Association mapping of quality traits in potato (*Solanum tuberosum* L.)

**Björn B. D'hoop · Maria João Paulo · Rolf A. Mank ·
Herman J. van Eck · Fred A. van Eeuwijk**

Abstract In this paper, we describe the assessment
of linkage disequilibrium and its decay in a collection
of potato cultivars. In addition, we report on a simple
regression based association mapping approach and
its results to quality traits in potato. We selected 221
tetraploid potato cultivars and progenitor lines, repre-
senting the global diversity in potato, with emphasis
on genetic variation for agro-morphological and qual-
ity traits. Phenotypic data for these agro-morphologi-
cal and quality traits were obtained from recent trials
performed by five breeding companies. The collection
was genotyped with 250 AFLP® markers from five
primer combinations. The genetic position of a subset
of the markers could be inferred from an ultra dense
potato map. Decay of linkage disequilibrium was esti-
mated by calculating the squared correlation between
pairs of markers using marker band intensities.
Marker-trait associations were investigated by fitting
single marker regression models for phenotypic
traits on marker band intensities with and without
correction for population structure. The paper illus-
trates the potential of association mapping in tetra-
ploid potato, because existing phenotypic data, a
modest number of AFLP markers, and a relatively
simple statistical analysis, allowed identifying inter-
esting associations.

Keywords AFLP · Association mapping · Linkage
disequilibrium · Potato · Quality

Introduction

Since quality demands of consumers and potato pro-
cessing industry have become increasingly stringent,
breeding efforts need to focus on quality traits in
potato. Association mapping, also known as linkage
disequilibrium mapping or gametic phase disequilib-
rium mapping—originally developed to study genetic
disorders in humans—is a recently emerging tool in
plant breeding research. Association mapping differs
from classical Quantitative Trait Locus (QTL) map-
ping. An important distinction is that no segregating
offspring have to be grown and phenotyped, since
association mapping can deal with collections of
existing cultivars. Another distinction is that parent
choice is less of a dilemma, nor a limitation on the
genetic diversity. A further difference is that pheno-
typing efforts may be reduced as existing phenotypic
data from national lists, gene banks and breeding
companies can be used in addition to or as a replace-

B. B. D'hoop (✉) · M. J. Paulo · H. J. van Eck
Laboratory of Plant Breeding, Wageningen University,
P.O. Box 386, 6700AJ Wageningen, The Netherlands
e-mail: bjorn.dhoop@wur.nl

R. A. Mank
Keygene N.V., Wageningen, The Netherlands

F. A. van Eeuwijk
Biometris, Wageningen University Research,
P.O. Box 100, 6700AC Wageningen, The Netherlands

ment of present day trials. Still, designed trials for the evaluation of selected germplasm will provide more reliable results. Lastly, some forms of association mapping, like the one presented in this paper and the one in Malosetti et al. (2007), provide a relatively simple approach to the identification of QTLs in tetraploids, comparable to the use of marker regression in segregating populations, but applied to a wider genetic background.

Association mapping has been and remains a popular research tool in human and animal genetics. Human disease genetics was the first area for which association mapping methodology was developed and where successes were achieved (Carlson et al. 2004; Jorde 2000; Lander and Schork 1994). In animal genetics, most concern was about the LD patterns in breeding populations to determine to what extent LD holds and the marker density required to fine map genes, e.g. in cattle (Farnir et al. 2000), pig (Nsengimana et al. 2004) and sheep (McRae et al. 2002). In plant genetics, however, fewer papers have been produced so far on association mapping, except for model plant systems like *Arabidopsis thaliana* (Hagenblad and Nordborg 2002; Nordborg et al. 2002), maize (Ching et al. 2002; Palaisa et al. 2003, 2004; Parisseaux and Bernardo 2004; Rafalski and Morgante 2004; Remington et al. 2001; Tenaillon et al. 2001) and rice (Garris et al. 2003; Lu et al. 2005; Semon et al. 2005), where association mapping is gaining importance due to the development of high-throughput marker systems and the availability of genome sequences. Other plant systems have been investigated for LD patterns and associations as well. In barley Kraakman et al. (2004, 2006) performed association analysis for agronomical, resistance related and morphological traits. Conifers were examined using the candidate gene approach to dissect complex traits (Neale and Savolainen 2004). Barnaud et al. (2006) reported on the LD pattern of grapevine. In *Lolium perenne*, Skøt (2005) found associations for heading date. LD patterns were examined deploying SNPs in soybean (Zhu et al. 2003), AFLPs in sugar beet (Kraft et al. 2000) and RFLPs in sugarcane (Jannoo et al. 1999). Kernel size and milling quality were fine mapped applying association mapping in wheat (Breseghello and Sorrells 2006). A complete overview on association mapping and the status of this approach in plants has been published recently (Gupta et al. 2005). In conclusion, although the amount of publications on association mapping in plant genetics is not as high as it is for animal or human genetics, there is a clear trend visible towards increasing number of plant species on which association mapping research has been performed with success.

Genetic and QTL studies in potato have predominantly been performed on diploid populations, e.g. leaf blight resistance QTLs have been detected in a di-haploid clone (Bradshaw et al. 2006), in a diploid *S. phureja* × *S. stenotomum* hybrid population (Costanzo et al. 2005; Simko et al. 2006), in diploid *S. vernei* (Sørensen et al. 2006), in diploid *S. microdontum* (Bisognin et al. 2005) and in diploid *S. tuberosum* (Visker et al. 2005). Cold sweetening QTLs have been mapped by Menendez et al. (2002) in diploid potato. Recently, tetraploid potato populations have been used as well. Simko and co-researchers mapped *Verticillium* resistance genes in tetraploid populations using various association mapping approaches, such as haplotype association—a whole genome approach—and candidate genes—a local approach (Simko et al. 2004a, b). With a limited number of targeted markers, Bormann et al. (2004) tagged QTLs for maturity corrected late blight resistance in tetraploid potato. However, the QTL studies in the tetraploid mapping population 12601ab1 × Stirling are unique in demonstrating the complexity of map construction and QTL analysis (Bradshaw et al. 2004; Bryan et al. 2004).

