ORIGINAL PAPER

Association of a CONSTANS-LIKE gene to flowering and height in autotetraploid alfalfa

Doris Herrmann · Philippe Barre · Sylvain Santoni · Bernadette Julier

Abstract In alfalfa (*Medicago sativa*), an autotetraploid forage legume, stem length is a major component of forage yield, quality and competing ability. In this species, flowering date is not a breeding criterion. Association mapping based on a candidate gene approach has given good results in plants, including autotetraploid species for which genetic analyses are complex. The role of a CONSTANS-LIKE gene, identified as a candidate for stem elongation and flowering date in the model legume *M. truncatula*, was tested for association with the same traits in alfalfa. Four hundred genotypes from ten cultivars were evaluated for stem height and flowering date in two locations during 4 years. They were genotyped with simple sequence repeat markers and a low structuration was noticed. Primers were designed to amplify and sequence two regions of the alfalfa gene homologous to CONSTANS-LIKE. Single nucleotide polymorphisms (SNPs) were detected and their allelic dose in each genotype was scored. Linkage disequilibrium within CONSTANS-LIKE rapidly decreased as expected. Eight SNPs with a frequency above 10% were detected over 1,010 bp (one SNP every 126 bp on average) in the 400 genotypes. This number was lower than observed in a neutral gene (a SNP every 31 bp on average). Highly significant associations of three SNPs to flowering date and stem height were identified. Each SNP explained up to 4.2% of the genetic variance. Thus, as in the model species, the CONSTANS-LIKE gene was shown to be involved in flowering date and stem height in alfalfa.

D. Herrmann · P. Barre · B. Julier (✉)
INRA, UR 4, Unité de Recherche Pluridisciplinaire
Prairies et Plantes Fourragères, BP 6, 86600 Lusignan, France
e-mail: bernadette.julier@lusignan.inra.fr

S. Santoni
INRA, UMR Diversité et Adaptation des Plantes Cultivées,
34060 Montpellier, France

Present Address:
D. Herrmann
EPFL, Coopération Indo-Suisse, CM 2 305, Station 10,
1015 Lausanne, Switzerland

Introduction

Although QTL mapping has been demonstrated to be a successful method to dissect genetic bases of agronomic traits in crops (Mansur et al. 1996; Brouwer et al. 2000; Studer et al. 2006), association mapping shows crucial advantages over QTL analyses. As association mapping is based on wide-basis populations, recombination events among loci have occurred over many generations enabling a more precise linkage of marker and target trait (Flint-Garcia et al. 2003). Association mapping is dependent on linkage disequilibrium (LD; the non-random occurrence of alleles at different loci), which is dependent on many factors including mating system, population substructure, selection or genetic drift (Flint-Garcia et al. 2003). In general, a shorter LD decay is observed for outcrossing species (e.g. *Lolium perenne*, Xing et al. 2007) than for selfing species (e.g. *Arabidopsis thaliana*, Nordborg et al. 2002). LD is strongly dependent on the studied population and the genome location (Remington et al. 2001; Auzanneau et al. 2007). In an association study, a genome scan or a candidate gene approach can be applied (Rafalski 2002).

In populations showing a long LD decay, both methods provide a low accuracy of the location of the mutation responsible for the target trait but the genome scan approach gives an exhaustive view of the genome as for QTL analyses. In populations showing a short LD decay, a candidate gene approach provides a high accuracy on the location of the causal mutation. Candidate genes can be selected on the basis of their function and their colocalisation with QTL for the target trait.

In the last decade, the investigation of length LD decay and calculation of association mapping have been performed in diploid plant species such as *Zea mays* (Remington et al. 2001), *L. perenne* (Skøt et al. 2007) and *A. thaliana* (Aranzana et al. 2005). In tetraploid species, little information is available. In potato (*Solanum tuberosum*), an outcrossing autotetraploid species propagated via clones, LD decay assessed on single nucleotide polymorphisms (SNPs) was long (Simko et al. 2006) and associations between quality traits and genes involved in carbohydrate metabolism or transport were reported (Li et al. 2008). An association was also reported between a gene acting in defence signalling and late blight (*Phytophthora infestans*) resistance (Pajerowska-Mukhtar et al. 2009). In alfalfa (*Medicago sativa*), an outcrossing autotetraploid species propagated via seeds, an association between AFLP markers and downy mildew resistance was detected based on two populations differing only for this resistance (Obert et al. 2000).

Alfalfa is the most cultivated forage legume over the world (Michaud et al. 1988). Cultivars are synthetic populations and contain a large genetic diversity for agronomic traits (Julier et al. 2000). Molecular markers have also revealed a wide within-population genetic diversity (Jenczewski et al. 1999; Flajoulot et al. 2005). Synthetic varieties with a high diversity and no genetic structure were described as promising populations for association studies (Auzanneau et al. 2007). In addition, the limited genetic differentiation within and among alfalfa cultivars (Flajoulot et al. 2005; Herrmann et al. 2008) is supposed to lower the risks of spurious association due to the structure of investigated populations whether they are synthetic cultivars or mixtures of cultivars (Flint-Garcia et al. 2003). In alfalfa, forage yield and quality are related to aerial morphogenesis. This set of traits also confers to the plants the ability to compete for light interception in dense canopies composed of alfalfa or alfalfa-grass mixtures. Genetic bases of such traits were studied in *Medicago truncatula*, a model species for legume crops that have a high synteny with alfalfa (Julier et al. 2003; Choi et al. 2004) and for which numerous genomic tools are available. In *M. truncatula*, a CONSTANS-LIKE gene was mapped in the region of a major QTL for flowering date (FD) and other morphological traits (Julier et al. 2007; Pierre et al. 2008). This candidate

gene was differentially expressed in two parental lines with contrasting flowering date (Pierre et al. 2010). The CONSTANS-LIKE family genes represent ideal candidates for association mapping. They were described in detail in *A. thaliana* (Putterill et al. 1995; Robson et al. 2001) and homologues of these genes were reported in other species (Armstead et al. 2005; Hecht et al. 2005). CONSTANS-LIKE promotes flowering in *A. thaliana* (Putterill et al. 1995) and association of members of this gene family to flowering date in several other species (Yuceer et al. 2002; Skøt et al. 2007) confirms the important impact of CONSTANS-LIKE on flowering.

The objective of this study was to evaluate the efficiency of an association mapping strategy based on a candidate gene in autotetraploid outcrossing species with a short LD decay. We tested the role of a candidate gene identified in the model species *M. truncatula*, CONSTANS-LIKE, in the variation for flowering date and stem length in alfalfa.

