker datasets, enabling construction of marker dense genetic linkage maps (van Os et al. 2006) and facilitating genome-wide association studies (GWAS) (D'hoop et al. 2008). GWAS are considered a valuable alternative to bi-parental crosses for the identification of quantitative trait loci (QTL) for three reasons. Association mapping using a panel of commercial cultivars increases the opportunity to detect superior alleles of relevant QTL within the breeders' gene pool. Second, GWAS has potentially a higher mapping resolution, and third manages the diversity of the tetraploid gene pool efficiently (Flint-Garcia et al. 2003; Jannink and Walsh 2002; Nordborg and Tavare 2002).

Successful pioneering association studies have been performed not only in human genetics, but also in selfing plant species like Arabidopsis (Nordborg et al. 2002; Aranzana et al. 2005; Zhao et al. 2007a, b), maize (Remington et al. 2001; Thornsberry et al. 2001; Yu et al. 2006) and rice (Lu et al. 2005; Mather et al. 2007). Currently, association studies are shifting away from model organisms towards crop species (Zhu et al. 2008). Association mapping identified QTL in various crop species, e.g. barley (Kraakman et al. 2004; Caldwell et al. 2006; Cockram et al. 2008), Brassica rapa (Zhao et al. 2007a, b), soybean (Wang et al. 2008), sugar beet (Kraft et al. 2000; Stich et al. 2008a, 2008c), sugarcane (Raboin et al. 2008) and wheat (Breseghello and Sorrells 2006a, b; Maccaferri et al. 2005; Stich et al. 2008b). The achieved successes clearly indicate that association mapping has become an important tool for QTL detection and that the application of associated markers in marker assisted selection is underway (Ersoz et al. 2007).

Association mapping has downsides as well. Linkage disequilibrium (LD) is affected by population structure due to drift, admixture and selection, the level of recombination, and allele frequencies. Consequently, association mapping results should be interpreted with care (Flint-Garcia et al. 2003; Gupta et al. 2005). Correction for population structure is essential to control the false discovery rate, but at the cost of losing true but confounded marker–trait associations (Jannink and Walsh 2002; Zhu et al. 2008; Breseghello and Sorrells 2006a, b; Mackay and Powell 2007). Population structure is mostly tackled by inspecting the association panel with STRUCTURE (Pritchard et al. 2000), a Bayesian approach that assigns group membership probabilities to genotypes using molecular marker information. Alternatively, multivariate approaches may be adopted to calculate genetic distances from marker data or pedigree information. Genetic distances can subsequently be used for cluster analysis (Kraakman et al. 2004; Simko et al. 2004a) or can directly impose structure on the variance–covariance matrix between genotypes in a mixed model environment (Parisseaux and Bernardo 2004; Malosetti et al. 2007). Results of Bayesian and multivariate approaches can also be used jointly in one model to correct for population structure as shown by Yu et al. (2006). Phenotypic analysis can precede or can be integrated with association mapping. Where phenotypic analysis precedes mapping, first genotypic (main) effects are estimated, which then are considered to be the response variable in association mapping (Stich et al. 2008b).

Studies on the inheritance of morphological and agro-ecological traits have, for obvious reasons, been performed mainly on diploid experimental material of limited agronomical value. QTL studies in diploid mapping populations aimed to improve pathogen resistance (e.g. Collins et al. 1999; Visker et al. 2003) or quality traits (e.g. Douches and Freyre 1994; Menendez et al. 2002; Werij et al. 2007). Only a few mapping and QTL studies have been reported using tetraploids (Bradshaw et al. 2004, 2008; Khu et al. 2008). Initial candidate gene association studies in tetraploid potato have been published recently (Li et al. 2008; Simko et al. 2004a, b; Malosetti et al. 2007; Gebhardt et al. 2004; Li et al. 2005). D'hoop et al. (2008) presented a proof-of-principle for GWAS in tetraploid potato, whereas this study reports on a comprehensive genome-wide association study in potato.

This study includes a larger number of genotypes and markers, including SSRs, and additional traits as compared with D'hoop et al. (2008). We compare two association panels using a large number of randomly generated markers. Population structure was corrected for by structuring the variance–covariance matrix of the random genotypic effects in a mixed model framework, using a relatedness matrix based on marker correlations. We detected new marker–trait associations and confirmed previously known associations for a variety of agro-morphological and quality traits.

Materials and methods

Plant material

A set of 430 tetraploid cultivars was collected, representative of commercial potato germplasm. Details on the composition of this set and selection criteria can be found in D'hoop et al. (2008). Tuber material was provided by five Dutch breeding companies and several genebanks (see Acknowledgments). Leaf material was harvested from greenhouse-grown and in vitro-grown genotypes, frozen with liquid nitrogen and stored at −80 °C until DNA extraction.

Phenotypic observations

Phenotypic data were collected for two partly overlapping association panels. The first panel contained 205 freely available historical and contemporary potato cultivars. Phenotypic values for this "academic panel" were obtained from a replicated two-location field trial in 2006. These trials resulted in a balanced dataset with few missing values. The phenotypic data of the second panel, referred to as the "industrial panel" integrated phenotypic data on 299 genotypes collected during company-specific clonal selection programs. For certain traits, observations made during clonal selection trials resulted in a somewhat unbalanced dataset, but other traits were phenotyped across many locations for many years. The 299 genotypes of the industrial panel included 190 recent breeds (five companies*38) and 109 standard cultivars. The overlapping 101 cultivars between the two panels illustrate that cultivars from the academic panel frequently served as standards in clonal selection trials. Further details about the characteristics of both panels can be found in D'hoop et al. (2011).

Table 1 For 19 quality traits heritabilities were estimated

Trait	Trait abbreviation	Academic panel	Industrial panel
After baking darkening[a]	ABD_Nov8c	0.68	0.48
	ABD_Feb8c	0.77	0.33
	ABD_Apr8c	0.71	0.38
	ABD_Apr4c	0.53	0.45
	ABD_24hNov8c	0.58	0.68
After cooking darkening[b]	ACD_1 h	0.88	0.52
	ACD_24 h	0.69	0.51
Flesh colour after cooking	CookColY	0.95	0.65
Cooking type	Cooking Type	0.85	0.56
Enzymatic browning	EnzBrow_30 min	0.8	0.58
Flesh colour (fresh)	FleshColY	0.92	0.63
Frying colour[a]	FryingCol_Nov8c	0.75	0.72
	FryingCol_Feb8c	0.88	0.78
	FryingCol_Apr8c	0.8	0.7
	FryingCol_Apr4c	0.74	0.68
Maturity	Maturity	0.85	0.85
Tuber shape	Shape	0.9	0.74
Tuber size	Size	0.79	0.52
Tuber specific gravity	Underwater weight	0.76	0.87

High heritabilities could be obtained during a single growing season in the academic panel, but multi-year-multi-location observations in the industrial panel showed lower heritabilities

[a] Trait abbreviation is followed by indication of the storage duration of samples after harvesting, where November, February and April indicated the month in which samples were analysed. Samples were kept at storage temperatures of 4 or 8 °C

[b] Non-enzymatic discolouration was evaluated 1 or 24 h after cooking

Trait values for 19 phenotypic characters were recorded following standard protocols of breeding companies, with emphasis on processing quality traits. An overview of trait names and/or abbreviations is presented in Table 1. Assay protocols are described in D'hoop et al. (2011), but brief trait descriptions are as follows: non-enzymatic discolouration of French fries or cooked tubers, through oxidation of iron-chlorogenic acid (Wang-Pruski and Nowak 2004) was observed after baking (ABD) and after cooking (ACD). Cold sweetening caused by the accumulation of reducing sugars was indirectly scored using the Maillard discolouration of French fries (FryingCol). Sample material for ABD and FryingCol phenotypes was subjected to varying storage regimes. Enzymatic discolouration caused by oxidation of phenols by *polyphenol oxidase* (Lærke et al. 2002), was recorded 30 min after grating tubers. The intensity of yellow pigmentation of tuber flesh was rated from light yellow 6 to orange 9. We only considered yellow pigmentation intensity to study the genetic control of tuber flesh yellowness in raw and cooked tuber tissue (FleshColY and CookColY). We excluded the white–yellow contrast, controlled by the well-known *Y*-locus on chromosome 3 (Bonierbale et al. 1988). Cooking Type, Maturity, Shape, Size and starch content (Underwater Weight) were recorded as well.

Phenotypic analysis of both association panels using mixed models resulted in the estimation of genotypic main effects that formed the basis for the association analyses in this paper, heritabilities and genetic correlations. Further details on the methodology used for these calculations and the obtained results are presented in D'hoop et al. (2011).

Molecular marker analysis

DNA extraction was according to van der Beek et al. (1992). DNA quality and concentration were visually examined using ethidiumbromide stained 1 % agarose gels. AFLP™ markers were generated according to Vos et al. (1995) using 26 EcoRI/MseI and 15 PstI/MseI primer combinations. Specifications about primer combinations, marker generation and processing can be found in D'hoop et al. (2010). With 53 primer pairs microsatellite peak patterns were generated and allele dosage data were pursued (0, 1, 2, 3, 4) to boost marker use efficiency. SSR specifications, the protocol used and the methodology applied to convert peak patterns into zygosity are given in D'hoop et al. (2010). Positions were retrieved from the ultra-dense potato map using the parental diploid genotypes

SH83-92-488 and RH89-039-16 as internal references (van Os et al. 2006). For the SSRs also literature was consulted for map positions.