Association mapping methodology in potato till now is mainly restricted to standard two sample t-tests or Mann-Whitney U tests. Gebhardt et al. (2004) for example, deployed the non-parametric Mann-Whitney U test for an association assay on 415 cultivars, genotyped with 5 PCR markers closely linked to previously mapped QTL for late blight resistance and plant maturity, with evaluation data for late blight resistance and maturity. Simko et al. (2004a, b) detected significant association between *Verticillium dahliae* resistance and an SSR marker closely linked to a candidate gene locus for *Verticillium* resistance (*StVe1*) in tetraploid potato using a standard two sample *t*-test. *Verticillium albo-atrum* resistance appeared associated with a specific haplotype of the same candidate gene locus in tetraploid potato, detected using haplotype association with the Wilcoxon two sample test. Recently, Malosetti et al. (2007) presented a general mixed model approach to LD mapping, using potato data for illustration, giving attention to the incorporation of relationships between genotypes,

whether induced by population substructure, pedigree or other, and the combination of data arising from multiple environments. Their approach is broadly applicable, without the need for specific software. The approach explained by Malosetti et al. (2007) is very similar to the mixed model approach of Yu et al. (2006). In the latter paper, the mixed methodology was illustrated by an example in maize.

In this paper, we will present preliminary results on the utility of a simple association mapping strategy for tetraploid potato using existing data from five potato breeding companies. With a limited number of markers and relatively simple regression models we detected marker-trait associations for a number of quality traits such as cold sweetening (Li et al. 2005; Menendez et al. 2002), after cooking darkening (ACD) (Friedman 1997; Wang-Pruski and Nowak 2004), enzymatic browning and/or blackspot bruising as catalysed by Polyphenol oxidase (PPO) (Lærke et al. 2002). These first results are very promising considering this was a preliminary experiment performed to sort out association mapping methodology before embarking on a more detailed and thorough follow up study that is presently undertaken.

Materials and methods

Plant material

With the aim to collect a representative subset of worldwide available potato germplasm, we selected 221 tetraploid potato cultivars and progenitor lines. The selection was based on (1) the acreage for seed potato production in the Netherlands, (2) the country of origin, (3) the year of commercial introduction, (4) the market niche (table, starch and processing), (5) phenotypic diversity for quality traits and (6) availability of the cultivars, especially for older genotypes. The material was kindly provided by Dutch breeding companies and several gene banks (see acknowledgements). Leaf material was harvested from greenhouse-grown and in-vitro-grown genotypes, was frozen with liquid nitrogen and stored at −80°C until DNA extraction.

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 5 well known Eco/Mse primer combinations: E + AAC/M + CAC, E + AAC/M + CAG, E + AAC/M + CCA, E + ACA/M + CAC, E + AGA/M + CAT. Fragments were separated using a capillary sequencer (MegaBACE 1000, Molecular Dynamics & Amersham, serial number 13757) according to van Eijk et al. (2004), each primer combination being labelled with either FAM, NED or JOE. The ROX channel was used for the MegaBACE™ ET900-Rox size standard from GE Healthcare (Amersham Biosciences).

Pseudo gel images were scored at Keygene N.V. Marker nomenclature was based on primer combination and fragment mobility.

Normalisation of signal intensity variation between capillaries due to DNA loading effects was performed on log-transformed band intensity values using the ANOVA procedure in GenStat, release 8.11 (VSN International Ltd., Oxford, UK).

Position information of AFLP markers was retrieved from the ultra dense potato map, using the parental diploid genotypes SH83-92-488 and RH89-039-16 as internal reference (van Os et al. 2006; http://potatodbase.dpw.wau.nl/UHDdata.html).

Phenotypic data collection

An overview of the agro-morphological and quality traits of this study is shown in Table 1.

The agro-morphological traits tuber shape, flesh colour and maturity are scored by breeders on an ordinal scale. Trait values do not imply breeder or consumer preference. The flesh colour shades go from white to orange and can be interpreted as a ranking. Under water weight was assessed on a continuous scale and was based on a 5 kg sample size.

Quality traits were all scored on an ordinal scale (Table 1). Cold sweetening was not measured as such, but through the highly correlated characteristics frying and chipping colour. All tubers were stored at 8°C before frying analysis. Frying colour was determined after 3 min of frying at 180°C, chipping colour after frying at 180°C until no more air bubbles appear. After baking darkening was determined following 2 min of pre-frying at 160°C, merely scoring the amount of greyness which has a different causal back-

Table 1 Overview of phenotypic traits, their data types, scales and the numbers of genotypes for which data were provided by five breeding companies for these 11 phenotypic traits

Trait	Type	Scale	Agrico	Averis	CMK	HZPC	Van Rijn
Tuber shape	Ordinal	1 = long, 5 = oval, 9 = round with intermediates	124	6	41	145	20
Flesh colour	Ordinal	4 = white, 5 = cream, 6 = light yellow, 7 = yellow, 8 = dark yellow, 9 = orange	123	6	41	145	20
Under water weight	Continuous	g per 5 kg tubers	78	6	41	142	20
Maturity	Ordinal	Minimum 1 (extremely early) to 9 (extremely late) with increments of 1	124	6	41	142	20
Cooking type	Ordinal	1 (A, firm) to 7 (D, very mealy) with increments of 1	110	–	36	141	20
After cooking darkening	Ordinal	Minimum 3 (very dark) to 8 (very light) with increments of 1	113	–	36	141	20
After baking darkening	Ordinal	Minimum 3 (very dark) to 8 (very light) with increments of 1	–	–	–	126	–
Frying colour	Ordinal	Minimum 3 (very dark) to 8 (very light) with increments of 1	111	–	16	124	18
Chipping colour	Ordinal	Minimum 3 (very dark) to 8 (very light) with increments of 1	110	–	31	109	16
Blackspot bruising	Ordinal	Minimum 0 (no tubers bruised) to maximum 50 (all tubers heavily bruised) with intermediates	–	6	41	130	19
Enzymatic browning	Ordinal	Minimum 3 (very dark) to 8 (very light) with increments of 1	–	–	–	52	–

Total number of genotypes was 221. Absence of information is indicated by –

ground than frying colour. After cooking darkening was phenotyped after steam cooking at 100°C, as the severity of greyness. To determine the bruising score, a sample of 50 tubers representative for the field plot was taken—when available—and shaken on a mechanical shaker. One week later, the tubers were peeled and subdivided into classes of severity of bruising (mild, medium, severe) according to size and frequency of blue spots. To determine the bruising score, tuber amounts per severity class were counted and weighed applying the following formula: (50/ number of tubers used) * ((number of mild + 2*number of medium + 3*number of severe)/3). For enzymatic browning, tubers were rasped and exposed to air and at a fixed time point enzymatic browning severity was scored. Cooking type was scored following steam cooking.

