CrossMark

Use of GBS markers to distinguish among lucerne varieties, with comparison to morphological traits

Bernadette Julier (iD) · Philippe Barre · Paola Lambroni · Sabrina Delaunay ·
Muriel Thomasset · Frédéric Lafaillette · Vincent Gensollen

Abstract Variety registration of lucerne often fails because of the difficulty to pass the distinction test. This test is presently based on phenotypic traits. We proposed to evaluate if high-throughput genotyping of varieties with genotyping-by-sequencing markers could help to distinguish among 20 varieties and if the genetic distances were correlated to phenotypic ones. Genotyping was performed either on 40 individuals or on 4 bulks of 100 plants per variety. At the methodological level, we obtained a high number of polymorphic markers, a precise determination of the allele dosage on the individuals, a high reproducibility of allele frequency estimated on bulks and a good correlation between allele frequencies scored on individuals or on bulks. Each pair of the 20 varieties was significantly distinct when using the markers. In addition, the varieties grouped in a way that was consistent with their genetic origin. Genetic distances were correlated to phenotypic distances obtained from data collected in the distinction tests. We suggest that molecular markers could be useful to assist distinction tests during the official trials for registration of lucerne varieties.

Keywords Alfalfa · Genetic distance · *Medicago sativa* · Molecular marker · Variety distinctness

B. Julier (✉) · P. Barre · P. Lambroni · S. Delaunay
P3F, INRA, 86600 Lusignan, France
e-mail: bernadette.julier@inra.fr

M. Thomasset
BioGEVES, Domaine du Magneraud, 17700 Surgères, France

F. Lafaillette
GEVES, Domaine de l'Anjouère, 49370 Erdre-en-Anjou, France

V. Gensollen
GEVES, 711 rue JF Breton, 34090 Montpellier, France

Introduction

Variety registration is based on two series of tests. Distinctness, uniformity and stability (DUS) tests aim at assessing whether candidate variety is (1) distinct from all varieties of common knowledge (here described as the reference collection), (2) genetically uniform (taking into account the reproductive mode) and (3) stable from one generation to the next one. Value for cultivation and use (VCU) tests measure the genetic progress for agronomical and quality traits, compared to control varieties of good value. These tests comply with the regulation adopted by the International Union for the Protection of New Varieties of Plants (UPOV). In the European Union, each member state is responsible for variety registration on its national list, but a variety that has been registered in one state is automatically included to the European list. The DUS tests are also used to grant plant breeders' rights. The Community Plant Variety Office (CPVO) oversees this intellectual property system for the European Union.

Lucerne (*Medicago sativa*) registration is facing severe difficulties due to recurrent failings to pass the

distinction test even though the VCU test is successful. As a consequence, some promising varieties are never registered and so the genetic progress cannot be delivered to the farmers.

Molecular markers are not allowed by UPOV rules for direct variety distinction, except if markers are specifically linked to a phenotypic trait (UPOV 2013). However, phenotypic and genotypic distances may be combined to manage the variety collection. This model aims to eliminate pairwise comparisons between candidate variety and variety of the effective collection, prior to the DUS growing trial. For this purpose, the "Distinctness plus" threshold is a distance calculated with molecular markers and phenotypic traits that undoubtedly separate a candidate variety from some other varieties without a need for direct phenotypic comparison (UPOV 2013). According to UPOV (UPOV 2013), a general trend for a positive correlation between genetic and phenotypic distances has to be established in order to combine molecular and morphological distances (Burstin and Charcosset 1997; Dillmann et al. 1997). This model, accepted by UPOV and based on the "Distinctness plus" threshold (Annex 2 of UPOV 2013) is currently used to manage maize (Zea mais) and spring barley (Hordeum vulgare) collections.

For lucerne DUS testing, phenotypic traits are recorded on spaced plants and small plot designs. In these experiments, the varieties of the reference collection for which seeds are provided by the maintainers (the effective collection of ~270 varieties among 600 varieties in the whole reference collection) and the candidate varieties are compared. To characterize the morphology of the varieties, the main traits are plant height (recorded at different stages of regrowth cycles), flower color, growth habit and regrowth recovery in spring, summer and autumn. In addition, disease tests are carried out under controlled conditions (UPOV 2005a). Based on these datasets, statistical analyses are performed to distinguish each candidate variety from each of the varieties of the effective collection: combined over years analysis for distinction (COYD) are calculated for each quantitative trait (UPOV 2016) and χ^2 tests are carried out for qualitative traits (UPOV 2002). The GAIA method, which is based on multi-trait data (both quantitative and qualitative), has recently been used to define groups of varieties so that all the varieties within a group are distinct from all the varieties of any other group (UPOV 2016).