Association analysis

A detailed report on population structure is presented in D'hoop et al. (2010). Various methods for identifying population structure roughly resulted in the same six population groups for our association panels, thereby indicating the necessity to account for genetic relationships in association analyses, to preclude false positives. We estimated a marker based genetic relatedness matrix that could effectively structure the variance–covariance matrix of the random genotypic effects, thereby following not only the approach of Malosetti et al. (2007), but also those of Kang et al. (2008) and Zhao et al. (2007a, b).

GenStat, 11th edition (VSN International Ltd., Oxford, UK) was used as software environment. We opted for a two-step association approach. First, random genotypic main effects, i.e. Best Linear Unbiased Predictors (BLUPs), were computed for both panels. Because the trait heritabilities of the two locations looked similar and because the phenotypic dataset was balanced, the academic panel allowed investigating both the average, consistent expression of traits and the differential expression, or the genotype-by-environment interaction (GEI). Taking the average of the genotypic main effects across both locations we studied consistent QTL expression, whereas the difference between the genotypic main effects across the two locations served to detect loci influenced by GEI. See D'hoop et al. (2011) for further details about the phenotypic analysis. Second, association analysis was performed with the BLUPs as entry values.

Three marker sets were designed: (1) a "mapped" set of 720 AFLPs, (2) a "comprehensive" set of 3364 AFLPs including AFLP markers without known map location, and (3) a set of 53 SSRs. We used normalised log-transformed band intensities for AFLPs and allele dosages for SSRs. Two association models were constructed and applied to each association panel. A "baseline" model: Response = Marker + error, and an "advanced" model where we do correct for population structure: Response = Marker + Genotype + error. In both models the random term Response refers to the set of BLUPs and the fixed term Marker to AFLP or SSR information. The random term Genotype is introduced in the advanced model to enable the variance–covariance matrix for the genotypic effects to be structured by a genetic relatedness matrix. This matrix was calculated for each association panel, specific for the included genotypes, using correlations between normalised log-transformed AFLP band intensities. We used the Wald test to assess significance of the marker–trait associations (Verbeke and Molenberghs 2000). In the results section, we only considered the associations that were obtained with a $-10\log P > 3$.

It is common practice in potato breeding to physically separate cultivars in field trials into three maturity classes: early, main crop and late. In D'hoop et al. (2011) it was demonstrated that the estimated BLUPs for the traits maturity and underwater weight were influenced by the correction for maturity class during phenotypic analysis. Because we used these BLUPs as input for association analysis, we foresaw a carry-over to the association results. To evaluate this possibility, we also used BLUPs obtained with a model without maturity class correction, as described in D'hoop et al. (2011).

Results

Phenotypic analysis

The academic panel contained 15,580 phenotypic observations regarding 19 agro-morphological and quality traits on 205 cultivars. These data were collected in a single year in a balanced replicated field trial on two soil types. For the same 19 traits 73,968 observations were collected by commercial breeders on 299 cultivars that belonged to the industrial panel. These data were collected across many years and locations, but were highly unbalanced. Breeders differentiate between market niches and, therefore, genotypic effects in the industrial panel were for some traits based on strongly different numbers of trials. For example starch cultivars were not evaluated for cooking properties or processing quality.

Our study takes full advantage of the clonal reproduction of potato genotypes, which allows replicated trait observations during a single growing season on different soil types, as well as observations across years in the academic and industrial panel, respectively. As a consequence, the environmental noise can be controlled and allowed to obtain strikingly high heritabilities using the academic panel, ranging from 53 to 95 % (Table 1). The industrial panel yielded lower heritabilities 33 to 87 %.

Molecular marker analysis

The potato genotypes of both association panels have been analysed with 41 AFLP primer combinations which resulted in 3364 AFLP fragments of which signal intensities were scored (the "comprehensive" set, see M&M). Variation in AFLP signal intensity can reflect zygosity polymorphisms with allele dosages ranging from Aaaa (simplex) to AAAA (quadruplex) (D'hoop et al. 2008). Genetic positions on the ultra-dense AFLP map of potato

(van Os et al. 2006) could be assigned to 720 out of these 3364 AFLP markers (the "mapped" set, see M&M). The remaining 2,644 fragments, only polymorphic in the genotypes of the association panels, but not in our mapping population, suggest that the number of haplotypes in the tetraploid potato germplasm is two- to threefold larger than the four alleles from the two diploid parents of this linkage map. Indeed resequencing of candidate genes in these genotypes showed the presence of 9 to 11 alleles per locus (Wolters et al. 2010). Similarly, SSR analysis of the potato genotypes led to the detection of 653 alleles from 53 SSR markers.

Association analysis

We restrict the main story line to the findings with the academic panel and the AFLP markers. The association results obtained with the industrial panel, the SSRs and the findings related to the impact of maturity class correction on association results are addressed in the discussion section.

Using the mapped AFLPs, the baseline model (see M&M) detected 161 associations for the academic panel. The advanced model (see M&M) reduced this number to 43 (Table 2). Figure 1 displays the associated loci detected with the advanced model along the potato genome. Tables S1 and S2 present details such as the $-\log10p$ value, the map position and the marker name of detected associations with the mapped set, respectively, for the academic and industrial panel. With the comprehensive marker set, the baseline model resulted in 701 marker–trait associations for the academic panel. The advanced model scaled this number down to 157 (Table 3). Tables S3 and S4 provide detailed information like $-\log10p$ value and marker name on found associations with the comprehensive set, for the academic and industrial panel, respectively.

Table 3 illustrates that with a more stringent significance threshold, the number of associations drops accordingly and that population structure correction reduces association numbers in a consistent manner, irrespective of significance threshold.

Correlated traits

Correlations between traits were calculated for both association panels (D'hoop et al. 2011). In short, across traits ABD correlates with ACD. Plant maturity correlates with underwater weight, and both correlate with Cooking Type and ABD and ACD. Within traits the various assays at different time points will correlate. These correlations are presented in Table 4. Due to trait by trait correlations, we anticipate to observe the same markers involved in marker–trait associations, suggesting shared QTL. Table 2 already lists associated loci between some traits within and across association panels, suggesting trait relationships. Table 4 explicitly quantifies the shared marker–trait associations for the comprehensive AFLP set for trait sets that appeared correlated.

Physiologically correlated traits like maturity type, underwater weight and cooking type were expected to show a considerable number of shared QTL. Surprisingly, with the advanced model no shared associations could be discerned (Tables 2, 3, and 4). But, when inspecting the results obtained with the baseline model, almost all marker–trait associations detected for maturity and cooking type were found shared with underwater weight (Table 4, Tables S1 to S4).

Traits with a similar underlying process appeared correlated as well. Flesh colour and cooking colour, both visually measuring carotenoid contents in tuber tissue, shared a QTL on chromosome 1 (8.8–9.2 cM) and on chromosome 3 at 31.5 cM (Table 2). A higher number of shared associations appeared with the baseline model (Table 4, Tables S1 to S4).

Other phenotypes were correlated because they represent repeated measurements of the same trait following different storage regimes. As shown in Table 2, the frying colour traits measured in April shared a QTL on chromosome 9 (15.8–16.6 cM), and those ones measured after storage at 8 °C shared a locus on chromosome 2 (3.7 cM). ABD traits were associated with the same QTL on chromosome 3 (31.5 cM) and 4 (27.5 cM), but no QTL was shared by all different ABD measurements. The detected QTL for after cooking darkening (ACD) on chromosome 3 at 56.5 cM and chromosome 11 at 35.5 cM could be linked with QTL for after baking darkening (ABD) on chromosome 3 at 31.5 cM and chromosome 11 at 44.6 cM. The ABD and ACD QTL mapped close by on chromosome 3, although not within the predicted LD decay range (D'hoop et al. 2010). Correction for population structure did not influence the identification of common QTL for processing quality traits. But trait by trait relationships are more obvious when the results of the baseline model are examined (Table 4, Tables S1 to S4).

Genotype-by-environment QTL

We identified genotype-by-environment interactions (GEI) for the academic panel using the difference between trait values as assessed under clay and sand conditions. In total, 26 GEI QTL were detected with the baseline model using the mapped AFLP set (Table 5). Population structure correction reduced the number of associated QTL to 21, a similar trend as observed for the main effect QTL (Table 5).

The number of GEI QTL was far lower than the number of main effect QTL (26 versus 161, Table 5). In four cases the QTL associated with a GEI were identical or equivalent

Table 2 The number of marker—trait associations obtained for the academic and industrial panel with the mapped marker set (19 traits and 720 marker loci)