Materials and methods

Plant material

Ten alfalfa cultivars were investigated, including two landraces (Genetic Resources Center, INRA, Lusignan, France): Provence and Flamande and eight modern varieties: Zenith, Barmed, Luzelle, Mercedes, Alpha, Symphonie, Cannelle and Harpe. Forty germinated seeds per cultivar were established in pots in a greenhouse in 2003 and three clonal replicates of the 400 genotypes were planted in a nursery in April 2004 in two environments: INRA, Lusignan (France) (46.4°N, 0.1°E) and Barenbrug Tourneur Recherches, Connantre (France) (48.7°N, 3.9°E). Space between plants was 0.7 m in Lusignan and 0.5 m in Connantre. Flowering date (date of opening of the first flower) was measured for the first two growth cycles, in May and June, respectively (c1 and c2), each year from 2004 to 2007 in Lusignan (Lus) and transformed into sum of degree-days. Sums of degree-days with a temperature basis of 0°C were measured from January first to measurement date for the first growth cycle and from date of first cut to the measurement date for the second growth cycle. Stem height (SH) was measured for the first two growth cycles in 2004–2007 in Lusignan and 2004–2006 in Connantre (Con) and defined as the height of the tallest stem at the harvest date.

Population structure of this plant material was investigated with 16 simple sequence repeat (SSR) markers: a large diversity was described and, with a global F_{ST} of only 0.012, no structure was observed among the cultivars (Herrmann et al. 2008).

Primer design for the amplification of CONSTANS-LIKE in alfalfa

The sequence of the candidate gene CONSTANS-LIKE of *M. truncatula* is available on http://www.medicago.org, and DNA and mRNA sequences were obtained in the two parental lines of a *M. truncatula* mapping population (Pierre et al. 2010). CONSTANS-LIKE is about 3,700 bp long (GenBank accessions GU080215 and GU080216 for the two parental lines, respectively) and contains four exons (Fig. 1). Primers were designed in the exons and amplification was tested in two alfalfa genotypes. For primer pairs providing a successful amplification, PCR products were sequenced. If the amplified product was larger than 2,000 bp, primers were redesigned in intron region, based on alfalfa sequences, in order to amplify PCR products of about 1,500 bp. Two primer pairs were finally selected (Table 1; Fig. 1). Primer pair *b* (beginning of gene) amplified the main part of exon 1 and the beginning of intron 1. PCR analyses were conducted in a total volume of 50 μl containing about 50 ng DNA, 1× PCR buffer including 1.5 mM MgCl$_2$, 0.2 μM of primers, 0.2 mM of each dNTP and 0.75 U Taq Polymerase (Qiagen). PCRs were performed on MJResearch PTC100 with the following conditions: 4 min at 94°C, 8 cycles of 1 min at 94°C, 1 min at 63°C (decreasing the temperature by 1°C after each cycle) and 1 min 30 s at 72°C, 27 cycles of 1 min at 94°C, 1 min at 56°C and 1 min 30 s at 72°C followed by a final extension of 5 min at 72°C. Two fragments of about 1,300 and 1,700 bp were detected on agarose gel. Sequencing with the forward primer (*b*F) revealed a clear sequence of about 500 bp after which indels in at least one of the four alleles generated an overlaying of several sequences (analogous to F and G in Figure S1, Electronic supplementary material). Primer pair *e* (end of gene) amplified exons 2–4 and introns 2 and 3. A touch-down

Table 1 Primers used for PCR amplification of regions *b* and *e* in CONSTANS-LIKE gene

Name	Primer sequence
Primer *b* forward (*b*F)	GCCAAAACTGCAAAAACG
Primer *b* reverse (*b*R)	CCAAATTTTCATGTCAATGCT
Primer *e* forward (*e*F)	TCTGGGATTTTYAAYTAYAG
Primer *e* reverse (*e*R)	CTCTCACTCTCTTTCTCG

PCR program was used as for primer pair *b* except that annealing temperature decreased from 58 to 50°C in the first eight cycles. Two fragments of about 1,500 and 1,700 bp were detected in agarose gel. Sequencing with forward primer revealed a clear sequence of about 500 bp before the overlaying of several alleles in the sequences was observed.

Mapping of the candidate gene

In order to be sure that the CONSTANS-LIKE gene amplified in alfalfa was the homologous gene of the one in *M. truncatula*, we mapped it in alfalfa. One hundred and twenty genotypes of the PERMED mapping population (Julier et al. 2008) were amplified using primer pair *e* according to the corresponding PCR protocol described above except that reaction was performed in a final volume of 10 μl. PCR products were separated on 1% agarose gels. The fragment of about 1,700 bp was monomorphic and the segregating fragment of about 1,500 bp was scored for the presence and absence. This marker was integrated in the genetic map using the TetraploidMap software (Hackett and Luo 2003).

Cloning and calculation of LD based on haplotypes

Forty genotypes of the variety Mercedes used in the association study were amplified using primer pair *e* according to the corresponding PCR protocol described above except that the elongation temperature was 68°C instead of 72°C, reactions were performed in a final volume of 25 μl and with 1 U of Platinum High Fidelity Taq and its 1× PCR buffer containing 2 mM MgSO$_4$ (Invitrogen). For each genotype, PCR product was purified using the MinElute PCR purification kit (Qiagen) and 3 μl of eluted product was cloned using the pGem-T easy Kit (Promega). Eight white colonies of each genotype were selected and directly amplified. PCR product was separated on a 1% agarose gel. If the PCR products of the eight colonies had the same length, one colony was selected for sequencing. For 20 genotypes, two different fragment lengths were observed in the eight colonies of a single genotype. In that case, two colonies with different fragment lengths were selected for sequencing. Overall, based on 39
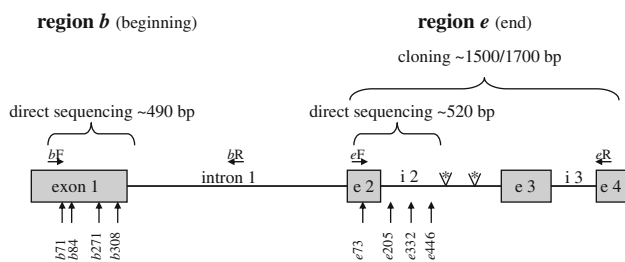

region *b* (beginning)   **region *e*** (end)