Phenotypic data came from field trials performed by five Dutch breeding companies (see acknowledgements). For a number of traits, the trait values for individual genotypes represented summary statistics across years and locations, calculated following company specific procedures that were not accessible for us. The phenotypic dataset was unbalanced because the breeding companies analysed different sets of cultivars and traits (Table 1).

Statistical procedures

Linkage disequilibrium (LD) between loci was quantified by the squared correlation coefficient, r^2, between the (logarithm of) band intensities (Flint-Garcia et al. 2003; Remington et al. 2001; Zhao et al. 2005). LD decay was investigated by plotting r^2 versus map distance.

We chose band intensities for the estimation of LD and marker-trait associations and not band presence/absence, because in a specially designed pilot experiment we found band intensity to be related to allele dosage (unpublished). The main conclusion of the pilot was a clear intensity difference between no allele and at least one allele, but not a very strong trend between intensity and allele dosage in the range of one to four alleles. We have experimented

with different functions representing the relation between band intensity and allele dosage, but in the end decided that at the moment of writing the most robust approach was to use (the logarithm of) band intensity as such. The use of band intensity produced results that were very closely the same to results using band presence/absence. In genetic terms this means that we worked with a model in which band presence is fully dominant, i.e., it does not matter whether a band corresponds to an allele dosage of 1, 2, 3 or 4, the effect on the phenotype will be the same for any dosage.

The problem of possible population structure in the potato cultivar collection was approached in two ways. Firstly, as an a posteriori approach, using MEGA 3.1 (Kumar et al. 2004), a neighbour joining dendrogram was calculated based on Euclidian distances of the AFLP band intensities. This dendrogram was subsequently inspected for possible clusters. These clusters would then indicate subpopulations. A contrasting, a priori approach consisted in assuming that the breeding companies that developed the cultivars represented subpopulations. In the association analysis, correction for population structure can then be achieved by inclusion of a factor whose levels are defined by the breeding companies.

A two step approach to association mapping was used. First, Best Linear Unbiased Estimators (BLUEs) were calculated per trait for the genotypes, using GenStat, release 8.11 (VSN International Ltd., Oxford, UK). For each trait an additive analysis of variance model was fitted, response = breeding company + genotype + error, where the breeding company factor referred to the company that provided the data for the particular trait (see Table 1). As the breeding companies differed in the set of genotypes they evaluated, the data were unbalanced and therefore the BLUEs for the genotypes were adjusted means.

The traits were all analysed as quantitative traits, assuming the error variation to be normally distributed with constant variance, although the majority of the traits was measured on an ordinal scale. Inspection of diagnostic plots for residuals indicated no strong violations of these assumptions.

After calculation of genotypic BLUEs, markers were screened for association with traits. For each trait and for each of the 149 markers, two models were fitted;

1) Response = marker + error
2) Response = breeder + marker + error

The marker information that entered the models consisted of the logarithm of the band intensity, so that the markers were introduced as continuous variables for which a slope parameter was estimated by standard least squares. The breeder term refers to the breeding company that developed the particular cultivar. This term is supposed to correct for possible population structure.

Results

LD decay and population structure

Markers which could be scored in an unambiguous manner were retained. In total 551 AFLP fragments were distinguished on 227 genotypes, including a three-fold repetition of the diploid mapping parents SH83-92-488 and RH89-039-16 (van Os et al. 2006), which were included as internal reference genotypes. From these 551 AFLP markers 149 markers could be assigned to mapped AFLP bands in the ultra dense potato map. We assumed that the position of these 149 markers in the diploid UHD map is essentially not different from the position of these bands in the collection of tetraploid genotypes. Besides position information, the AFLP bands represent alleles on two homologous chromosomes from the female and male mapping parents. Although male and female cM positions may differ due to differences in recombination in male and female meiosis, we did not attempt to assign sex-averaged recombination distances between the markers. Markers from the same homology group were tested for LD while ignoring from which homologous chromosome they were derived.

Linkage disequilibrium (LD) decay across the potato genome was calculated deploying r^2, the squared correlation between the (log-) band intensities of AFLP marker loci. A graphical representation, plotting the resulting r^2-values in relation to the UHD map distance in centiMorgan (cM) between the AFLP markers, was constructed. Linkage disequilibrium between markers mapped on different non-homologous chromosomes is embedded in the same graph by using ∞ cM as symbolic distance between these unlinked markers (Fig. 1). In this AFLP marker
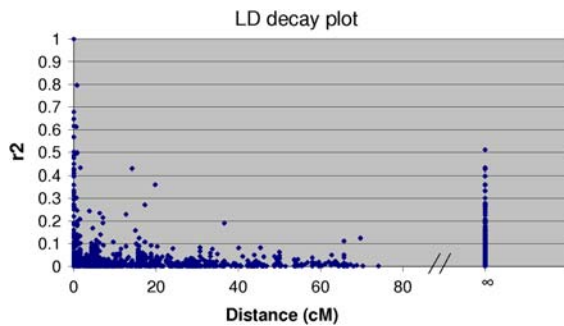
Fig. 1 Linkage disequilibrium (LD) decay plot. The squared correlation between paired marker intensities on the y-axis is plotted against the distance between pairs of markers in centi-Morgan (cM) on the x-axis. LD between pairs of markers on different non-homologous chromosomes is represented by the symbol ∞. The commonly accepted reference value for r^2 of 0.1, indicates LD to decay at about 3 cM. The critical test value for r^2 is 0.03 for a test level of 0.01. Using this critical value as threshold, LD decays at about 8 cM

dataset r^2 seems to drop below 0.1 at about 3 cM, which is promising for LD mapping in potato and its applications in marker assisted breeding. For the current sample size of 221 and a significance level of 0.01, the critical value for r^2 is 0.03. Applying this value as reference, LD seems to disappear after about 8 cM.