Trait	Number of associated markers ($p \leq 0.001$) for the academic panel (main effect QTL only)			Number of associated loci ($p \leq 0.001$) for the industrial panel			Literature—diploid and tetraploid QTL mapping studies		Literature—association mapping studies		
	Baseline	Advanced	Chromosome (cM)	Baseline	Advanced	Chromosome (cM)	Chromosome (candidate gene)	Reference	Number of loci	Chromosome (cM)	Reference
ABD_Nov8c	6	1	4 (27.5)	23	1	10 (0)	–	–	10	1 (9.2, 24.9), 3, 4 (27.5, 30, 32.9), 6, 7 (55.3), 11	D'hoop et al. (2008)
ABD_Feb8c	5	1	3 (31.5)	7	0	–	–	–	10	1 (9.2, 24.9), 3, 4 (27.5, 30, 32.9), 6, 7 (55.3), 11	D'hoop et al. (2008)
ABD_Apr8c	10	5	3 (31.5), 4 (0), 9 (15.8), 11 (19.4, 44.6)	2	0	–	–	–	10	1 (9.2, 24.9), 3, 4 (27.5, 30, 32.9), 6, 7 (55.3), 11	D'hoop et al. (2008)
ABD_Apr4c	5	3	1 (67.2), 9 (59.6), 10 (65.6)	5	2	1 (24.9), 8 (6.9)	–	–	–	–	–
ABD_24hNov8c	5	1	4 (27.5)	10	2	1 (49.2), 6(37)	–	–	10	1 (9.2, 24.9), 3, 4 (27.5, 30, 32.9), 6, 7 (55.3), 11	D'hoop et al. (2008)
ACD_1 h	8	2	3 (56.5), 11 (35.5)	19	0	–	1 (CIS), 2, 4, 6, 7 (ACO), 11	Bradshaw et al. (2008), Chen et al. (2001), Wang-Pruski and Nowak (2004)	7	1 (0.8, 24.9), 2, 4 (22.8), 5, 7 (54.5), 10	D'hoop et al. (2008)
ACD_24 h	0	0	–	0	0	–	1 (CIS), 2, 4, 6, 7 (ACO), 11	Bradshaw et al. (2008), Chen et al. (2001), Wang-Pruski and Nowak (2004)	7	1 (0.8, 24.9), 2, 4 (22.8), 5, 7 (54.5), 10	D'hoop et al. (2008)
CookColY	20	3	1 (9.2), 3 (31.5), 8 (11.8)	8	0	–	–	–	–	–	–
Cooking Type	5	2	2 (2.3), 11 (35.5)	18	2	1 (9.2), 6 (22.9)	9 (StTLRP)	Kloosterman et al. (2010)	6	1, 2, 6, 10	D'hoop et al. (2008)
EnzBrow_30 min	4	3	5 (36.4, 37.2), 6 (10.7)	1	1	1 (56.5)	1 (ED-1), 3 (ED-3), 8 (ED-8a, ED-8b, PPO-III)	Bachem et al. (1994), Werij et al. (2007), Bonierbale et al. (1988), Jacobs et al. (1995), Sliwka et al. (2008)	6	1, 3, 4, 7, 8 (10.6), 11	D'hoop et al. (2008), Urbany et al. (2011)
FleshColY	12	5	1 (8.8), 5 (8.5), 6 (38.5), 9 (16.6), 12 (8.8)	39	3	3 (31.5), 6 (13.7, 38.5)	3 (*Y/y*-locus), 4, 10, 12	Bonierbale et al. (1988), Van Eck et al. (1994a)	6	1, 2, 3 (31.5), 4, 6, 7	D'hoop et al. (2008)

Table 2 continued

Trait	Number of associated markers ($p \leq 0.001$) for the academic panel (main effect QTL only)			Number of associated loci ($p \leq 0.001$) for the industrial panel			Literature—diploid and tetraploid QTL mapping studies		Literature—association mapping studies		
	Baseline	Advanced	Chromosome (cM)	Baseline	Advanced	Chromosome (cM)	Chromosome (candidate gene)	Reference	Number of loci	Chromosome (cM)	Reference
FryingCol_Nov8c	7	0	–	33	1	10 (57.6)	1 (Sug1b), 2 (Sug2a), 6, 7 (Sug7a,c,e, SPS), 11	Bradshaw et al. (2008), Menendez et al. (2002)	3	1 (4.6), 2 (3.7), 3, 7 (63), 8, 12	D'hoop et al. (2008), Li et al. (2013), Fischer et al. (2013)
FryingCol_Feb8c	10	2	2 (3.7, 8.4)	29	1	4 (0.0)	1 (Sug1b), 2 (Sug2a), 6, 7 (Sug7a,c,e, SPS), 11	Bradshaw et al. (2008), Menendez et al. (2002)	3	1 (4.6), 2 (3.7), 7 (63)	D'hoop et al. (2008)
FryingCol_Apr8c	9	7	1 (43.8), 2 (3.7), 4 (19.7), 5 (31.3), 6 (51.8), 9 (15.8, 16.6)	17	2	6 (35), 9 (15.8)	1 (Sug1b), 2 (Sug2a), 6, 7 (Sug7a,c,e, SPS), 11	Bradshaw et al. (2008), Menendez et al. (2002)	3	1 (4.6), 2 (3.7), 7 (63)	D'hoop et al. (2008)
FryingCol_Apr4c	5	1	9 (16.6)	13	5	1 (24.9), 6 (6.6), 7 (53.9), 9 (15.8), 12 (43.6)	1 (Sug1b), 2 (Sug2a), 6, 7 (Sug7a,c,e, SPS), 11	Bradshaw et al. (2008), Menendez et al. (2002)	–	–	–
Maturity	10	2	1 (74.9), 3 (44)	66	1	2 (7.5)	3, 5	Sliwka et al. (2008), Bradshaw et al. (2004), Visker et al. (2005)	11	1, 4, 5 (16.2), 6, 9, 10	D'hoop et al. (2008), Gebhardt et al. (2004), Kloosterman et al. (2013)
Shape	10	4	2 (3.7), 10 (0), 11 (3.8), 12 (39.8)	27	1	6 (51.8)	2, 5, 7, 10 (Ro/ro), 11, 12	Bradshaw et al. (2008), Sliwka et al. (2008), Sørensen et al. (2006)), van Eck et al. (1994b)	5	2, 4, 11	D'hoop et al. (2008)
Size	4	1	8 (75.5)	7	1	4 (15.9)	5	Bradshaw et al. (2008)	–	–	–
Underwater weight	26	0	–	90	1	2 (62.8)	1 (ts(s)), 2, 4 (ts(i)), 5 (ts(a)), 9 (ts(o)), …	Gebhardt et al. (2005)	6	1 (8.4, 9.2, 24.9), 4 (21.3), 5 (16.2), 9 (46)	D'hoop et al. (2008), Li et al. (2013), Fischer et al. (2013)

Map positions are provided for the remaining marker loci following population structure correction (advanced model). The last two columns allow comparing our loci with QTL described in literature. References presented on the right can be found at the bottom of this table

to QTL associated with main effects (Table 5). ABD had a GEI QTL on chromosome 3 at 29.5 cM ($-10\log p = 3.4$) and a main effect QTL at 31.5 cM ($-10\log p = 7.1$). Because the main effect QTL had a much higher $-10\log p$, the effect due to soil type seemed negligible. A second example for ABD was found on chromosome 11 at 44.6 cM where a main effect QTL ($-10\log p = 5.1$) and a GEI QTL ($-10\log p = 3.8$) reside. For flesh colour we found a locus on chromosome 9 at 16.6 cM with $-10\log p$ of 3.1 for the main effect versus 3.2 for the GEI QTL, indicating that this QTL should be explained only as a genotype-by-location interaction. Likewise, cooking colour had a main effect QTL on chromosome 8 (11.8 cM) in the proximity of a GEI QTL at 25 cM. For the ABD traits GEI QTL were expected because during phenotypic analysis, the estimated variance component for GEI was high in comparison with the genetic variance component (D'hoop et al. 2011). Similarly, the absence of GEI QTL for frying colour traits neatly matches with their observed variance components for GEI and main effect following phenotypic analysis (D'hoop et al. 2011).

Discussion

We conclude that phenotypic and marker data observed in 205 tetraploid potato cultivars (academic panel) are a valuable dataset which allowed the detection of highly significant marker–trait associations which should allow fixation of beneficial alleles in marker-assisted potato breeding. Trait by trait correlations anticipated through preceding phenotypic analysis and background knowledge were confirmed by shared marker–trait associations. Population structure correction had a severe diminishing effect on the number of detected associations. Most likely not all removed associations were false positives. Genotype-by-environment interactions were found leading to a more accurate evaluation of coinciding main effect QTL. Our obtained results appear robust and reliable because of the consistent detection of marker–trait associations across datasets (see discussion) and will likely have impact on current day breeding practices.

Industrial panel versus academic panel

One of the major advantages of GWAS compared with conventional QTL mapping using full-sib offspring is that it can be applied to panels with prior availability of phenotyping data. Most potato genotypes have been tested for many years on many locations during clonal selection and in international variety trials by commercial breeding companies. Therefore, it is interesting to compare the single-year academic panel with the vast amount of phenotypic data in the industrial panel. In Table 1 it is shown that almost without exception the heritabilities for the industrial panel dropped by ten to forty per cent. More locations and years of testing the industrial panel is expected to improve the predictions of genotypic trait values. However, there are more differences in variance analysis and experimental design between the academic and industrial panel (D'hoop et al. 2011) that may explain the lower heritabilities observed in the industrial panel. The academic panel included a more diverse sample of the potato gene pool, including heirloom and non-Dutch cultivars. Furthermore, the academic panel was planted in such a way that cultivars of similar maturity classes were in the same blocks, to avoid interactions (shading) between early and taller late cultivars. The wider diversity and the maturity correction are considered as the main causes of the higher heritabilities observed in the academic panel.

The industrial panel reflected similar trends in association numbers as observed with the academic panel. With the comprehensive set 2875 associations have been detected of which 437 associations had a known map location. Population structure correction scaled these numbers down to 150 and 24, respectively (Tables 2 and 3). Similar trait relationships appeared using the industrial panel, but when considering the advanced model they were somewhat less abundant (Table 4). This was partly attributable to the more severe effect of population structure correction within this panel.

Population structure correction had an equally lowering effect on the number of reported associations with SSR data as it had with AFLPs. For the academic panel numbers went down from 15 to 11, and for the industrial panel from 99 to 13 (Table S5). The SSR results also revealed trait relationships within each panel. A shared locus for ABD traits was found on chromosome 4 (STI012) with the academic panel, but this was not the case in the industrial panel. On chromosome 8 a shared locus for ABD and ACD traits was identified with the industrial panel (Table S5), but not in the academic panel. Differences between the detected QTL across both panels could be due to different genotype sets, and furthermore, the tubers for phenotyping were derived from different field trials and subjected to different storage regimes or treatments. Therefore, complete congruence between ABD and ACD QTL or panels is not expected.