cloning ~1500/1700 bp

direct sequencing ~490 bp   direct sequencing ~520 bp

*b*F   *b*R   *e*F   *e*R

| exon 1 | intron 1 | e 2 | i 2 | e 3 | i 3 | e 4 |

b71 b84 b271 b308   e73 e205 e332 e446

Fig. 1 Structure of the CONSTANS-LIKE gene (∼3,700 bp) (GenBank accessions GU080215 and GU080216). Parts of the gene, for which sequences were screened for polymorphisms based on cloning (59 alleles of 39 genotypes) or direct sequencing (400 genotypes), are indicated with *curly brackets*. *Horizontal arrows* indicate selected primers for amplification of *regions b* and *e*. *Vertical arrows* indicate location of SNPs (allelic frequency higher than 0.1) detected on 400 genotypes and direct sequencing. *Asterisks* indicate major inserts of about 200 bp

genotypes, 59 alleles were sequenced in both directions. Sequencing was performed on PCR products purified using the AMPure® kit (Agencourt®, Massachusets, USA); BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied BioSystem, California, USA) was used according to the standard protocol and the reactions were purified with the CleanSEQ® kit (Agencourt) and separated on an 3130xl Genetic Analyzer (Applied BioSystem). The sequencing primers were the same as the ones used for PCR amplification. Sequences were aligned using Staden (http://staden.sourceforge.net/) and forward and reverse sequences were manually combined.

Polymorphisms were detected manually in the chromatograms. Calculation of LD based on haplotypes as a function of squared correlations of allelic frequencies (r^2) was performed using DnaSP (Rozas et al. 2003) and TASSEL (http://www.maizegenetics.net/) software. P values for r^2 were adjusted using a Bonferroni correction.

Direct sequencing, calculation of LD based on genotypes and association analyses

The 400 genotypes of the 10 cultivars were amplified using the two primer pairs b and e and the corresponding PCR protocol described above. PCR products were sequenced with direct sequencing in the forward direction as above. Sequences were aligned using Staden (http://staden.source forge.net/) and chromatograms were manually screened for polymorphisms. In autotetraploid plants, the maximum of information on SNP polymorphism is given by the estimation of allelic doses. Polymorphisms were scored for the presence or absence. In addition, the dose of the alternative base (or indel) was estimated and resulted in five classes of genotype for each polymorphism (Figure S1, Electronic supplementary material). If a genotype showed a clear indel whose length could unambiguously be identified, this genotype was kept for scoring and indel was considered for coding the following SNPs (F and G in Figure S1, Electronic supplementary material). If a genotype showed an unclear indel, of which length could not be identified, genotype was no longer scored and missing values were noted for the following SNPs. Scoring was stopped if the majority (>50%) of the genotypes showed an indel in at least one of the four alleles.

For each SNP, the frequency of the alternative base over 400 genotypes (genotypic frequency) and frequency over the 1,600 alleles with alternative base (allelic frequency) were calculated. To calculate LD based on genotypes, the alternative base was scored for the presence or absence of SNPs occurring with an allelic frequency higher than 0.1. For each pair of SNPs, LD was tested: a 2 × 2 contingency table was built as described by Julier (2009) and a Fisher exact test was computed using the FREQ procedure of SAS

statistical software package (version 8.1, SAS Institute Inc. 2000). Resulting probabilities for LD were divided by the number of pairs of SNPs to apply a Bonferroni correction. Distances between SNPs within regions b and e, respectively, were given by alfalfa sequences. Distance between SNPs between regions b and e was based on the sequence of CONSTANS-LIKE gene in *M. truncatula*.

Association calculations were performed using the SNPs with an allelic frequency higher than 0.1. For flowering date and stem height, each measurement was considered and means over all growth cycles were calculated (FDglobal and SHglobal, respectively). Analyses were based on two different statistical methods. Firstly, a one-way analysis of variance was calculated with each SNP as factor and measurement as variable. In this additive model (Gallais 2003), the factor SNP had up to five levels (absence or presence in a dose varying from 1 to 4). Secondly, for each measurement and all SNPs, a stepwise regression was performed using the REG procedure (option stepwise selection, $P < 0.05$ for entry into the model) of the SAS statistical software package (SAS Institute Inc. 2000). Average values and standard deviations of genotypic classes were calculated. Differences among average values were tested for significance using the Newman–Keuls test in Statistica software (version 8.0, StatSoft Inc. 2007).

In order to evaluate population differentiation for CONSTANS-LIKE, for SNPs with an allelic frequency higher than 0.1, the fixation index F_{ST} was determined over all ten cultivars and for pairs of cultivars using Gene4x software (Ronfort et al. 1998).

Amplification of a glutamate-synthase gene

In order to compare the sequence polymorphism of CONSTANS-LIKE to that of a neutral gene, a gene coding for glutamate-synthase and described to be neutral (Muller et al. 2006) was sequenced. It was amplified by PCR on a subset of 50 genotypes (4–6 genotypes per cultivar) using the protocol described in Muller et al. PCR products were sequenced with direct sequencing in the reverse direction (see protocols above). Sequences were aligned using Staden software (http://staden.sourceforge.net/) and chromatograms were screened for polymorphisms based on the coding system described above. Sequences included a part of intron 7, exon 8 and a part of intron 8 (Vance et al. 1995).

Results

Structure of CONSTANS-LIKE in alfalfa

Regions b and e of the CONSTANS-LIKE gene were amplified and PCR products of the 400 genotypes were

directly sequenced in forward direction (Fig. 1). For region *b*, the first 490 bp of the sequences was unambiguously analysed, then an overlaying of alleles in the sequences was observed for all genotypes. In these 490 bp and the 400 genotypes, only one insert of 1 bp occurring in one genotype was identified. For region *e*, 520 bp of the sequences was integrated in analyses since a major indel which was detected at least once in more than 80% of the genotypes made unambiguous interpretation of a longer sequence impossible. Cloning confirmed this major insert at a distance of about 560 bp to the forward primer. This indel showed a complex structure. Ten alleles out of 59 cloned alleles showed an insert of 227 bp and four alleles showed the same insert in reverse direction. For the other 45 alleles, a deletion varying between 227 and 243 bp was observed. A second major insert, at a distance of about 980 bp to the forward primer, was detected in 11 cloned alleles (Fig. 1). Fourteen indels were observed in the 400 sequences within the 520 bp of region *e*. The length of four indels, which occurred in the last third of the sequence, was not identifiable and missing values were noted henceforward for the 15 affected genotypes. Except for one indel (14 bp), other indels were short (1–3 bp) and genotypes were kept for scoring while considering the indel in the coding sequence (F and G in Figure S1, Electronic supplementary material). All these 14 indels showed a genotypic frequency below 0.07 and in total only 71 genotypes were affected.