Population structure of our cultivar and progenitor line collection was examined, using the neighbour joining facilities of the package MEGA 3.1 for clustering (Kumar et al. 2004) and defining distance between genotypes by the Euclidian distance on the basis of the AFLP (log-) band intensities. Although there is a suggestion of some clusters of cultivars sharing country of origin or market niche, the neighbour joining analysis revealed no clear structural relationships, as demonstrated by the very short branches separating clusters from the central point of the astral projection of the dendrogram (Fig. 2).

Marker-trait associations

Numbers of marker loci associated with traits at a test level of 0.01 are given for the models with and without correction for breeding origin of the material in Table 2. The breeder term was significant for all traits. As expected, the number of marker-trait associations typically decreased after inclusion of the correction for breeder. Since all 149 AFLP markers in the genotypic dataset had a known map location, the

loci associated with particular traits could be allocated to a potato chromosome. Table 2 offers an overview per trait of the detected loci together with their chromosome allocations.

Table 2 also presents a comparison between results described in literature and our association mapping results. In most instances, detected QTLs coincide with chromosomes for which associations with the respective trait were previously reported. For cooking type and tuber shape, our QTL chromosomes differed from literature reports. For maturity, after cooking darkening, chipping colour and enzymatic browning more loci were detected than described in literature. For the other traits like flesh colour, under water weight and frying colour less loci were detected than reported elsewhere. This paper presents the first QTLs reported for after baking darkening and black spot bruising.

Illustration for two marker trait associations

Choosing a significance level of 0.01, which involves a very modest multiple testing correction, eleven AFLP marker loci were detected for maturity, scattered over six chromosomes of the potato genome, namely 1, 4, 5, 6, 9 and 10 (Table 2). The observed distribution of maturity is depicted in Fig. 3. The histogram in Fig. 3 illustrates the spectrum of earliness as determined throughout the collection of 221 potato cultivars and progenitor lines. Unfortunately, there was no maturity data available for all 221 genotypes, only 155 genotypes were represented in the phenotypic dataset.

Because literature has reported several times on a major QTL influencing maturity residing on chromosome 5, we were interested in the effect of a chromosome 5 AFLP marker associated with maturity in our collection of genotypes. Figure 4 illustrates the effect of this chromosome 5 marker in the context of the observed data when portrayed by boxplots; marker presence will result in an earlier ripening phenotype. Observed difference between the medians of genotypes with and without band was 0.83 units. The estimate for the effect of the band on maturity in the association model was 0.97 units with an estimated standard error of 0.20. The F-test for association of this marker with maturity had a P-value of 10e−6.

Frying colour, an indicator for cold sweetening, appeared associated with three AFLP marker loci at the significance level of 0.01, spread over three chro-
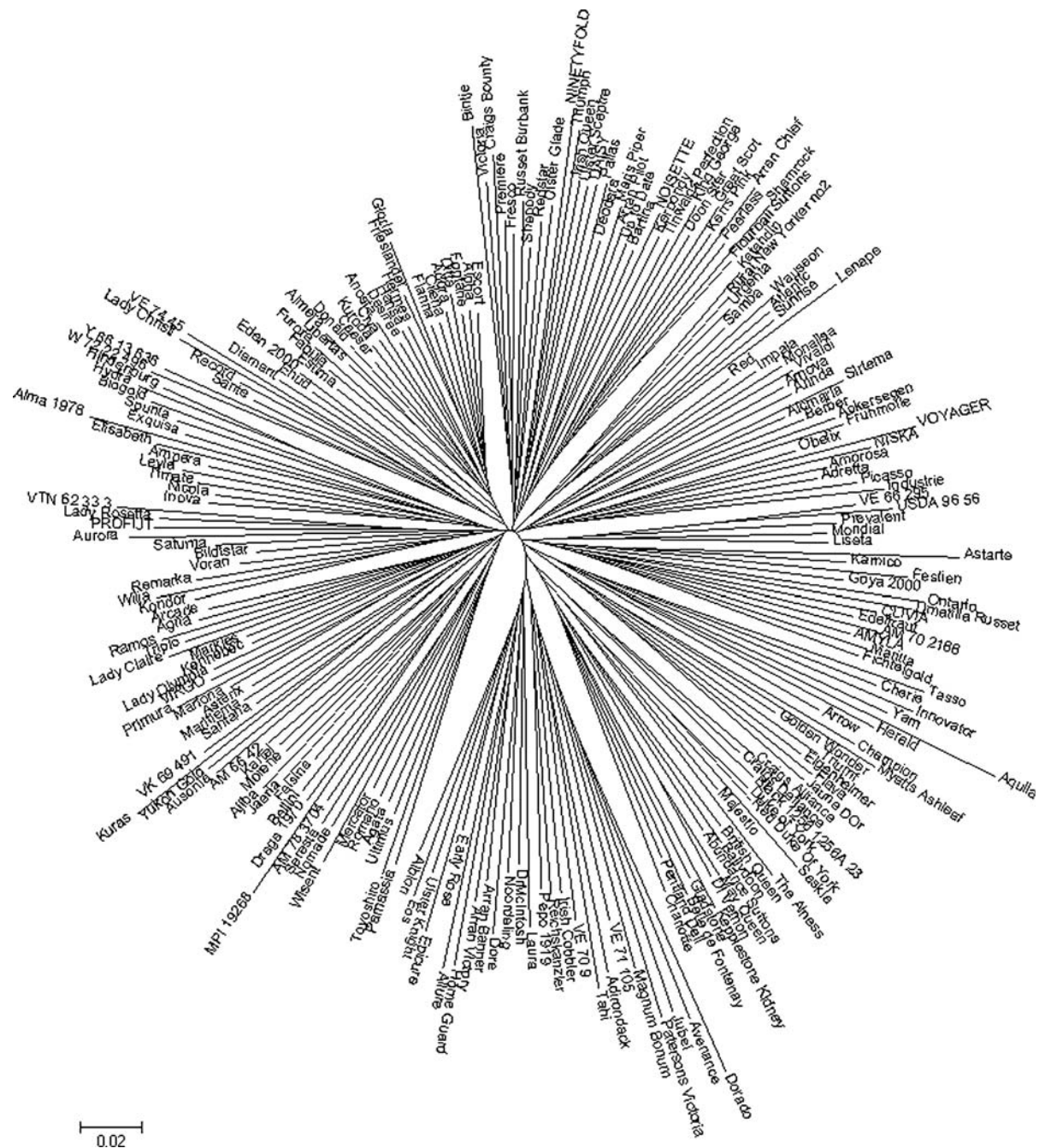
Fig. 2 Neighbour joining dendrogram visualised with MEGA3.1 (Kumar et al. 2004). Euclidian distances between band intensities for 149 AFLP markers defined distance between genotypes. The population of genotypes comprised 221 potato cultivars and progenitor lines