Considerably more marker–trait associations were detected in the industrial panel using the baseline model, irrespective of the marker set used. However, this difference in the number of marker–trait associations between the panels vanished or even reverted upon correction for population structure (Tables 2, 3 and S5). These initially inflated numbers of associations in the industrial panel were most

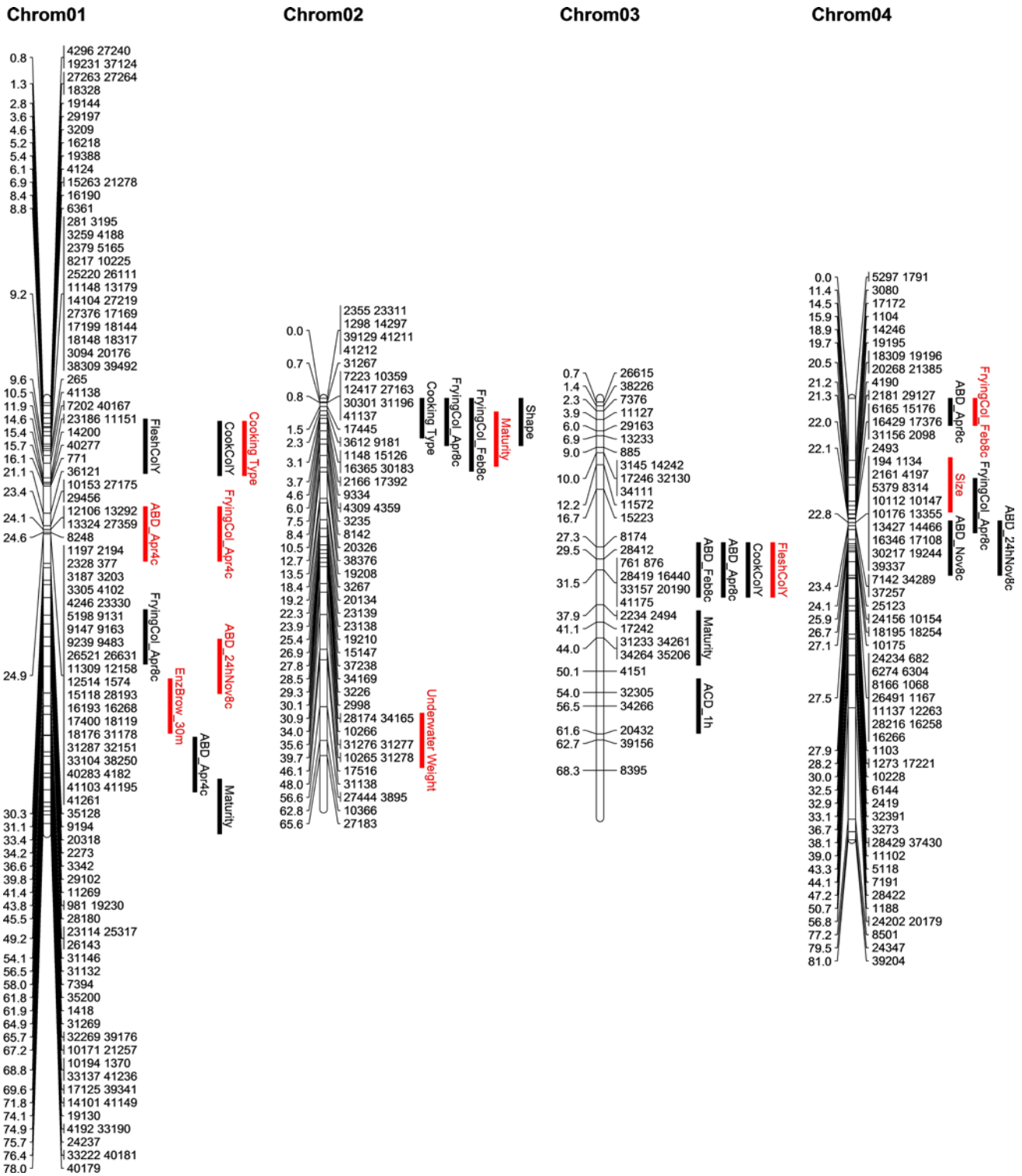

Fig. 1 Genome-wide distribution of QTL detected with the advanced model. Borders of QTL loci were indicated by adding or subtracting the estimated overall LD decay of 5 cM (D'hoop et al. 2010) to the associated marker positions. QTL in *black* were obtained with the academic panel, QTL in *red* with the industrial panel (colour figure online)

likely due to (1) closer relatedness of recent cultivars present in the industrial panel and (2) breeding for different market niches in the past decades caused a diversification in gene pools from breeding companies, reflecting additional phenotypic variation—at least for some traits—contained in the industrial panel (D'hoop et al. 2011).

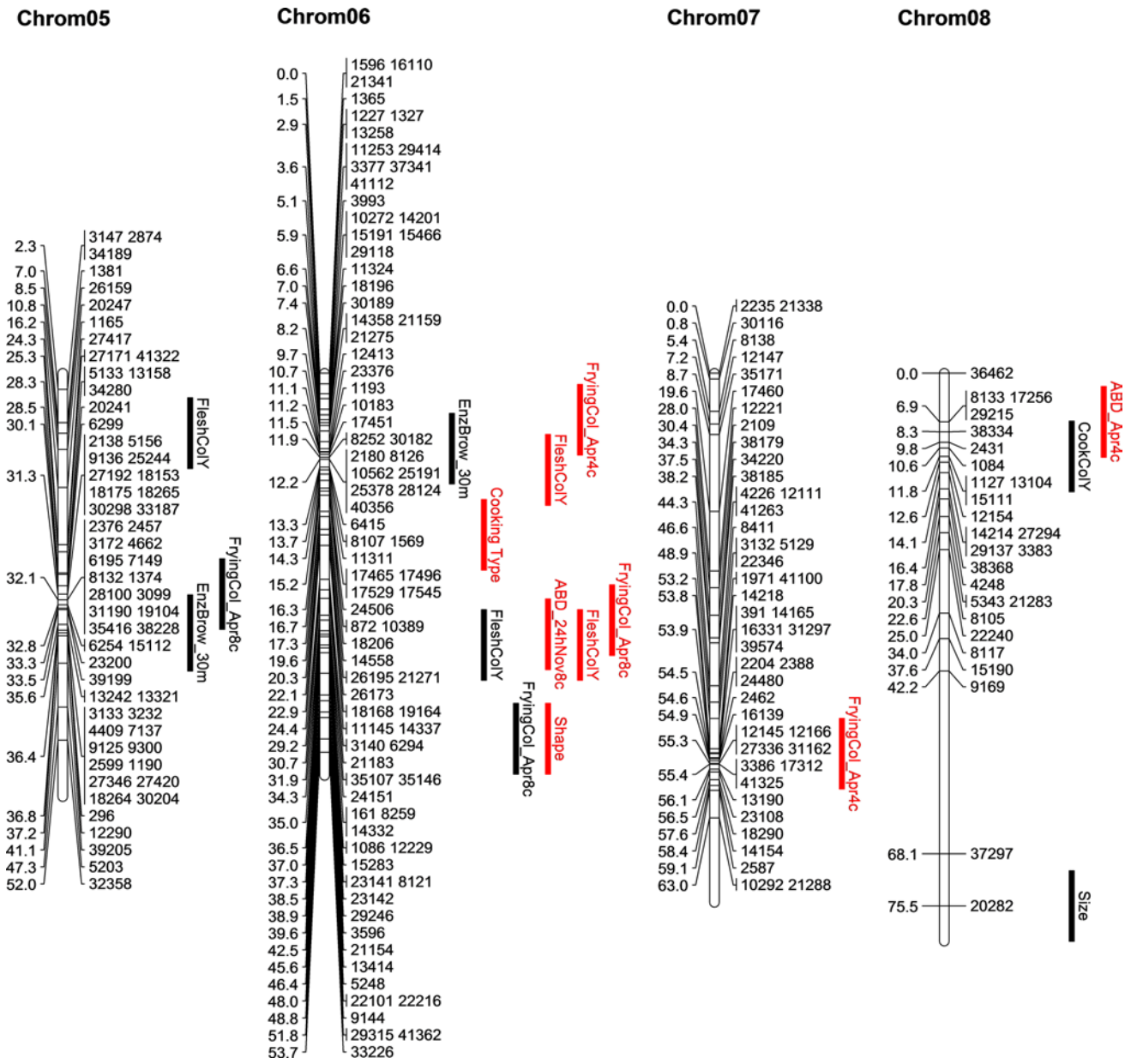

Fig. 1 continued

At first sight, only very few identically associated loci could be discerned between traits in both panels (Table 2). Only FleshColY with a locus on chromosome 6 at 38.5 cM and FryingCol_Apr8c and FryingCol_Apr4c (chromosome 9:15.8–16.6 cM) showed coinciding or closely coinciding loci. But this lack of correspondence was mainly artificial because we only listed mapped loci that were retained after correction for population structure. The results obtained with the baseline model and the comprehensive AFLP set showed much more consistency between both panels. The majority of the associations detected for traits with high heritabilities (Table 1) were found shared between both panels albeit with varying *p* values (Tables S3 and S4). The baseline results obtained with the mapped set confirmed this trend (Tables S1 and S2). The SSRs did not allow similar observations because very little associated loci were detected.