Most studies of lucerne genetic diversity have been carried out with a low number of markers because only low-throughput technologies [random-amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), microsatellites] have been available until recently. These studies have shown that it was possible to separate wild and cultivated materials, sativa or falcata subspecies, or dormant and non-dormant cultivars (Ariss and Vandemark 2007; Bhandari et al. 2011; Crochemore et al. 1996; Ilhan et al. 2016; Maureira et al. 2004), but the structure among cultivated accessions has been usually weak (Bhandari et al. 2011; Flajoulot et al. 2005; Herrmann et al. 2018; Maureira et al. 2004) and poorly correlated to the structure obtained with phenotypic traits (Annicchiarico et al. 2016; Herrmann et al. 2018; Qiang et al. 2015). In recent years, high-throughput genotyping methods have been established on lucerne: a 10-K single nucleotide polymorphism (SNP) array (Li et al. 2014a) and genotyping-by-sequencing (GBS; Annicchiarico et al. 2016; Li et al. 2014b; Rocher et al. 2015) that have each yielded a few thousand markers. By using such a high number of markers, the structure of the cultivated material was close to that expected from breeding origins (Li et al. 2014a) and the distinctness among lucerne varieties was easier than with microsatellite markers (Annicchiarico et al. 2016).

Genetic diversity analyses have usually been conducted on populations in which a number of individuals were genotyped. Since lucerne is an autotetraploid species, the determination of allele dosage at the individual level is impossible with dominant markers such as AFLP, but feasible with some codominant markers in optimized polymerase chain reaction (PCR) and revelation conditions (Flajoulot et al. 2005). With GBS markers, only three classes of genotypes are recovered, the two monomorphic, quadruplex ones (AAAA or BBBB) and one heterozygous class, comprising simplex (ABBB), duplex (AABB) and triplex (AAAB) genotypes (Li et al. 2014b; Rocher et al. 2015). The allele frequency calculated at the population level from individual data for which the allele dosage was not properly determined may cause bias.

In some studies, when using low-throughput genotyping methods (AFLP for example), pools of individuals were used (Bhandari et al. 2011) while making the assumption that the presence of a marker revealed a high frequency of this marker in the population. With high-throughput SNP genotyping based on sequencing,

the ability to quantify the allele frequency was theoretically demonstrated (Raineri et al. 2012) and the repeatability was tested on bulks of heterozygous diploid populations of perennial ryegrass (*Lolium perenne*; Byrne et al. 2013). This method was used in tetraploid alfalfa (Annicchiarico et al. 2016), but the accuracy of allele frequency estimated on bulks, compared to a frequency based on a set of individuals per population, was never tested.

In this study, we have tested several aspects of GBS technology, including the ability to score the allele dosage at the individual level on an autotetraploid species, the accuracy of the estimation of allele frequency on pools of individuals, the ability to distinguish varieties from a set of 20 varieties and the correlation between the distinctness given by GBS markers and that provided by phenotypic traits during DUS tests.

Material and methods

Material

Twenty varieties have been chosen in the lucerne reference collection to cover all the range of diversity for autumn dormancy, flower color, disease (anthracnose: *Colletotrichum trifolii* and Verticillium: *Verticillium alboatrum*) and pest (stem nematode: *Ditylenchus dipsaci*) resistances (Online Resource 1). All varieties except one (Vernal) were registered in Europe. All these varieties were genotyped in bulks of 100 plants. For 3 of these varieties, 40 individuals were individually genotyped (Herrmann et al. 2010): Galaxie and Félicia, both of dormancy 4, that are known to be phenotypically very similar, and Barmed, dormancy 7, that is phenotypically different from Galaxie and Félicia.