The sequence similarities of *b* and *e* regions with other species were checked. For the consensus sequence of the 490 bp in region *b* and for the corresponding sequence of amino acids, 96% similarity was observed with the CONSTANS-LIKE gene of *M. truncatula* (http://blast.ncbi.nlm. nih.gov/Blast.cgi). Blast with the *A. thaliana* genome identified an *E* value of 8E-5 for the zinc finger protein CONSTANS-LIKE 14 (O22800, COL14_ARATH; http://www.uniprot.org). For the consensus sequence of the 520 bp scored in region *e*, a 93% similarity was observed with the corresponding part of the gene in *M. truncatula*.

The marker defined by the amplification of region *e* segregated in proportion of 3:1. As expected, it was mapped at the bottom of chromosome 7 in an alfalfa map at a position of 76 and 52 cM from the top of the map of each parent, respectively (data not shown).

SNPs in sequenced regions

For the two sequences of 490 and 520 bp in regions *b* and *e*, 15 SNPs with a genotypic frequency (calculated over the 400 genotypes) higher than 0.1 were detected and 8 of them showed an allelic frequency (calculated over the 1,600 alleles carried by the 400 genotypes) higher than 0.1 (Table 2). These eight SNPs were detected on average every 126 bp and showed an average allelic frequency of 0.22. Over the 15 SNPs with a genotypic frequency higher than 0.1, 11 were located in exons. Four of them did not change amino acid sequence (Table 2). In region *e*, 59 alleles of 39 genotypes of the variety Mercedes were cloned and sequenced in both directions. Twenty-one SNPs were detected in more than one allele, of which 12 SNPs

Table 2 SNPs with a genotypic frequency higher than 0.1 detected by direct sequencing in regions *b* and *e* of CONSTANS-LIKE in 400 genotypes (1,600 alleles)

Name	Location	Abundant base	Alternative base	Amino acid change	Genotypic frequency	Allelic frequency
*b*71	exon 1	C	T	P → S	0.40	0.13
*b*84	exon 1	C	T	S → F	0.89	0.47
*b*92	exon 1	C	A	H → N	0.22	0.06
*b*156	exon 1	C	T	T → I	0.25	0.08
*b*193	exon 1	C	T	stay F	0.20	0.06
*b*270	exon 1	C	T	A → V	0.10	0.02
*b*271	exon 1	G	T	stay A	0.45	0.16
*b*308	exon 1	C	T	stay L	0.67	0.27
*b*318	exon 1	A	G	N → S	0.27	0.08
*e*73	exon 2	T	C	Y → H	0.47	0.17
*e*114	exon 2	A	G	stay S	0.19	0.08
*e*135	intron 2	G	A	–	0.11	0.04
*e*205	intron 2	G	T	–	0.44	0.13
*e*332	intron 2	A	T	–	0.76	0.32
*e*446	intron 2	G	T	–	0.27	0.11

Genotypic frequency frequency of the genotypes, over the 400 genotypes, in which at least one of the four alleles was the alternative base; *allelic frequency* frequency of the alleles, over the 1,600 alleles, with the alternative base; numbers in names indicate SNP positions; SNPs in bold have allelic frequency higher than 0.10

were detected with an allelic frequency higher than 0.1 (Table 3). For the part of the region *e* that was sequenced after cloning and directly sequenced, the same seven SNPs were detected. However, *e*135 was only detected in the cloned alleles as a singleton and *e*354 was only detected in 9 of the 400 genotypes based on direct sequencing (Tables 2, 3).

For the 220 bp of the glutamate-synthase gene directly sequenced in a subset of 50 genotypes, 7 SNPs with an allelic frequency higher than 0.1 were detected, 2 in intron 8, 1 in exon 8 and 4 in intron 7. On average, a SNP was observed every 31 bp and the SNPs showed an allelic frequency of 0.41 on average.

Linkage disequilibrium

Linkage disequilibrium was estimated using the 12 SNPs with an allelic frequency higher than 0.1, determined on 59 cloned alleles of the variety Mercedes (Table 3). For the calculations of distance between the SNPs, the two major inserts (each ∼200 bp) were not taken into account, as majority of haplotypes showed neither of these inserts. Fifty-eight of the 66 pairwise comparisons showed squared correlations of allelic frequencies (r^2) lower than 0.35 and

Table 3 SNPs detected in more than one genotype after cloning in CONSTANS-LIKE gene in 59 alleles of 39 genotypes

Name	Frequency	Location
*e*16	0.19	exon 2
*e*73	0.17	exon 2
*e*114	0.08	exon 2
*e*205	0.07	intron 2
*e*332	0.46	intron 2
*e*354	0.03	intron 2
*e*446	0.15	intron 2
*e*504	0.08	intron 2
*e*872	0.03	intron 2
*e*923	0.27	intron 2
*e*937	0.10	intron 2
*e*975	0.19	intron 2
*e*1170	0.19	intron 2
*e*1344	0.03	intron 2
*e*1371	0.07	intron 2
*e*1701	0.34	intron 3
*e*1740	0.10	intron 3
*e*1829	0.08	intron 3
*e*1830	0.08	intron 3
*e*1935	0.44	exon 4
*e*1948	0.12	exon 4

Frequency frequency of the alleles, over 59 alleles, with the alternative base; numbers in names indicate SNP position

all pairs with a distance larger than 700 bp had a r^2 lower than 0.25 (Fig. 2). Thirty-one pairs showed significant ($P < 0.05$) linkage disequilibrium (r^2), of which 11 remained significant after a Bonferroni correction. Only one of the significant pairs (e332/e1935) showed a distance longer than 700 bp.

For calculations of LD based on the 400 genotypes, the eight SNPs with an allelic frequency higher than 0.1 were integrated (Table 2). Six of the 28 pairs of SNPs showed significant LD after a Bonferroni correction, with four pairs showing a distance between SNPs longer than 1,000 bp (Fig. 3). A particular low value of P (8E-90) was observed between *b*271 and *e*205.