Population structure correction

Population structure correction had a severe deflating effect on numbers of associations (Tables 2, 3, 4, 5, 6, and 6, Tables S1 to S4). As expected, the most significant associations found with the baseline model reappeared with the advanced

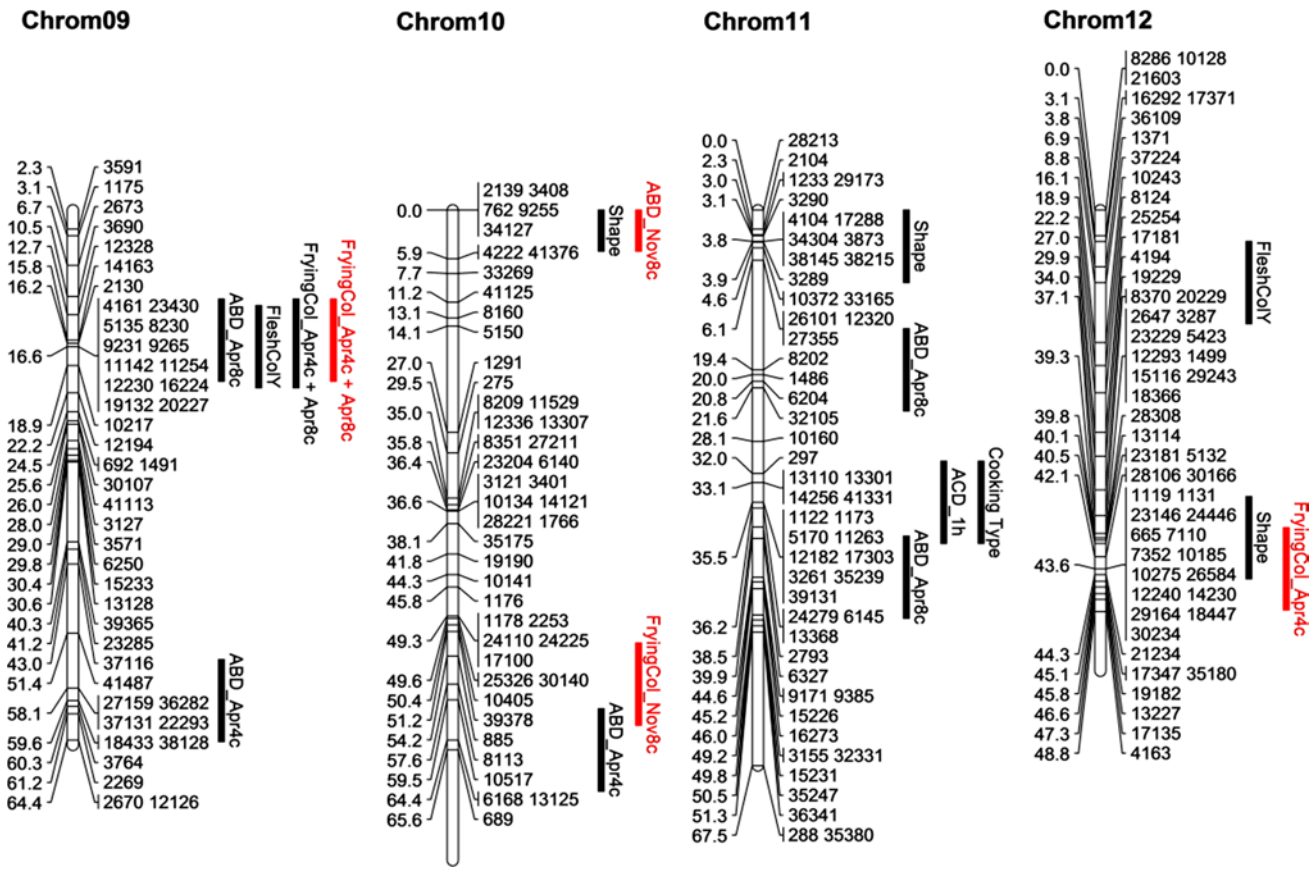

Fig. 1 continued

model, usually with a less significant p value (Tables S1 to S4). The marker–trait associations that were discarded or dropped below a predefined significance threshold, e.g. $p \leq 0.001$, when correcting for population structure are not necessarily meaningless. Consequently, QTL obtained with a baseline model as opposed to an advanced model require a careful inspection of a subset of associations with the highest explained variance and significance threshold.

In exceptional cases new associations can appear upon population structure correction. This was observed in 25 cases for the academic panel and in 23 cases for the industrial panel (Tables S3 and S4). Exactly 93.8 % of these "new" associations were detected with the baseline model at a less stringent significance threshold of $p \leq 0.05$. The remainder of the "new" associations were never observed as significant before and illustrate the need to inspect newly identified associations equally carefully.

The importance of maturity

Phenotypic variation for certain traits is often not independent. In Arabidopsis it has been observed that a large number of QTL co-localize on chromosome 2 (Alonso-Blanco et al. 1999) because of an overarching, pleiotropic effect of the erecta locus affecting many other phenotypes (van Zanten et al. 2009). The same effect is caused by dwarfing or short straw genes in many cereal crops (Thomas et al. 1995). In potato it is commonly known that plant maturity, localized on the short arm of chromosome 5, has an overarching effect on many phenotypes such as yield, starch yield, as measured by underwater weight, as well as cooking type which in turn is caused by tuber starch content. However, our results did not show an association accumulation on chromosome 5. The maturity locus on chromosome 5 did not even remain significant when the advanced model was applied (Table 2). There are two factors that play a role in this.

First, in potato field trials it is common practice to plant cultivars separated into blocks according to their maturity class. In D'hoop et al. (2011) it was shown that maturity class correction during phenotypic analysis influences the estimation of genotypic main effects and GEI effects of phenotypic traits. Because these effects were used as response values in the association models, we anticipated some influence on association results. With a model without maturity class correction the 19 phenotypes of the

Table 3 Quantification of associations detected with the comprehensive set

Trait	Number of loci ($p \leq 0.001$) academic panel main effects			Number of loci ($p \leq 0.0001$) academic panel main effects			Number of loci ($p \leq 0.00001$) academic panel main effects			Number of loci ($p \leq 0.001$) industrial panel			Number of loci ($p \leq 0.0001$) industrial panel			Number of loci ($p \leq 0.00001$) industrial panel		
	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced
ABD_Nov8c	24	3	2	3	0	0	0	0	0	245	8	7	107	3	3	47	1	1
ABD_Feb8c	29	5	5	8	2	2	2	1	1	60	10	10	19	1	1	6	0	0
ABD_Apr8c	26	11	7	6	2	2	2	1	1	20	8	8	3	0	0	0	0	0
ABD_Apr4c	7	6	5	1	2	2	1	1	1	23	8	5	6	0	0	1	0	0
ABD_24hNov8c	38	10	9	10	1	1	0	0	0	45	13	11	19	5	5	6	1	1
ACD_1 h	49	13	13	14	3	3	3	1	1	128	10	6	46	3	3	24	1	1
ACD_24 h	9	8	8	1	1	1	1	1	1	3	2	1	0	0	0	0	0	0
CookColY	66	14	10	21	1	1	8	0	0	38	3	3	11	0	0	2	0	0
Cooking type	38	12	11	13	1	1	5	0	0	103	7	7	46	2	2	17	0	0
EnzBrow_30 min	14	7	6	2	1	1	0	0	0	2	3	2	0	0	0	0	0	0
FleshColY	28	12	11	10	1	1	2	0	0	248	10	9	122	3	3	65	1	1
FryingCol_Nov8c	27	5	3	5	0	0	0	0	0	206	7	5	95	0	0	36	0	0
FryingCol_Feb8c	48	5	3	12	1	1	4	0	0	185	9	7	75	3	3	22	0	0
FryingCol_Apr8c	33	14	8	15	1	1	4	0	0	137	6	6	53	2	2	18	1	1
FryingCol_Apr4c	24	5	5	7	1	1	4	0	0	75	15	11	27	5	5	10	3	3
Maturity	30	7	6	7	3	3	2	0	0	504	9	9	314	3	3	201	1	1
Shape	48	12	11	13	3	3	7	0	0	198	8	8	93	2	2	45	0	0
Size	11	4	4	3	1	1	0	0	0	21	3	3	4	1	1	0	0	0
Underwater weight	152	4	4	72	1	1	32	0	0	634	11	9	428	1	1	279	0	0
Total	701	157	131	223	26	26	77	5	5	2,875	150	127	1,468	34	34	779	9	9

The left part describes the academic panel results for subsequent stringency levels for *p*-values, and the right part does the same for the industrial panel. The third column each time indicates how many of the loci found with the advanced model were previously significant with the baseline model

Table 4 Overview of shared marker–trait associations between different sets of traits, found with the comprehensive marker set

Traits	Number of shared marker–trait associations ($p \leq 0.001$) academic panel, only main effect			Genetic correlation between traits derived from phenotypic analysis	Number of shared marker–trait associations ($p \leq 0.001$) industrial panel			Genetic correlation between traits derived from phenotypic analysis
	Baseline	Advanced	Across models		Baseline	Advanced	Across models	
ABD traits	35	4	14	0.5	52	4	15	0.57
ACD traits	2	1	2	0.77	0	0	0	0.54
ABD and ACD traits	2	2	2	0.38	47	2	7	0.49
Frying colour traits	31	5	7	0.59	170	3	13	0.73
FleshColY and CookingColY	12	1	7	0.6	19	0	4	0.86
Maturity and underwater weight	14	0	0	−0.34	351	0	8	−0.36
Maturity and cooking type	4	0	0	−0.37	36	0	2	−0.38
Cooking type and underwater weight	16	1	6	0.52	80	2	9	0.83
Maturity, cooking type and underwater weight	1	0	0	0.41	36	0	2	0.52