Genotyping

For each variety, seeds were sown in cell trays in a greenhouse. For individual genotyping, young leaflets were sampled, dried in silica gel and DNA was extracted using the protocol of Doyle and Doyle (1987) adapted for 96-well plates, with sodium bisulfite instead of beta-mercaptoethanol as antioxidant. For bulk genotyping, each variety was represented by 4 bulks of 100 different plants. A leaflet of each of the 100 plants was sampled, the bulk of leaflets (42 to 190 mg depending on the bulk) were dried in silica gel and DNA was extracted as

described above. Eight GBS libraries were built, with each one containing 15 individuals and 10 bulks. The construction of the libraries followed the protocols described by Elshire et al. (2011). DNA (100 ng) was digested by the restriction enzyme ApeKI (NEB, R0643S) and ligated to a specific barcode adaptor and a common adaptor. Digested DNA of 25 samples was pooled (5 μl each) and cleaned up with the QIAquick PCR purification kit (Qiagen, cat. no. 28104). For PCR, 2.5 μl of DNA was mixed with 0.4 μM of each primer, 2 mM $MgCl_2$, 0.2 mM dNTP, 1X PCR buffer and 1.25 u Taq (ThermoScientific, #EP0402) in a final volume of 50 μl. Temperature cycling consisted of 72 °C for 5 min, 95 °C for 30 s followed by 18 cycles of 95 °C for 10 s, 65 °C for 30 s, 72 °C for 30 s with a final Taq extension step at 72 °C for 5 min. The PCR was replicated five times and pooled. The libraries were cleaned up using the QIAquick PCR purification kit, checked for quality and diluted to obtain a concentration of 2 nM. Each library was sequenced on a lane of an Illumina HiSeq 3000 at the genomic platform GeT-PlaGe of INRA (INRA Toulouse, France). A 150-base, single-end protocol was used.

A Linux pipeline was developed to obtain the allele dosage of the individuals and the allele frequencies in the bulks. The five successive steps were common to both analyses: i) demultiplexing, ii) trimming of the sequences (discarding of the adaptors with scythe, minimum quality score of 30 and minimum read length of 60 with sickle); iii) alignment of the sequences on *M. truncatula* reference genome sequence, version 4.0 with bwa aln, iv) local realignment with GATK RealignerTargetCreator and IndelRealigner), v) variant calling with GATK HaplotypeCaller (GVCF mode). The GVCF files were merged with GATK GenotypeGVSFs, either according to variety for individuals or including all pools for bulks. The biallele variants with an alignment quality higher than 30 were selected using GATK SelectVariants. Then SNPs with a read depth of at least 30 and a quality of at least 30 were retained. The alleles were coded as 1 for the reference alleles (present in *M. truncatula* sequence) and 0 for the alternative alleles. Regarding the distribution of the reference allele frequency within each individual in the varieties (Online Resource 2), the genotype of each individual was defined by: 0 for individuals with a frequency between 0 and 0.06; 1 for individuals with a frequency between 0.06 and 0.36, 2 for individuals with a frequency between 0.36 and 0.62; 3 for individuals

with a frequency between 0.62 and 0.92; 4 for individuals with a frequency between 0.92 and 1; NA for the rest. For pools, only SNPs with an allele frequency higher than 0.1 and lower than 0.9 were kept. A mean allele frequency was calculated from the four bulks.

Phenotypic data

The 20 varieties are part of the varieties tested in official DUS tests. The French office in charge of lucerne DUS tests (GEVES: Groupe d'Etude et de contrôle des Variétés Et des Semences) provided us with the data collected on these varieties in two locations (Lavalette near Montpellier and Anjouère near Angers) in 2013 and 2014. GEVES, as recommended by UPOV (UPOV 2002, 2008), calculated synthetic statistics called combined-over-year distinctness (COYD) from quantitative morphological traits (UPOV 2016). Thirteen traits were used: growth habit, several measurements of plant height in each season, date of flowering, stem length at flowering and dormancy (UPOV 2005a). In the UPOV regulation, each time a pair of varieties has a significant COYD at a probability of 1% for at least one trait, the varieties are judged as distinct. In this study, a 20 × 20 matrix containing the number of traits for which the COYD for each pair of varieties was significant in the two locations was built. It contained values ranging from 0 (no significant trait) to 13 (2 different varieties for the 13 different traits). In addition, the GAIA software (UPOV 2016) was used to calculate a phenotypic distance with five qualitative and quantitative traits (flower color, dormancy, resistance to stem nematode, *Colletotrichum trifolii* and *Verticillium alboatrum*). The GAIA phenotypic distance between two varieties is a weighted sum of distances on each observed characteristic. GAIA distances are not yet used for variety distinction by GEVES on lucerne, but should be so from 2018.