Association to flowering date and stem height

An association between SNPs *b*271 and *e*205 and flowering date of the second growth cycle at Lusignan in 2005 (FDc2Lus05) was detected in analysis of variance. By stepwise regression, a single SNP, *b*271, was associated to this trait. Four measurements for stem height (SHc2Lus04, SHc2Lus05, SHc1Con04, SHglobal) showed associations to these two SNPs (Table 4). For SNP *b*271, genotypes that were monomorphic for the abundant base (GGGG) significantly showed later flowering date and lower stem height compared to the genotypes that had three doses of the alternative base (GTTT; Table 5). In addition, an association between *b*71 and several measurements of flowering date and stem height was observed in analysis of variance, which was confirmed for SHc2Lus04 in stepwise regression (Table 4). For SHc2Lus04, genotypes that were monomorphic for the abundant base (CCCC) showed significantly higher stem height than the genotypes with three
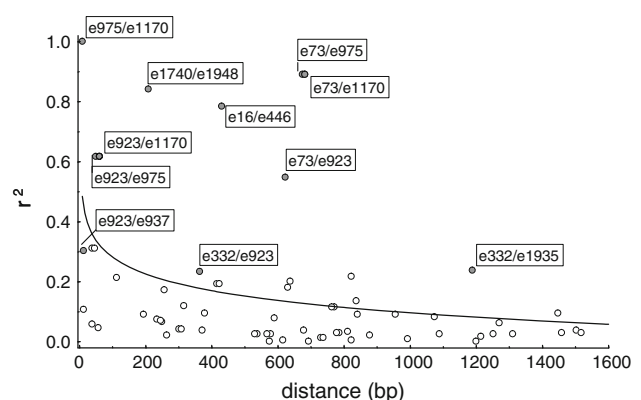

Fig. 2 Decay of linkage disequilibrium (LD) based on haplotypes: squared correlations of allelic frequencies (r^2) as a function of distance in bp. LD was calculated with pairs of 12 SNPs with an allelic frequency higher than 0.1 based on 59 cloned alleles of 39 genotypes. *Grey dots* are used to show significant ($P < 0.05$) r^2 after Bonferroni correction for the pairs of SNP indicated in *boxes* and *curve* shows logarithmic regression
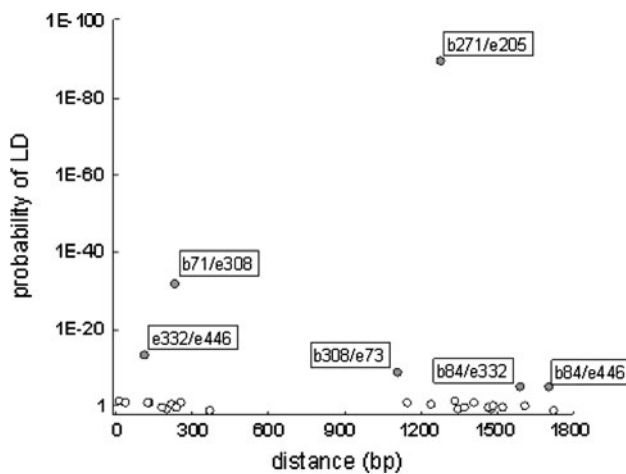
Fig. 3 Decay of linkage disequilibrium (LD) based on genotypes: probability of linkage equilibrium (P) as a function of distance in bp. LD was calculated with pairs of eight SNPs of an allelic frequency higher than 0.1 based on 400 genotypes. *Grey dots* are used to show significant ($P < 0.05$) r^2 after Bonferroni correction for the pairs of SNP indicated in *boxes*

doses of the alternative base (CTTT; Table 5). Interestingly, the plants of genotype GTTT for SNP *b271* were different from the plants of genotype CTTT for SNP *b71*.

Structure of cultivars

Global F_{ST} among the ten cultivars based on the eight SNPs of CONSTANS-LIKE with an allelic frequency higher than 0.1 was 0.0043 and 80% of F_{ST} between pairs of cultivars were not significantly ($P > 0.05$) different.

Discussion

Direct sequencing or sequencing after cloning?

In autotetraploid and heterozygous species, recovery of all four haplotype sequences and SNP doses face several constraints. Ideally, the sequencing of 20 clones of a gene fragment in one genotype gives access to both the haplotype sequence and the dose of each allele with a probability of 95%. But usually, all alleles are not equally cloned. Moreover, this method is expensive and time consuming, so direct sequencing of PCR products could be a solution. Direct sequencing has the major advantage that the information of four alleles is available in one sequence and several SNPs can be genotyped at the same time. But indels present in heterozygous state impede the reading of long sequences. In potato (Pajerowska-Mukhtar et al. 2009), the presence of indels was mentioned for only 3 of the 22 studied genes, which enabled the reading of long sequences (up to 900 bp) in 19 genes. In CONSTANS-LIKE gene of

alfalfa, direct sequencing was only possible for about 500 bp. Interpretation of chromatograms for tetraploid species is more difficult than for diploid ones, as the dose of alleles has to be estimated for each polymorphic site. Due to indels, manual interpretation of the chromatograms using direct sequencing was labour intensive, but we succeeded to score the dose of alleles (Figure S1, Electronic supplementary material). The cloned alleles confirmed the scoring of the SNPs by comparing SNPs detected with the two methods in the same genomic region. Furthermore, cloning allowed a selection of a region for direct sequencing with few insertions and a maximum number of SNPs. However, no software exists for haplotype reconstruction from tetraploid genotypes as for diploid ones (PHASE, Stephens and Donnelly 2003). LD calculations are therefore based on genotypes and not on haplotypes, which leads to a severe loss of information. For sequences based on cloned alleles, the haplotypes are known. A combination of the two methods allowed us to maximise information content with given financial resources. With the cloning of the alleles in one cultivar, we were able to calculate a reliable estimation of LD based on haplotypes. With direct sequencing of the two regions, we obtained the information on 1,600 alleles over a sequence length of 1,010 bp with a reasonable expense. In conclusion, direct sequencing worked but new methods should be developed in order to optimise the identification of SNPs (high throughput resequencing) and to score their presence and dose (SNPlex, Veracode).

Validation of CONSTANS-LIKE gene and determination of LD

Considering the genetic characteristics (allogamy and autotetraploidy) of alfalfa, an association mapping strategy using a candidate gene approach seemed appropriate. However, an unambiguous amplification of the candidate gene in the target species had to be validated and a short LD decay in the candidate gene had to be confirmed.

The amplification of the candidate gene in the target species had to be assured in order to avoid an amplification of a member of the same gene family in another part of the genome. This is particularly important for a CONSTANS-LIKE gene, for which numerous members of the gene family were reported for different species (Robson et al. 2001; Hecht et al. 2005; Griffiths et al. 2003). The high similarity of the amplified sequences in alfalfa to the sequence of CONSTANS-LIKE in *M. truncatula* and the mapping of the amplified parts on the expected chromosome proved a correct amplification.