Left side academic panel, right side industrial panel. Across models presents numbers of shared associations irrespective of the used model. The genetic correlations between trait sets, obtained during phenotypic analysis are provided as well

Table 5 Association results obtained with the academic panel and the mapped set with special attention to GEI QTL

Trait	Number of associated loci with $p \leq 0.001$ for the academic panel: main effect QTL			Number of associated loci with $p \leq 0.001$ for the academic panel: GEI QTL		
	Baseline	Advanced	Chromosome (cM)	Baseline	Advanced	Chromosome (cM)
ABD_Nov8c	6	1	4 (27.5)	0	0	–
ABD_Feb8c	5	1	3 (31.5)	1	1	3 (29.5)
ABD_Apr8c	10	5	3 (31.5), 4 (0), 9 (15.8), 11 (19.4, 44.6)	2	3	6 (15.2), 11 (44.6), 12 (43.6)
ABD_Apr4c	5	3	1 (67.2), 9 (59.6), 10 (65.6)	0	0	–
ABD_24hNov8c	5	1	4 (27.5)	1	1	10 (0)
ACD_1 h	8	2	3 (56.5), 11 (35.5)	1	1	12 (43.6)
ACD_24 h	0	0	–	0	0	–
CookColY	20	3	1 (9.2), 3 (31.5), 8 (11.8)	1	1	8 (25)
Cooking Type	5	2	2 (2.3), 11 (35.5)	1	1	10 (36.4)
EnzBrow_30 min	4	3	5 (36.4, 37.2), 6 (10.7)	1	1	1 (24.9)
FleshColY	12	5	1 (8.8), 5 (8.5), 6 (38.5), 9 (16.6), 12 (8.8)	5	2	9 (16.6), 10 (14.1)
FryingCol_Nov8c	7	0	–	1	1	6 (15.2)
FryingCol_Feb8c	10	2	2 (3.7, 8.4)	0	0	–
FryingCol_Apr8c	9	7	1 (43.8), 2 (3.7), 4 (19.7), 5 (31.3), 6 (51.8), 9 (15.8, 6.6)	0	0	–
FryingCol_Apr4c	5	1	9 (16.6)	0	0	–
Maturity	10	2	1 (74.9), 3 (44)	3	1	4 (22.8)
Shape	10	4	2 (3.7), 10 (0), 11 (3.8), 12 (39.8)	2	2	1 (24.9), 7 (55.3)
Size	4	1	8 (75.5)	4	4	2 (0.8), 10 (35, 59.5), 12 (3.1)
Underwater weight	26	0	–	3	2	7 (56.5), 12 (46.6)

On the left side the main effect QTL are reported, on the right the GEI QTL. Chromosomal positions are provided for associations found with the advanced model

Table 6 Effect of maturity class correction during phenotypic analysis on the number of detected associations for the academic panel, using the comprehensive set

Trait	Number of loci ($p \leq 0.001$) with maturity class correction			Number of loci ($p \leq 0.001$) without maturity class correction		
	Baseline	Advanced	Baseline persistent in advanced	Baseline	Advanced	Baseline persistent in advanced
ABD_Nov8c	24	3	2	27	5	1
ABD_Feb8c	29	5	5	41	6	4
ABD_Apr8c	26	11	7	34	5	3
ABD_Apr4c	7	6	5	3	9	1
ABD_24hNov8c	38	10	9	36	6	3
ACD_1 h	49	13	13	70	8	4
ACD_24 h	9	8	8	49	6	3
CookColY	66	14	10	71	19	3
Cooking Type	38	12	11	38	5	3
EnzBrow_30 min	14	7	6	23	8	3
FleshColY	28	12	11	27	2	1
FryingCol_Nov8c	27	5	3	31	7	2
FryingCol_Feb8c	48	5	3	74	8	4
FryingCol_Apr8c	33	14	8	48	5	3
FryingCol_Apr4c	24	5	5	31	4	1
Maturity	30	7	6	85	3	3
Shape	48	12	11	52	6	4
Size	11	4	4	36	4	0
Underwater weight	152	4	4	201	2	2
Total	701	157	131	977	108	48

academic panel were reanalysed and the resulting BLUPs used as input for association mapping. In Table 6 the effect of maturity class correction on the association results is quantified. The two most affected traits during phenotypic analysis (D'hoop et al. 2011), maturity and underwater weight, also appeared heavily affected in association analysis (Table 6). The number of detected associations with the baseline model increased for maturity from 30 to 85 and for underwater weight from 152 to 201. For most of the other traits an increasing effect in detected associations could be discerned as well, but not so strong and only effective for the baseline model. These findings demonstrate that we effectively removed maturity class variation from the estimated genetic variance components of nearly all phenotypes during phenotypic analysis (D'hoop et al. 2011), enabling the identification of marker–trait associations not coinciding with the maturity locus.

Maturity class correction removed the majority of the maturity type differences between cultivars during phenotypic analysis, but the baseline model still detected one marker–trait association for maturity on chromosome 5 for the academic panel (Table S1). Following population structure correction this association became insignificant. As such, population structure was the second factor leading to the absence of chromosome 5 QTL for maturity. Also with the industrial panel where no maturity class correction was applied during phenotypic analysis, population structure correction made all chromosome 5 associations for maturity non-significant (Table S2). In D'hoop et al. (2010) we showed that population structure mainly coincided with underwater weight differences between subgroups in our association panels and because an earlier maturity type is negatively correlated with a higher underwater weight (Table 4) also indirectly with maturity type differences.

Trait–trait correlations and pleiotropic effects

For the majority of traits for which shared marker–trait associations were observed a plausible explanation can be given. The genetic correlations between quality and agro-morphological traits estimated during phenotypic analysis seemed indeed high, i.e. well above $r = 0.3$ (Table 4), but these correlations can be explained from common underlying biological mechanisms and processes. Concerning quality traits, common marker–trait associations were predictable for the ABD traits and the frying colour traits. These two trait sets measured exactly the same discolouration type following exactly the same protocol; the only difference lay in

the storage specifications. Likewise, the shared associations between ABD and ACD traits were reasonable since both observed a similar discolouration process. Shared QTL for flesh colour and cooking colour were equally expected: both visually measured carotenoids in tuber tissue, but separated by one processing step. The shared loci between maturity and underwater weight or cooking type are due to a physiological correlation. Cultivars displaying late maturity can accumulate more starch and thus a higher underwater weight. Recently, the major effect QTL on chromosome 5 for plant maturity *StCDF* was cloned (Kloosterman et al. 2013), and it seems that associations reported here point to minor effect QTL, but are most likely false positives. This explains that in Table 4 no shared markers were identified between Matury and Underwater Weight or Cooking Type in the academic panel, whereas shared QTL in the industrial panel could be due to an insufficient correction for population structure in this highly structured material. Higher underwater weight correlates with a more sloughing cooking type. Both traits are under more polygenic control. Therefore, more marker associations are shared between these traits Table 4, which could lie in or in the vicinity of, or are in high LD with candidate genes that have pleiotropic effects on these traits.

ABD and ACD traits collectively point to a QTL on chromosome 3, where also flesh and cooking colour have marker–trait associations. This may point to a true QTL or to the inability to accurately phenotype these discolouration traits. It may be more difficult for the breeder's eye to detect discolouration on an intense yellow background than on a white background. Accurate phenotyping methodologies, like the colorimetric method for ACD measurement proposed by Wang-Pruski and Nowak (2004), would likely improve the signal-to-noise ratio for discolouration traits.

Genotype-by-environment QTL

Ten GEI QTL for agro-morphological traits like underwater weight, shape, size, maturity and cooking type were found with the advanced model (Table 5). Location effects seemed credible for agro-morphological traits because soil can influence a cultivar's performance. Also, the variance components for genotype-by-location interaction, estimated during phenotypic analysis, were not negligible for traits like size and underwater weight (D'hoop et al. 2011). Clay has a better water retaining capacity than sand which could lead to an environment favouring a later maturity class and thus a higher underwater weight and possibly a larger tuber size. Soil type can influence tuber shape: tubers are generally rounder in clay soil than in sandy soil. The effect of environment on cooking type can be seen as a result of its relationship with underwater weight. The six GEI QTL detected for quality traits (ACD, ABD, enzymatic browning and frying colour, see Table 5) were probably related to the exceptional summer during the growing season of 2006. Abundant rain towards the end of the growing season followed a period of drought which resulted in secondary growth. Surprisingly, three GEI QTL were detected for yellowness of tuber flesh and cooked tuber tissue (Table 5). Abundant rainfall on sandy soil may have eroded hills, exposing tubers to sunlight which caused tuber greening. This tuber greening effect may have led to biased observations on flesh colour traits on the sandy soil.