mosomes of the potato genome, namely 1, 2 and 7 (Table 2). The distribution of this trait is depicted in Fig. 5, where the histogram reflects the spectrum of frying colour as determined throughout the collection of 221 potato cultivars and progenitor lines. Unluckily, there was again no frying colour data available on all 221 genotypes, only 145 genotypes were represented in the phenotypic dataset.

Table 2 This table presents on the left side the number of detected marker-trait associations deploying the model without (Simple model) and deploying the model with correction for population structure (Structure model) together with their chromosomal location, based on the UHD map, for the 11 traits studied
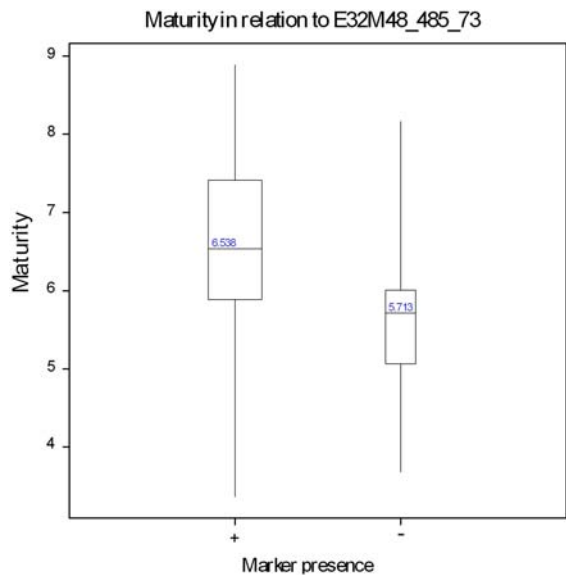
Trait	Number of associated loci with P-value ≤ 0.01		Chromosomal location of loci detected with "structure model"	Literature		Reference
	Simple model	"Structure model"		Number of loci	Chromosomal location	
Tuber shape	6	5	2, 4, 11	1	10	van Eck et al. (1994b)
Flesh colour	7	6	1, 2, 3 (31.5 cM), 4, 6, 7	5	3 (Y-locus, 20 cM), 10	van Eck et al. (1994a); Bonierbale et al. (1988); Jacobs et al. (1995)
Under water weight	16	6	1 (8.4, 9.2 & 24.9 cM), 4 (21.3 cM), 5 (16.2 cM), 9(46 cM)	–	1 (ts(s), 5 cM), 4 (ts(i), 7 cM), 5 (ts(a), 14 cM), 9 (ts(o), 24 cM),...	Gebhardt et al. (2005)
Maturity	13	11	1, 4, 5 (16.2 cM), 6, 9, 10	2	3, 5 (GP21, 15 cM)	Gebhardt et al. (2004); Visker et al. (2005)
Cooking type	6	6	1, 2, 6, 10	StTLRP candidate gene	9	Kloosterman 2006
After cooking darkening	17	7	1 (0.8 & 24.9 cM), 2, 4 (22.8 cM), 5, 7 (54.5 cM), 10	ACO, CIS, HQT candidate genes, ...	1 (CIS, 42 cM), 7 (ACO, 42 cM), ...	Chen et al. (2001); Wang-Pruski and Nowak (2004)
After baking darkening	10	10	1 (9.2 & 24.9 cM), 3, 4 (27.5, 30 & 32.9 cM), 6, 7 (55.3 cM), 11	–	–	
Frying colour	4	3	1 (4.6 cM), 2 (3.7 cM), 7 (63 cM)	24	1(Sug1 b, 10 cM), 2(Sug2a, 9 cM), 7 (Sug7a,c,e, SPS, 60 cM), ...	Menendez et al. (2002)
Chipping colour	4	6	1, 2 (3.7 cM), 4 (27.5 cM), 7, 10 (49.3 cM)	6	2 (110.1), 4 (H14.1, 35 cM), 5, 10 (G13.3, 50 cM)	Douches and Freyre (1994)
Blackspot bruising	5	2	1, 3	–	–	
Enzymatic browning	6	6	1, 4, 7, 8 (10.6 cM), 11	POT candidate genes	8 (PPO-III, 25 cM)	Bachem et al. (1994)