LD based on haplotypes decayed shortly and after 700 bp only one pair of SNPs was significant. However, LD decay probably occurred at a shorter distance than

Table 4 Mean values and standard deviation (SD) for flowering dates (in °C.D) and stem height (in cm), and significant association of SNPs with allelic frequency higher than 0.1 to flowering date and stem height based on 400 genotypes

SNP:	FDc1Lus04	FDc1Lus05	FDc1Lus06	FDc1Lus07	FDc2Lus04	FDc2Lus05	FDc2Lus06	FDc2Lus07	FDglobal	SHc1Con04	SHc1Con05	SHc1Con06
Mean:	1,519	1,159	930	1,260	661	363	376	535	836	65.5	52.0	74.3
SD:	71	52	18	77	111	45	39	91	38	8.8	7.2	6.0
Analysis of variance												
b71	–	2.2*	–	–	2.4*	–	–	–	–	–	–	–
b84	–	–	–	–	–	–	–	–	–	–	–	–
b271	–	–	–	–	–	3.9**	–	–	–	–	–	–
b308	–	–	–	–	–	–	–	–	–	–	–	–
e73	–	–	–	–	4.2**	–	–	–	–	–	–	–
e205	–	–	–	–	–	4.1**	–	–	–	2.8*	–	–
e332	–	–	–	–	–	–	–	–	–	–	–	–
e446	–	–	–	–	–	–	–	–	–	–	–	–
Stepwise regression												
b71	–	–	–	–	–	–	–	–	–	–	–	–
b84	–	–	–	–	–	–	–	–	–	–	–	–
b271	–	–	–	–	–	3.6***	–	–	–	–	–	–
b308	–	–	–	–	–	–	–	–	–	–	–	–
e73	–	–	–	–	2.5**	–	–	–	1.2*	–	–	–
e205	–	–	–	–	–	–	–	–	–	–	–	–
e332	–	–	–	–	–	–	–	–	–	–	–	–
e446	–	–	–	–	–	–	–	–	–	–	–	–

SNP:	SHc1Lus04	SHc1Lus05	SHc1Lus06	SHc1Lus07	SHc2Con04	SHc2Con05	SHc2Con06	SHc2Lus04	SHc2Lus05	SHc2Lus06	SHc2Lus07	SHglobal
Mean:	85.8	97.7	75.6	74.8	42.4	83.3	73.2	78.9	55.8	91.3	79.1	73.6
SD:	8.0	11.2	8.4	9.0	6.0	8.7	7.0	9.6	5.4	6.8	8.8	5.41
Analysis of variance												
b71	–	–	–	–	–	–	–	3.0**	2.1*	–	–	–
b84	–	–	–	–	–	–	–	–	–	–	–	–
b271	–	–	–	–	–	–	–	2.7*	3.8**	–	–	–
b308	–	–	–	–	–	–	2.6*	4.0**	–	–	–	–
e73	–	–	–	–	–	–	–	3.1*	–	–	–	–
e205	–	–	–	–	–	–	–	3.8**	2.8*	–	–	2.6*
e332	–	–	–	–	–	–	–	–	–	–	–	–
e446	–	–	–	2.9*	–	–	–	–	–	–	–	–
Stepwise regression												
b71	–	–	–	–	–	–	–	2.0**	–	–	1.8*	–
b84	–	–	–	–	–	1.1*	–	–	–	1.1*	–	–

Table 4 continued

SNP:	SHc1Lus04	SHc1Lus05	SHc1Lus06	SHc1Lus07	SHc2Con04	SHc2Con05	SHc2Con06	SHc2Lus04	SHc2Lus05	SHc2Lus06	SHc2Lus07	SHglobal
Mean:	85.8	97.7	75.6	74.8	42.4	83.3	73.2	78.9	55.8	91.3	79.1	73.6
SD:	8.0	11.2	8.4	9.0	6.0	8.7	7.0	9.6	5.4	6.8	8.8	5.41
b271	–	–	–	–	–	–	–	–	–	–	–	–
b308	–	–	–	–	–	–	–	–	–	–	–	–
e73	–	–	–	–	–	–	–	–	–	–	–	–
e205	–	–	–	–	–	–	–	1.4*	2.0**	–	–	–
e332	–	–	–	–	–	–	–	–	–	–	1.2*	–
e446	–	–	–	–	1.1*	–	–	–	–	1.5*	–	–

Numbers indicate percentage of variance explained and are based on 400 genotypes and on one-way ANOVA or stepwise regression, respectively

FD flowering date, *SH* stem height, *c1/2* cycle 1/2, *Lus* Lusignan, *Con* Connantre, *global* average on 8 and 14 measurements of flowering and height, respectively

* $P < 0.05$

** $P < 0.01$

*** $P < 0.001$

700 bp, as significant r^2 values were rare. A distance of LD decay below 1,000 bp was confirmed by LD calculations based on genotypes, although 4 of the 16 pairs of SNPs with a distance longer than 1,000 bp showed significant P values. As calculations were based on genotypes, P values were probably underestimated and P values of haplotypes should have been higher. However, a longer LD decay cannot be excluded as calculations were limited to distances below 1,800 bp and only 66 and 28 pairs of SNPs were integrated, for the haplotypes and the genotypes, respectively. The observed short distance for LD decay was comparable to distances reported in the outcrossing species *L. perenne* (Xing et al. 2007) and even lower than in the outcrossing species maize (Remington et al. 2001). However, the genetic region or the diversity of populations also plays a role (Auzanneau et al. 2007). This last reason may explain the difference of the LD observed between alfalfa in this study and autotetraploid potato, for which the longer LD decay was explained by the clonal reproduction system and the resulting low number of meiotic generations between the populations (Simko et al. 2006). Finally, the unambiguous amplification of the candidate gene in alfalfa and the short LD decay made a candidate gene approach for an association study appropriate.

Association of CONSTANS-LIKE gene to flowering date and stem height

Compared to glutamate-synthase gene, which is expected to be neutral, a lower number of SNPs was detected in CONSTANS-LIKE gene and the SNPs had a lower allelic frequency. Therefore, CONSTANS-LIKE has been subjected to strong selective pressure against mutation and probably has an important function. Similarly in *L. perenne*, the number of polymorphisms in a CONSTANS-LIKE homologue was low compared to another candidate gene (Skøt et al. 2007). This low sequence polymorphism made the detection of association to morphological traits more difficult. Furthermore, a strong interaction between genotypes and environment was observed for morphological traits resulting in a low heritability of the traits (Herrmann et al., unpublished), which may further hinder the detection of associations.