Previous studies confirmed

In Table 2 our results are compared with traditional QTL mapping studies in potato at the diploid and tetraploid levels and with previous association studies. Many more previously reported QTL were confirmed, but here we discuss only the QTL that remained after population structure correction. For the ABD loci a clear overlap was revealed. In D'hoop et al. (2008) chromosomes 3, 6 and 11 were equally found to be involved in ABD and for chromosome 4 there was a perfect match at 27.5 cM. Several associations for ABD were detected on chromosome 1, but they did not all reside in the vicinity of previously reported QTL, except for the locus at 24.9 cM obtained with the industrial panel. Possibly, due to LD which extends farther than 12 cM on chromosome 1 (D'hoop et al. 2010), the same haplotype may have been identified by the distantly located QTL on chromosome 1. A locus for ACD on chromosome 11 (between 5 and 19 cM) has been described by Bradshaw et al. (2008); perhaps the same haplotype was detected in our study with a QTL at 35.5 cM. Interestingly, our study detected a QTL on chromosome 11 (19.4 cM) for ABD. Since ABD and ACD measure similar discolouration processes we may have identified the same locus. The loci for cooking type on chromosomes 2 and 6 were in agreement with our previous work (D'hoop et al. 2008). Werij et al. (2007) combined candidate gene analysis with QTL and expression analysis and mapped enzymatic discolouration on chromosome 1 between 32.7 and 46.6 cM. We found a QTL on the same chromosome at 56.5 cM. With an independent association panel marker–trait associations for enzymatic browning were reported on chromosome 1 as well (D'hoop et al. 2008). Sliwka et al. (2008) identified a flesh colour locus on the top of chromosome 12, whereas the classica *Y*-locus encoding beta-carotene hydroxylase (Chy2) resides on chromosome 3 (Bonierbale et al. 1988). Wolters et al. (2010) nicely illustrated that Chy2 allele 3 is in fact the key determinant for flesh colour, also in tetraploid potato. Other flesh colour QTL were detected on chromosomes 1 and 6 (D'hoop et al. 2008). We obtained flesh colour QTL on chromosomes 1, 3, 6 and 12 as well. The combined loci for Maillard discolouration (Frying colour) on chromosomes 1, 2, 6 and 7 have all been reported

before (Bradshaw et al. 2008; D'hoop et al. 2008; Menendez et al. 2002), even with a perfect match on chromosome 2 at 3.7 cM. As in our study, QTL for maturity have been observed on chromosome 3 (Visker et al. 2005) and chromosome 1 (D'hoop et al. 2008). We detected tuber shape QTL on chromosomes 2 (3.7 cM), 10 (0 cM), 11 (3.8 cM) and 12 (39.8 cM), similar to what has been found by Bradshaw et al. (2008), Sliwka et al. (2008) and D'hoop et al. (2008) for chromosome 2 and 11, and by Sørensen et al. (2006) for chromosome 12. The Ro/ro locus on chromosome 10 was previously defined by van Eck et al. (1994a). Finally, the obtained QTL for underwater weight on chromosome 2 was observed by Gebhardt et al. (2005) as well.

Underwater weight

To illustrate the impact of our association mapping results, we discuss one agronomically important trait in terms of explained variance: underwater weight. Underwater weight is a key determinant for the market niche of a cultivar, plays a central role in many processing quality traits (Haase 2003) and appears significantly correlated with tuber yield (Maris 1969). Tables S3 and S4 show that in total 662 different marker–trait associations were found for underwater weight. Following rigorous filtering by allowing only for these associations that were identified with both association panels, this number dropped to 126. When map location pointed to two associated markers residing within 1 cM of the genetic map of potato (van Os et al. 2006), we considered them as redundant and subsequently dropped the least significant marker. This brought the total down to 121. Likely, a big part of these associations may still be redundant, because we do not know the map position of all markers. To further exclude redundant associations we performed a forward selection strategy (Montgomery et al. 2001) to obtain a final QTL set. Two QTL sets were obtained and they explained 42 and 63 % of the total phenotypic variation for the academic and industrial panel, respectively. The academic set contained ten QTL of which three were in common with the 14 QTL obtained for the industrial panel. We are planning to re-sequence these QTL regions so that allele-specific markers can be developed for use in marker-assisted selection. Allele-specific markers could predict underwater weight performance at seedling stage and may allow a substantial increase in cultivar performance for underwater weight.

Prospects

As this paper illustrates, genome-wide association mapping is a valuable tool to identify candidate trait loci for complex traits, also in a complex crop like potato. QTL that were repeatedly observed when analysing different association panels are surely candidate loci for further studies. If these loci would be confirmed with conventional linkage mapping, their reliability would increase, and it would qualify them as candidate regions for in-depth research. Extra confidence by linking expression or metabolite data to associations may be opportune prior to embarking into costly translational research, where one aims to translate identified QTL into easy accessible PCR markers, amenable to marker-assisted breeding (Kloosterman et al. 2010).

Now that the potato genome sequence has been released (http://www.potatogenome.net/images/2/2e/PGSC_Press_Release_0909.pdf), and more potato genotypes are planned to be sequenced in the near future (SOL-100, http://solgenomics.net/static_content/solanaceaeproject/docs/SOL_newsletter_Mar_10.pdf), full haplotype information on candidate regions will soon be available. This would allow an unambiguous assessment of diversity for particular regions in potato, next to detection of rare alleles that may have been missed. On the one hand this could lead to allele-specific sequence assisted breeding. On the other hand, genomic selection, which appeared extremely successful in animal breeding as reviewed recently by Goddard and Hayes (2009), may become feasible as well.

Acknowledgments We thank Agrico Research B. V. (Bant, Netherlands), Averis Seeds B. V. (Valthermond, Netherlands), C. Meijer B. V. (Rilland, Netherlands), HZPC Holland B. V. (Metslawier, Netherlands) and KWS Potato B. V. (Emmeloord, Netherlands) for providing tuber material, sharing phenotypic data and for their help in the academic trial. We acknowledge the genebanks Agriculture and Agri-Food (Canada), Arche Noah (Austria), IPK Gatersleben (Germany), INRA (France), SASA (Scotland), Teagasc (Ireland) and USDA-ARS (USA) for providing scarce heirloom potato cultivars. This study was financed as project P8 and PP2 by the Centre for BioSystems Genomics (CBSG), which is part of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research for financing this study. Fred van Eeuwijk and Paul Keizer worked wholly and João Paulo partially on this paper within the context of the CBSG projects BB9 and BB12.

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

References

Alonso-Blanco C, Blankestijn de Vries H, Hanhart CJ, Koornneef M (1999) Natural allelic variation at seed size loci in relation to other life history traits of *Arabidopsis thaliana*. Proc Natl Acad Sci USA 96:4710–4717

Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M et al (2005) Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLoS Genet 1:e60

Bachem CWB, Speckmann G-J, van der Linde PCG, Verheggen FTM, Hunt MD et al (1994) Antisense expression of polyphenol

oxidase genes inhibits enzymatic browning in potato tubers. Nat Biotechnol 12:1101–1105

Bonierbale MW, Plaisted RL, Tanksley SD (1988) RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120:1095–1103

Bradshaw JE, Pande B, Bryan GJ, Hackett CA, McLean K et al (2004) Interval mapping of quantitative trait loci for resistance to late blight [*Phytophthora infestans* (Mont.) de Bary], height and maturity in a tetraploid population of potato (*Solanum tuberosum* subsp *tuberosum*). Genetics 168:983–995

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 116:193–211

Breseghello F, Sorrells ME (2006a) Association analysis as a strategy for improvement of quantitative traits in plants. CropSci 46:1323–1330

Breseghello F, Sorrells ME (2006b) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. Genetics 172:1165–1177

Caldwell KS, Russell J, Langridge P, Powell W (2006) Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. Genetics 172:557–567

Chen X, Salamini F, Gebhardt C (2001) A potato molecular-function map for carbohydrate metabolism and transport. Theor Appl Genet 102:284–295

Cockram J, White J, Leigh FJ, Lea VJ, Chiapparino E et al (2008) Association mapping of partitioning loci in barley. BMC Genet 9:16

Collins A, Milbourne D, Ramsay L, Meyer R, Chatot-Balandras C, Oberhagemann P, De Jong W, Gebhardt C, Bonnel E, Waugh R (1999) QTL for field resistance to late blight in potato are strongly correlated with maturity and vigour. Mol Breed 5:387–398

D'hoop BB, Paulo MJ, Mank R, van Eck HJ, van Eeuwijk FA (2008) Association mapping of quality traits in potato (*Solanum tuberosum* L.). Euphytica 161:47–60

D'hoop BB, Paulo MJ, Kowitwanich K, Sengers M, Visser RGF et al (2010) Population structure and linkage disequilibrium unravelled in tetraploid potato. Theor Appl Genet 121:1151–1170

D'hoop BB, Paulo MJ, Visser RGF, van Eck HJ, van Eeuwijk FA (2011) Phenotypic analyses of multi-environment data for two diverse tetraploid potato collections; comparing an academic panel with an industrial panel. Potato Res 54:157–181

Douches DS, Freyre R (1994) Identification of genetic factors influencing chip color in diploid potato (*Solanum* spp). Am Potato J 71:581–590

Ersoz ES, Yu J, Buckler ES (2007) Applications of linkage disequilibrium and association mapping in crop plants. In: Varshney RK, Tuberosa R (eds) Genomics-assisted crop improvement. Springer Verlag, Germany, pp 97–119

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Plant Biol 54:357–374

Fischer M, Schreiber L, Colby T, Kuckenberg M, Tacke E, Hofferbert H-R, Schmidt J, Gebhardt C (2013) Novel candidate genes influencing natural variation in potato tuber cold sweetening identified by comparative proteomics and association mapping. BMC Plant Biol 13(1):113

Gebhardt C, Ballvora A, Walkemeier B, Oberhagemann P, Schuler 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 13:93–102

Gebhardt C, Menendez CM, Chen X, Li L, Schäfer-Pregl R et al (2005) Genomic approaches for the improvement of tuber quality traits in potato. Acta Hortic 684:85–92

Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10:381–391

Gupta PK, Rustgi S, Kulwal PL (2005) Linkage disequilibrium and association studies in higher plants: present status and future prospects. Plant Mol Biol 57:461–485

Haase N (2003) Estimation of dry matter and starch concentration in potatoes by determination of underwater weight and near infrared spectroscopy. Potato Res 46:117–127

Jacobs JME, van Eck HJ, Arens P, Verkerk-Bakker B, Lintel Hekkert B et al (1995) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. Theor Appl Genet 91:289–300