On the right side a comparison is presented with associations described in literature. Precise chromosomal locations are presented whenever there is a chromosome match between literature and our association mapping results
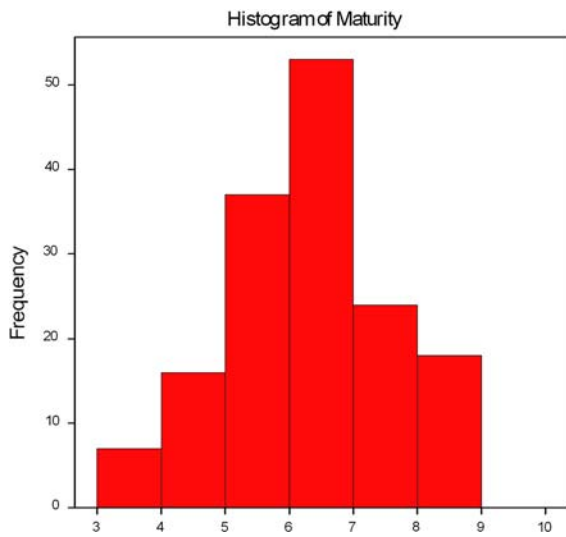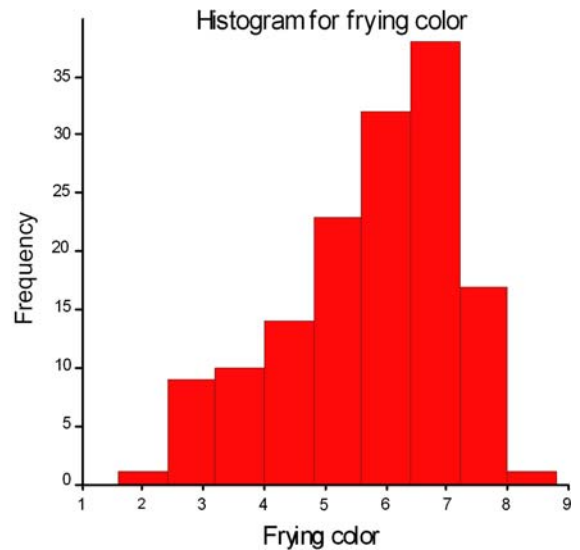
Fig. 3 Histogram for maturity. Number of observations was 155

Fig. 5 Histogram for frying colour. Number of observations was 145

Fig. 4 Boxplots for maturity using presence and absence of an associated AFLP marker on chromosome 5 to define the groups. The presence and absence of the specific marker was determined using peak detection on the pseudo gel image. The median values are written in the boxes, the difference in median between genotypes with and without band was 0.83 points on the maturity scale, meaning that a genotype having this marker band will in general be earlier ripening in the field

Literature has reported on the location of cold sweetening influencing loci throughout the whole potato genome (Menendez et al. 2002). Therefore we decided to look at the effect of the molecular marker

showing the strongest significant association with frying colour in our collection of genotypes. Figure 6 illustrates the effect of this chromosome 7 marker. As can be deduced from the boxplot, marker presence will result in a slightly better frying quality, more specific a less dark frying colour, meaning a less severe cold sweetening phenotype. Observed difference in median frying colour between genotypes with and without the band at this marker was 0.38 units. The effect of band presence for this marker was estimated to be 0.37 units with a standard error of 0.13. The F-test for association for this marker had a P-value of 0.004.

Discussion

Association mapping approach

When comparing LD mapping in tetraploid potato with QTL analysis on segregating populations, LD mapping is less cumbersome than standard QTL mapping: it is less time-consuming as no segregating offspring need to be grown and phenotyped and no parent choice needs to be made. For LD mapping, the relationships between marker and trait loci are not merely explained by genetic distance, but also by the number of meiotic generations separating genotypes from each other. On the one hand, this renders the
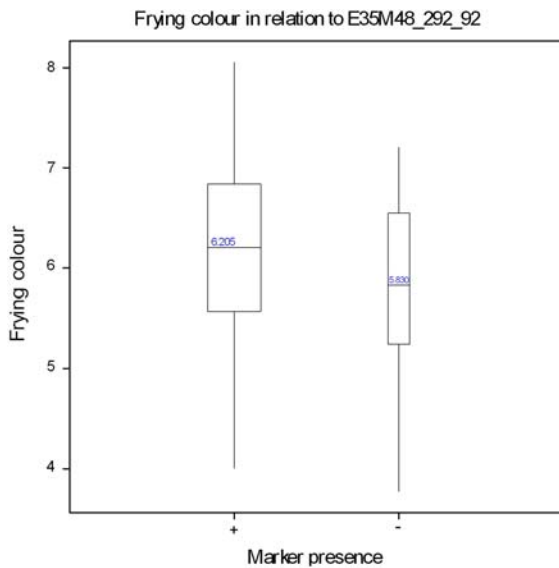
Fig. 6 Boxplots for cold sweetening using presence and absence of an associated AFLP marker on chromosome 7 to define the groups. The presence and absence of the specific marker was determined using peak detection on the pseudo gel image. The median values are written in the boxes. The difference in median between genotypes with and without the band was 0.38 points on the frying colour scale, meaning that a genotype having this marker band will in general have a less dark frying colour, i.e. a higher frying quality, than a genotype not expressing this marker band

genetics more complex, because population admixture, selection or genetic drift, can bias detected associations. On the other hand, due to recombination events occurring each meiotic generation, stretches of linkage disequilibrium will tend to be somewhat shorter, which can be advantageous to pinpoint genetic markers closer to a trait locus of interest. In LD studies in potato, the role of recombination may be restricted. Gebhardt et al. (2004) showed for a subset of German potato germplasm that there are only 4–5 meiotic generations separating heirloom potatoes from their recent descendants. In contrast, our collection contains exceptional cultivars, like for example Biogold being separated by 10 meiotic generations from his ancestors (van Berloo et al. 2007).

LD decay

We estimated the linkage disequilibrium pattern of the potato genome deploying r^2 between (log-) band intensities, thereby following an approach similar to Remington et al. (2001) and Flint-Garcia et al.