However, we observed a highly significant link between flowering date in 1 year and location and SNPs in CONSTANS-LIKE, with significantly different mean values of genotypic classes. This association seems to confirm the impact of CONSTANS-LIKE in the flowering pathway for alfalfa, which was reported in several other species (Putterill et al. 1995; Robson et al. 2001; Yuceer et al. 2002; Skøt et al. 2007). Associations were also found between SNPs in CONSTANS-LIKE and stem length, a trait that was not genetically correlated to flowering date in

Table 5 Average phenotypic values for genotypic classes of SNPs calculated over 400 genotypes

Genotype	FDc2Lus05			SHc2Lus05			SHc2Lus04		
	N	Average	SD	N	Average	SD	N	Average	SD
Total	388	362.72	44.50	389	55.828	5.421	388	78.855	9.605
SNP *b271*									
GTTT	7	325.06[a]	43.77	7	59.548[a]	3.173	7	85.238[a]	7.783
GGTT	57	346.73[ab]	46.27	57	55.692[b]	5.007	57	79.858[ab]	10.302
GGGT	110	364.10[b]	40.34	111	57.186[ab]	5.650	110	80.401[ab]	9.126
GGGG	214	367.50[b]	44.82	214	55.038[b]	5.302	214	77.584[b]	9.535
SNP *b71*									
CTTT	9	350.83[a]	49.76	9	58.259[a]	5.537	9	72.855[a]	10.624
CCTT	31	361.08[a]	45.68	31	58.016[a]	6.288	30	74.767[ab]	11.303
CCCT	115	371.87[a]	43.74	115	55.272[a]	5.015	115	78.456[ab]	8.988
CCCC	233	358.88[a]	44.10	234	55.718[a]	5.427	234	79.805[b]	9.455

N number of genotypes, *average* average values, *SD* standard deviation; *FDc2Lus05/SHc2Lus05/SHc2Lus04* measurements (for detailed description see Table 4)

Significant ($P < 0.5$) differences among average values are indicated with different letter

alfalfa. In the genomic region where CONSTANS-LIKE was mapped in *M. truncatula*, QTLs were observed for both flowering date and branch length (Julier et al. 2007). In addition, a CONSTANS-LIKE gene of *A. thaliana* showed an effect on stem height in *A. thaliana* and in potato which was transformed with this gene (Simon et al. 1996; Martinez-Garcia et al. 2002). It was proposed that CONSTANS-LIKE gene affects flowering date and stem elongation through distinct mechanisms (Gonzalez-Schain and Suarez-Lopez 2008). The associations of SNPs in CONSTANS-LIKE to stem height in different years were therefore not spurious associations, but could be explained by the additional role of CONSTANS-LIKE. These additional associations confirmed the functionality of this CONSTANS-LIKE gene in alfalfa.

The two SNPs *b271* and *e205* showed a strong LD based on genotypes and therefore only one or the other was kept in stepwise regression. Considering SNPs with a frequency lower than 10%, other strong LDs were observed between SNPs separated by a long distance (not shown). This may suggest that a haplotype consisting of *b271*, *e205* and probably of other SNPs was associated with the traits rather than a single SNP. Such an association was observed in other species (Skøt et al. 2007).

There was no evidence that *b271* (synonymous exchange of amino acid) or *e205* (in intron) was functional. In addition, they were not located in the regions of the B-boxes or in the CTT domain, which were described to be functionally important in CONSTANS-LIKE gene (Putterill et al. 1995; Robson et al. 2001). In contrast, *b71* was situated in the second B-box and *e1948* was situated at the beginning of the CTT domain and both showed a non-synonymous exchange of amino acids (*b71*: P/S, *e1948*: T/A)

and were therefore candidates for a functional SNP. However, only *b71* was genotyped on all 400 genotypes and showed a significant association to stem height.

No structure among the ten cultivars based on the SNPs in CONSTANS-LIKE was observed, so the significant difference of height among the cultivars was probably based on other genes. Polymorphism in CONSTANS-LIKE gene could be used to generate an additional source of variation in stem height in alfalfa breeding. It could be interesting to create genotypes that are monomorphic for the alternative bases of the significant SNPs, as these genotypes were not available in this study. Comparing flowering date and stem height of these genotypes with those of genotypes that are monomorphic for the abundant base would complete the information on allele effects.

In this study, which was performed on a population with a short LD decay, the role of CONSTANS-LIKE gene in flowering date and stem height is demonstrated. In contrast, in a species such as potato with a long LD decay, the association between a candidate gene and a trait can be indirect, the loci showing association with a trait being physically linked to functional genes (Pajerowska-Mukhtar et al. 2009).

In conclusion, this is, to the authors' knowledge, the first association study based on SNPs identified in a candidate gene in an autotetraploid species with a short LD decay. Scoring of SNPs with the allelic doses after direct sequencing was possible. Unambiguous amplification of a candidate gene of *M. truncatula* in alfalfa was obtained. LD decay was, as expected, short and similar to that calculated with other outcrossing species. Association study based on SNPs in a candidate gene is therefore a very promising approach for future projects in tetraploid species

with a short LD decay in general and in particular in alfalfa, for which the sequences of numerous candidate genes are known in the model species *M. truncatula*. The use of larger populations could be more efficient to obtain enough individuals in all five genotypic classes at each SNP. The candidate gene CONSTANS-LIKE was subject to strong selection pressure and therefore functionally important in alfalfa but only a limited number of SNPs was detectable. However, a weak but significant association between three SNPs of the CONSTANS-LIKE gene and the flowering date and height was demonstrated.

References

Aranzana MJ, Kim S, Zhao KY, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang CL, Toomajian C, Traw B, Zheng HG, Bergelson J, Dean C, Marjoram P, Nordborg M (2005) Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. PLoS Genet 1:531–539

Armstead IP, Skøt L, Turner LB, Skøt K, Donnison IS, Humphreys MO, King IP (2005) Identification of perennial ryegrass (*Lolium perenne* (L.)) and meadow fescue (*Festuca pratensis* (Huds.)) candidate orthologous sequences to the rice Hd1(Se1) and barley HvCO1 CONSTANS-like genes through comparative mapping and microsynteny. New Phytol 167:239–247

Auzanneau J, Huyghe C, Julier B, Barre P (2007) Linkage disequilibrium in synthetic varieties of perennial ryegrass. Theor Appl Genet 115:837–847

Brouwer DJ, Duke SH, Osborn TC (2000) Mapping genetic factors associated with winter hardiness, fail growth, and freezing injury in autotetraploid alfalfa. Crop Sci 40:1387–1396

Choi HK, Kim D, Uhm T, Limpens E, Lim H, Mun JH, Kalo P, Penmetsa RV, Seres A, Kulikova O, Roe BA, Bisseling T, Kiss GB, Cook DR (2004) A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. Genetics 166:1463–1502

Flajoulot S, Ronfort J, Baudouin P, Barre P, Huguet T, Huyghe C, Julier B (2005) Genetic diversity among alfalfa (*Medicago sativa*) cultivars coming from a single breeding program, using SSR markers. Theor Appl Genet 111:1420–1429