Jannink JL, Walsh B (2002) Association mapping in plant populations. In: Kang MS (ed) Quantitative genetics, genomics and plant breeding, chap 5. CABI, UK, pp 59–68

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D et al (2008) Efficient control of population structure in model organism association mapping. Genetics 178:1709–1723

Khu DM, Lorenzen J, Hackett CA, Love SL (2008) Interval mapping of quantitative trait loci for corky ringspot disease resistance in a tetraploid population of potato (*Solanum tuberosum* subsp *tuberosum*). Am J Potato Res 85:129–139

Kloosterman B, Oortwijn M, Uitdewilligen J, America T, de Vos R et al (2010) From QTL to candidate gene: genetical genomics of simple and complex traits in potato using a pooling strategy. BMC Genom 11:158

Kloosterman B, Abelenda JA, del Mar Carretero Gomez M, Oortwijn M, de Boer JM, Kowitwanich K, Horvath BM, van Eck HJ, Smaczniak C, Prat S, Visser RGF, Bachem CWB (2013) Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature 495(7440):246–250

Kraakman ATW, Niks RE, van den Berg PMMM, Stam P, van Eeuwijk FA (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168:435–446

Kraft T, Hansen M, Nilsson N-O (2000) Linkage disequilibrium and fingerprinting in sugar beet. Theor Appl Genet 101:323–326

Lærke PE, Christiansen J, Veierskov B (2002) Colour of blackspot bruises in potato tubers during growth and storage compared to their discolouration potential. Postharvest Biol Technol 26:99–111

Li L, Strahwald J, Hofferbert HR, Lubeck J, Tacke E et al (2005) DNA variation at the invertase locus invGE/GF is associated with tuber quality traits in populations of potato breeding clones. Genetics 170:813–821

Li L, Paulo MJ, Strahwald J, Lübeck J, Hofferbert HR et al (2008) Natural DNA variation at candidate loci is associated with potato chip color, tuber starch content, yield and starch yield. Theor Appl Genet 116:1167–1181

Li L, Tacke E, Hofferbert H-R, Lübeck J, Strahwald J, Draffehn AM, Walkemeier B, Gebhardt C (2013) Validation of candidate gene markers for marker-assisted selection of potato cultivars with improved tuber quality. Theor Appl Genet 126(4):1039–1052

Lindhout P, Meijer D, Schotte T, Hutten RCB, Visser RGF, van Eck HJ (2011) Towards F_1 hybrid seed potato breeding. Potato Res 54(4):301–312

Lu H, Redus MA, Coburn JR, Rutger JN, McCouch SR et al (2005) Population structure and breeding patterns of 145 US rice cultivars based on SSR marker analysis. CropSci 45:66–76

Maccaferri M, Sanguineti MC, Noli E, Tuberosa R (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. Mol Breed 15:271–289

Mackay I, Powell W (2007) Methods for linkage disequilibrium mapping in crops. Trends Plant Sci 12:57–63

Malosetti M, van der Linden CG, Vosman B, van Eeuwijk FA (2007) A mixed-model approach to association mapping using pedigree information with an illustration of resistance to *Phytophthora infestans* in potato. Genetics 175:879–889

Maris B (1969) Studies on maturity, yield, underwater weight and some other characters of potato progenies. Euphytica 18:297–319

Mather KA, Caicedo AL, Polato NR, Olsen KM, McCouch S et al (2007) The extent of linkage disequilibrium in rice (*Oryza sativa* L.). Genetics 177:2223–2232

Menendez CM, Ritter E, Schäfer-Pregl R, Walkemeier B, Kalde A et al (2002) Cold sweetening in diploid potato: mapping quantitative trait loci and candidate genes. Genetics 162:1423–1434

Montgomery DC, Peck EA, Vining GG (2001) Introduction to linear regression analysis. John Wiley & Sons Inc., New York, p 641

Nordborg M, Tavare S (2002) Linkage disequilibrium: what history has to tell us. Trends Genet 18:83–90

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J et al (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat Genet 30:190–193

Parisseaux B, Bernardo R (2004) In silico mapping of quantitative trait loci in maize. Theor Appl Genet 109:508–514

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959

Raboin L-M, Pauquet J, Butterfield M, D'Hont A, Glaszmann J-C (2008) Analysis of genome wide linkage disequilibrium in the highly polyploid sugarcane. Theor Appl Genet 116:701–714

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR et al (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98:11479–11484

Simko I, Costanzo S, Haynes KG, Christ BJ, Jones RW (2004a) Linkage disequilibrium mapping of a *Verticillium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor Appl Genet 108:217–224

Simko I, Haynes KG, Ewing EE, Costanzo S, Christ BJ et al (2004b) Mapping genes for resistance to *Verticillium albo-atrum* in tetraploid and diploid potato populations using haplotype association tests and genetic linkage analysis. Mol Genet Genomics 271:522–531

Sliwka J, Wasilewicz-Flis I, Jakuczun H, Gebhardt C (2008) Tagging quantitative trait loci for dormancy, tuber shape, regularity of tuber shape, eye depth and flesh colour in diploid potato originated from six Solanum species. Plant Breed 127:49–55

Sørensen KK, Madsen MH, Kirk HG, Madsen DK, Torp AM (2006) Linkage and quantitative trait locus mapping of foliage late blight resistance in the wild species *Solanum vernei*. Plant Breed 125:268–276

Stich B, Melchinger A, Heckenberger M, Möhring J, Schechert A, Piepho H-P (2008a) Association mapping in multiple segregating populations of sugar beet (*Beta vulgaris* L.). Theor Appl Genet 117:1167–1179

Stich B, Möhring J, Piepho H-P, Heckenberger M, Buckler ES et al (2008b) Comparison of mixed-model approaches for association mapping. Genetics 178:1745–1754

Stich B, Piepho H-P, Schulz B, Melchinger A (2008c) Multi-trait association mapping in sugar beet (*Beta vulgaris* L.). Theor Appl Genet 117:947–954

Thomas WTB, Powell W, Waugh R, Chalmers KJ, Barua UM, Jack P, Lea V, Forster BP, Swanston JS, Ellis RP, Hanson PR, Lance RCM (1995) Detection of quantitative trait loci for agronomic, yield, grain and disease characters in spring barley (*Hordeum vulgare* L.). Theor Appl Genet 91:1037–1047

Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D et al (2001) Dwarf8 polymorphisms associate with variation in flowering time. Nat Genet 28:286–289

Urbany C, Stich B, Schmidt L, Simon L, Berding H, Junghans H, Niehoff K-H, Braun A, Tacke E, Hofferbert H-R, Lübeck J, Strahwald J, Gebhardt C (2011) Association genetics in Solanum tuberosum provides new insights into potato tuber bruising and enzymatic tissue discoloration. BMC Genom 12:7

Van der Beek JG, Verkerk R, Zabel P, Lindhout P (1992) Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: Cf9 (resistance to *Cladosporium fulvum*) on chromosome 1. Theor Appl Genet 84:106–112

Van Eck HJ, Jacobs JME, Stam P, Ton J, Stiekema WJ et al (1994a) Multiple alleles for tuber shape in diploid potato detected by qualitative and quantitative genetic analysis using RFLPs. Genetics 137:303–309

Van Eck HJ, Jacobs JME, van den Berg PMMM, Stiekema WJ, Jacobsen E (1994b) The inheritance of anthocyanin pigmentation in potato (*Solanum tuberosum* L.) and mapping of tuber skin colour loci using RFLPs. Heredity 73:410–421

Van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ et al (2006) Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. Genetics 173:1075–1087

Van Zanten M, Snoek LB, Proveniers MCG, Peeters AJM (2009) The many functions of ERECTA. Trends Plant Sci 14:214–218

Verbeke G, Molenberghs G (2000) Mixed models for longitudinal data. Springer, Berlin, p 568

Visker MHPW, Keizer LCP, van Eck HJ, Jacobsen E, Colon LT et al (2003) Can the QTL for late blight resistance on potato chromosome 5 be attributed to foliage maturity type? Theor Appl Genet 106:317–325

Visker MHPW, Heilersig HJB, Kodde LP, van de Weg WE, Voorrips RE et al (2005) Genetic linkage of QTLs for late blight resistance and foliage maturity type in six related potato progenies. Euphytica 143:189–199

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T et al (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wang J, McClean PE, Lee R, Goos RJ, Helms T (2008) Association mapping of iron deficiency chlorosis loci in soybean (*Glycine max* L. Merr.) advanced breeding lines. Theor Appl Genet 116:777–787

Wang-Pruski G, Nowak J (2004) Potato after-cooking darkening. Am J Potato Res 81:7–16

Werij JS, Kloosterman B, Celis-Gamboa C, de Vos CHR, America T et al (2007) Unravelling enzymatic discolouration in potato through a combined approach of candidate genes, QTL, and expression analysis. Theor Appl Genet 115:245–252

Wolters A-M, Uitdewilligen J, Kloosterman B, Hutten RCB, Visser RGF et al (2010) Identification of alleles of carotenoid pathway genes important for zeaxanthin accumulation in potato tubers. Plant Mol Biol 73:659–671

Yu J, Pressoir G, Briggs WH, Vroh BI, Yamasaki M et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zhao J, Paulo MJ, Jamar D, Lou P, van Eeuwijk FA et al (2007a) Association mapping of leaf traits, flowering time, and phytate content in Brassica rapa. Genome 50:963–973

Zhao KY, Aranzana MJ, Kim S, Lister C, Shindo C et al (2007b) An Arabidopsis example of association mapping in structured samples. PLoS Genet 3:e4

Zhu C, Gore M, Buckler ES, Yu J (2008) Status and prospects of association mapping in plants. Plant Genome 1:5–20