(2003). Applying the commonly used reference value of 0.1 for r^2, we found that LD seems to drop at about 3 cM. LD decay in potato has already been addressed by Gebhardt et al. (2004), where they examined four DNA markers within 1 cM on chromosome V. For two markers residing within 0.3 cM LD was maintained, whereas for markers being separated by 0.6 cM LD had decreased and for markers being 0.9 cM apart linkage equilibrium had been reached. Simko et al. (2006) found indications of a fast decaying linkage disequilibrium on short range ($r^2 = 0.208$ at 1 kb) but then afterwards a slower decreasing LD ($r^2 = 0.137$ at 70 kb) was found. They conclude from their data that LD decayed below 0.10 at about 10 cM genetic distance.

Population structure

We examined population structure with the neighbour joining algorithm using Euclidian distance on the AFLP marker band intensities. This revealed no clear clusters, although some geographically isolated cultivars (USA and UK) and some market niche cultivars (starch processing) tended to group together (Fig. 2). Cultivars known to be mutants of one another, for example Duke of York (Eersteling) and Red Duke of York (Rode Eersteling) or cultivars known to be closely related to each other, such as Craig's Defiance and Craig's Alliance or Arran Banner and Arran Victory, clustered nicely together. Certainly, the clustering of cultivars could change if more marker data had been included in the cluster analysis, but still, this dataset, although small, sets a trend and more or less confirms what could be expected, namely that other continents and different market niches use to some extent a different gene pool.

Our correction for origin of germplasm in the form of inclusion of breeding company in the association analyses was successful in that this term was judged to be significant in the models for all phenotypic traits. Thus, it made sense to include this correction for population structure in the models for association.

Relations between traits

As illustrated in Table 2, some different types of discolouration have loci on the same chromosomes, namely chromosomes 2 and 7. For the traits chipping

colour and frying colour this confirmed our expectations, since both traits are used as indirect indicators of cold sweetening. Moreover, both traits were correlated in our dataset ($r = 0.49$) and the detected loci on chromosomes 2 and 7, at 3.7 cM and 63.0 cM respectively, seemed pleiotropic because they were associated with the same AFLP markers. Contrarily, we did not expect after cooking darkening to share loci with cold sweetening nor with enzymatic browning, since these traits are caused by different biochemical processes. Indeed, no associated markers were found for those traits that showed association with more than one trait, thereby providing more evidence for close linkage of different QTLs than pleiotropic effects of a single QTL. Nevertheless, with the current precision for marker positions, it is not possible to draw firm conclusions on pleiotropy versus linkage as long as the markers are within 5–10 cM from each other.

The traits after cooking darkening and after baking darkening shared loci on chromosomes 1, 4 and 7 (see Table 2). Except for one marker tagging chromosome 1, the detected loci were not associated with the same AFLP markers. Nevertheless, the markers of these associated loci map closely together on the UHD map which could indicate a common QTL. This confirms our expectations, since both traits are assumed to be caused by the same process (see M&M), the only difference between both traits lies in the temperature to which the samples have been exposed before darkening took place.

Prospects

We used AFLP markers for an association study in potato using a compilation of existing data belonging to different breeding companies. We found in total 69 associated AFLP marker loci for 11 phenotypic traits at the significance level of 0.01, spread throughout the entire potato genome (Table 2), showing that even this simply constructed phenotypic historic data set in combination with a modest number of markers of the widely available and easy to use AFLP marker system and a relatively straightforward statistical analysis was able to produce interesting results for applied breeding. Our association approach identified markers that could be useful in the development of marker assisted breeding strategies. These initial results obtained with association mapping should be compared with the power to detect loci involved in quality traits with conventional QTL mapping approaches (both at the diploid and at the tetraploid level). At this moment we regret that in spite of the number of QTL studies on pathogen resistance, the number of QTL studies on quality traits is limited. We anticipate that association mapping will appear a valid method to analyse complex traits in tetraploid potato and will be used along with diploid QTL mapping studies. We can, however, compare the input required by either approach. Our association approach and QTL mapping in diploid mapping populations are both relatively simple in methodology and modestly demanding with respect to genotyping. This is in contrast to the genotyping and methodological demands for QTL mapping using tetraploids (Luo et al. 2006). A similar conclusion was obtained by Malosetti et al. (2007). Although these promising results were based on existing phenotypic data, we will invest in an improved phenotypic dataset. This will be achieved through a field trial with replicates including all cultivars and progenitor lines of the present study. Furthermore, we aim to expand the current marker dataset with additional AFLP markers, but also candidate genes and microsatellites. To further improve on the methodological part of the association approach, research is going on with more advanced mixed models that are able to correct for pedigree relationships and can introduce different marker types into the model. Finally, we aim at obtaining allele dosage observations and estimates for inclusion in our genetic models.

which is part of the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research.

References

Bachem CWB, Speckmann G-J, van der Linde PCG, Verheggen FTM, Hunt MD, Steffens JC, Zabeau M (1994) Antisense expression of polyphenol oxidase genes inhibits enzymatic browning in potato tubers. Bio-technology 12:1101–1105

Barnaud AA, Lacombe TT, Doligez AA (2006) Linkage disequilibrium in cultivated grapevine, *Vitis vinifera* L. Theor Appl Genet 112:708–716

Bisognin DA, Douches DS, Buszka L, Bryan G, Wang D (2005) Mapping late blight resistance in *Solanum microdontum* bitter. Crop Sci 45:340–345

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

Bormann CA, Rickert AM, Ruiz RAC, Paal J, Lubeck J, Strahwald J, Buhr K, Gebhardt C (2004) Tagging quantitative trait loci for maturity-corrected late blight resistance in tetraploid potato with PCR-based candidate gene markers. Mol Plant Microbe Interact 17:1126–1138

Bradshaw JE, Pande B, Bryan GJ, Hackett CA, McKlean K, Stewart HE, Waugh R (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, Lowe R, McLean K, Stewart HE, Tierney I, Vilaro MDR, Bryan GJ (2006) Detection of a quantitative trait locus for both foliage and tuber resistance to late blight [*Phytophthora infestans* (Mont.) de Bary] on chromosome 4 of a dihaploid potato clone (*Solanum tuberosum* subsp. tuberosum). Theor Appl Genet 113:943–951