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

Gallais A (2003) Quantitative genetics and breeding methods in autopolyploid plants. INRA Editions, Paris, France

Gonzalez-Schain ND, Suarez-Lopez P (2008) CONSTANS delays flowering and affects tuber yield in potato. Biol Plant 52:251–258

Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis. Plant Physiol 131:1855–1867

Hackett CA, Luo ZW (2003) TetraploidMap: construction of a linkage map in autotetraploid species. J Hered 94:358–359

Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltran JP, Rameau C, Weller JL (2005) Conservation of *Arabidopsis* flowering genes in model legumes. Plant Physiol 137:1420–1434

Herrmann D, Flajoulot S, Barre P, Huyghe C, Ronfort J, Julier B (2008) Comparison of morphological traits and SSR markers to analyze genetic diversity of alfalfa cultivars. Joint meeting of the 41st North American Alfalfa Improvement Conference and 20th Trifolium Conference, Dallas

Jenczewski E, Prosperi JM, Ronfort J (1999) Differentiation between natural and cultivated populations of *Medicago sativa* (*Leguminosae*) from Spain: analysis with random amplified polymorphic DNA (RAPD) markers and comparison to allozymes. Mol Ecol 8:1317–1330

Julier B (2009) A program to test linkage disequilibrium between loci in autotetraploid species. Mol Ecol Resour 9:746–748

Julier B, Huyghe C, Ecalle C (2000) Within- and among-cultivar genetic variation in alfalfa: forage quality, morphology, and yield. Crop Sci 40:365–369

Julier B, Flajoulot S, Barre P, Cardinet G, Santoni S, Huguet T, Huyghe C (2003) Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. BMC Plant Biol 3:9

Julier B, Huguet T, Chardon F, Ayadi R, Pierre JB, Prosperi JM, Barre P, Huyghe C (2007) Identification of quantitative trait loci influencing aerial morphogenesis in the model legume *Medicago truncatula*. Theor Appl Genet 114:1391–1406

Julier B, Bernard K, Abdelguerfi LM, Udupa SM, Semiani Y, Sfarni O, Elboutahiri N, Gasmi F, Barre B, Huguet T, Lelièvre F (2008) QTL for water stress resistance and water use efficiency in alfalfa. Joint meeting of the 41st North American Alfalfa Improvement Conference and 20th Trifolium Conference, Dallas

Li L, Paulo MJ, Strahwald J, Lubeck J, Hofferbert HR, Tacke E, Junghans H, Wunder J, Draffehn A, van Eeuwijk F, Gebhardt C (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

Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic mapping of agronomic traits using recombinant inbred lines of soybean. Crop Sci 36:1327–1336

Martinez-Garcia JF, Virgos-Soler A, Prat S (2002) Control of photoperiod-regulated tuberization in potato by the *Arabidopsis* flowering-time gene CONSTANS. Proc Natl Acad Sci USA 99:15211–15216

Michaud R, Lehman WF, Rumbauch MD (1988) World distribution and historical development. In: Hanson AA, Barnes DK, Hill RR (eds) Alfalfa and alfalfa improvement. ASA, CSSA, SSSA, Madison, WI, pp 25–91

Muller MH, Poncet C, Prosperi JM, Santoni S, Ronfort J (2006) Domestication history in the *Medicago sativa* species complex: inferences from nuclear sequence polymorphism. Mol Ecol 15:1589–1602

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

Obert DE, Skinner DZ, Stuteville DL (2000) Association of AFLP markers with downy mildew resistance in autotetraploid alfalfa. Mol Breed 6:287–294

Pajerowska-Mukhtar K, Stich B, Achenbach U, Ballvora A, Lubeck J, Strahwald J, Tacke E, Hofferbert HR, Ilarionova E, Bellin D, Walkemeier B, Basekow R, Kersten B, Gebhardt C (2009) Single nucleotide polymorphisms in the Allene Oxide Synthase 2 gene are associated with field resistance to late blight in

populations of tetraploid potato cultivars. Genetics 181:1115–1127

Pierre JB, Huguet T, Barre P, Huyghe C, Julier B (2008) Detection of QTLs for flowering date in three mapping populations of the model legume species *Medicago truncatula*. Theor Appl Genet 117:609–620

Pierre JB, Bogard M, Herrmann D, Huyghe C, Julier B (2010) Constans gene as a candidate to explain genetic variation for flowering date in *Medicago truncatula*. Mol Breed (in press)

Putterill J, Robson F, Lee K, Simon R, Coupland G (1995) The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc-finger transcription factors. Cell 80:847–857

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. Curr Opin Plant Biol 5:94–100

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

Robson F, Costa MMR, Hepworth SR, Vizir I, Pineiro M, Reeves PH, Putterill J, Coupland G (2001) Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28:619–631

Ronfort JL, Jenczewski E, Bataillon T, Rousset F (1998) Analysis of population structure in autotetraploid species. Genetics 150:921–930

Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496–2497

Simko I, Haynes KG, Jones RW (2006) Assessment of linkage disequilibrium in potato genome with single nucleotide polymorphism markers. Genetics 173:2237–2245

Simon R, Igeno MI, Coupland G (1996) Activation of floral meristem identity genes in *Arabidopsis*. Nature 384:59–62

Skøt L, Humphreys J, Humphreys MO, Thorogood D, Gallagher J, Sanderson R, Armstead IP, Thomas ID (2007) Association of candidate genes with flowering time and water-soluble carbohydrate content in *Lolium perenne* (L.). Genetics 177:535–547

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73:1162–1169

Studer B, Boller B, Herrmann D, Bauer E, Posselt UK, Widmer F, Koelliker R (2006) Genetic mapping reveals a single major QTL for bacterial wilt resistance in Italian ryegrass (*Lolium multiflorum* Lam.). Theor Appl Genet 113:661–671

Vance CP, Miller SS, Gregerson RG, Samac DA, Robinson DL, Gantt JS (1995) Alfalfa NADH-dependent glutamate synthase—structure of the gene and importance in symbiotic N_2 fixation. Plant J 8:345–358

Xing Y, Frei U, Schejbel B, Asp T, Lübberstedt T (2007) Nucleotide diversity and linkage disequilibrium in 11 expressed resistance candidate genes in *Lolium perenne*. BMC Plant Biol 7:43

Yuceer C, Harkess RL, Land SB, Luthe DS (2002) Structure and developmental regulation of CONSTANS-LIKE genes isolated from *Populus deltoides*. Plant Sci 163:615–625