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

Bryan GJ, McLean K, Pande B, Purvis A, Hackett CA, Bradshaw JE, Waugh R (2004) Genetic dissection of H3-mediated polygenic PCN resistance in a heterozygous autotetraploid potato population. Mol Breed 14:105–116

Carlson CS, Eberle MA, Kruglyak L, Nickerson DA (2004) Mapping complex disease loci in whole-genome association studies. Nature 429:446–452

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

Ching A, Caldwell KS, Jung M, Dolan M, Smith OS, Tingey S, Morgante M, Rafalski AJ (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet 3: Art. No. 19 Oct 7 2002

Costanzo S, Simko I, Christ BJ, Haynes KG (2005) QTL analysis of late blight resistance in a diploid potato family of *Solanum phureja* × S-stenotomum. Theor Appl Genet 111:609–617

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

Farnir F, Coppieters W, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Mni M, Nezer C, Simon P, Vanmanshoven P, Wagenaar D, Georges M (2000) Extensive genome-wide linkage disequilibrium in cattle. Genome Res 10:220–227

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

Friedman M (1997) Chemistry, biochemistry, and dietary role of potato polyphenols. A review. J Agric Food Chem 45:1523–1540

Garris AJ, McCouch SR, Kresovich S (2003) Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (*Oryza sativa* L.). Genetics 165:759–769

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, Salamini F (2005) Genomic approaches for the improvement of tuber quality traits in potato. Acta Hortic 684:85–92

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

Hagenblad J, Nordborg M (2002) Sequence variation and haplotype structure surrounding the flowering time locus FRI in *Arabidopsis thaliana*. Genetics 161:289–298

Jacobs JME, Eck HJ, Arens P, Verkerk-Bakker B, Hekkert BL, Bastiaanssen HJM, El-Kharbotly A, Pereira A, Jacobsen E, Stiekema WJ (1995) A genetic map of potato (*Solanum tuberosum*) integrating molecular markers, including transposons, and classical markers. Theor Appl Genet 91: 289–300

Jannoo N, Grivet L, Dookun A, D'Hont A, Glaszmann JC (1999) Linkage disequilibrium among modern sugarcane cultivars. Theor Appl Genet 99:1053–1060

Jorde LB (2000) Linkage disequilibrium and the search for complex disease genes. Genome Res 10:1435–1444

Kloosterman B (2006) Transcriptomic analysis of potato tuber development and tuber quality traits using microarray technology. Dissertation, Wageningen University

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

Kraakman ATW, Martinez F, Mussiraliev B, Van Eeuwijk FA, Niks RE (2006) Linkage disequilibrium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. Mol Breed 17:41–58

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

Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform 5:150–163

Lander ES, Schork NJ (1994) Genetic dissection of complex traits. Science 265:2037–2048

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

Li L, Strahwald J, Hofferbert HR, Lubeck J, Tacke E, Junghans H, Wunder J, Gebhardt C (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

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

Luo ZW, Zhang Z, Leach L, Zhang RM, Bradshaw JE, Kearsey MJ (2006) Constructing genetic linkage maps under a tetrasomic model. Genetics 172:2635–2645

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

McRae AF, McEwan JC, Dodds KG, Wilson T, Crawford AM, Slate J (2002) Linkage disequilibrium in domestic sheep. Genetics 160:1113–1122

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

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9:325–330

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat Genet 30:190–193

Nsengimana J, Baret B, Haley CS, Visscher PM (2004) Linkage disequilibrium in the domesticated pig. Genetics 166:1395–1404

Palaisa KA, Morgante M, Williams M, Rafalski A (2003) Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. Plant Cell 15:1795–1806

Palaisa K, Morgante M, Tingey S, Rafalski A (2004) Long-range patterns of diversity and linkage disequilibrium surrounding the maize Y1 gene are indicative of an asymmetric selective sweep. Proc Natl Acad Sci USA 101:9885–9890

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

Rafalski A, Morgante M (2004) Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. Trends Genet 20:103–111

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98:11479–11484

Semon M, Nielsen R, Jones MP, McCouch SR (2005) The population structure of African cultivated rice *Oryza glaberrima* (Steud.): evidence for elevated levels of linkage disequilibrium caused by admixture with *O. sativa* and ecological adaptation. Genetics 169:1639–1647

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, Jones RW (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

Simko I, Costanzo S, Ramanjulu V, Christ BJ, Haynes KG (2006) Mapping polygenes for tuber resistance to late blight in a diploid *Solanum phureja* × *S. stenotomum* hybrid population. Plant Breed 125:385–389

Skøt LL (2005) An association mapping approach to identify flowering time genes in natural populations of *Lolium perenne* (L.). Mol Breed 15:233–245

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

Tenaillon MI, Sawkins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. mays L.). Proc Natl Acad Sci USA 98:9161–9166

van Berloo R, Hutten RCB, van Eck HJ, Visser RGF (2007) An online potato pedigree database resource. Potato Research, in press

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

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

van Eijk MJ, Broekhof JL, van der Poel HJ, Hogers RC, Schneiders H, Kamerbeek J, Verstege E, van Aart JW, Geerlings H, Buntjer JB, van Oeveren AJ, Vos P (2004) SNPWave: a flexible multiplexed SNP genotyping technology. Nucleic Acids Res 32 (4): Art. No. e47 Feb 2004

van Os H, Andrzejewski S, Bakker E, Barrena I, Bryan GJ, Caromel B, Ghareeb B, Isidore E, De Jong W, Van Koert P, Lefebvre V, Milbourne D, Ritter E, Rouppe Van Der Voort JNAM, Rousselle-Bourgeois F, Van Vliet J, Waugh R, Visser RGF, Bakker J, Van Eck HJ (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

Visker MHPW, Heilersig HJB, Kodde LP, Van De Weg WE, Voorrips RE, Struik PC, Colon LT (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, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

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

Yu J, Pressoir G, Briggs WH, Vroh Bi I, 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 38:203–208

Zhao H, Nettleton D, Soller M, Dekkers JCM (2005) Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between markers and QTL. Genet Res 86:77–87

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. Genetics 163:1123–1134