Statistical analyses

The accuracy of the estimation of allele frequency on bulks of individuals was assessed by the calculation of two components, precision and trueness [ISO 5725-6:1994 (E) 1994]. Precision was evaluated by the coefficients of linear regression between the allele frequencies obtained on pairs of bulks of the same variety and the interval of prediction. The trueness of the allele frequency estimated on bulks was measured by the coefficients of linear regression between the allele frequency average on the 4 bulks and the allele frequency measured on the 40 individuals and the interval of prediction. Trueness was evaluated on the three varieties that were sequenced in both individual and bulk samples.

The structure among the set of 20 varieties was illustrated with a principal component analysis (PCA) calculated with all polymorphic SNPs. The same analysis was further performed on subgroups of varieties. The dudi.pca function of R software (The R Development Core Team 2004) was used. On the data collected at the individual level for three varieties, we performed a PCA and then a discriminant analysis principal component (DAPC) with the function dapc of the adegenet package of R (Jombart et al. 2016).

The possibility to distinguish the varieties was based on the calculation of F_{ST} index among pairs of varieties [stamppFst function of the StAMPP package (Pembleton et al. 2013)] with 1000 permutations. An additional test was conducted with an analysis of molecular variance (AMOVA, stamppAmova function of the StAMPP package) with 10,000 permutations calculated from the Nei distance between each pair of varieties (stamppNei function of the StAMPP package). To evaluate the number of markers required to separate 2 varieties, 100 subsamplings of 50, 100, 150, 500, 1000, 5000 and 10,000 markers were carried out for the most similar pair of varieties; P value of AMOVA and F_{ST} were calculated for each subsample and their mean and standard deviation were drawn as a function of marker number.

The correlation between Nei distance or F_{ST} index and COYD or GAIA distances were calculated and tested with a Mantel test (mantel.test of R) with 1000 permutations.

Results

The mean number of reads, after trimming and mapping, for each individual or bulk varied from 3.7 to 4.6 million reads.

Accuracy of the allele frequency in the bulks of individuals

For the bulks, 39,424 polymorphic markers were obtained. The frequency of missing data in a bulk averaged

5.2%. The reference alleles (present in the *M. truncatula* sequence and coded as 1) were more frequent than the alternative alleles (coded as 0) and few alleles had an intermediate frequency (Online Resource 3). The correlation between the allele frequencies of each pair of the 4 bulks of 100 plants showed a highly significant correlation (r between 0.985 and 0.997, $P < 0.001$) for each of the 20 varieties (Online Resource 3 for Galaxie). The regressions between the allele frequencies of each pair of the four bulks were characterized by slopes and ordinates at the origin that averaged 0.993 and 0.006, respectively, and a mean interval of prediction of 0.116. The precision of allele frequencies estimated on bulks was thus very high.

For the three varieties genotyped at the individual level, the number of polymorphic markers ranged from 52,479 for Félicia to 58,528 for Barmed. The allele frequencies obtained on the 4 bulks of each variety were highly correlated to the frequencies calculated on the 40 individuals ($r = 0.995$, 0.996, 0.996, $P < 0.001$, for the 3 varieties Galaxie, Félicia and Barmed, respectively; Fig. 1 for Galaxie). The regressions were characterized by slopes between 0.987 and 0.991, ordinates at the origin between 0.0034 and 0.0066 and intervals of

prediction of the allele frequency between 0.118 and 0.137 for the three varieties. The trueness of allele frequencies estimated on bulks was thus very good compared to allele frequencies estimated on a set of individuals.

Variety distinctness based on molecular markers

For the three varieties analyzed at the individual level, the PCA (Online Resource 4A) showed that, as expected from phenotypic data, Barmed individuals were grouped and were separated from Galaxie and Félicia individuals, while Galaxie and Félicia individuals were partly overlaid. The DPAC analysis (Online Resource 4) separated Galaxie from Felicia, two varieties that are phenotypically similar.

The PCA with the 20 varieties clearly showed that 5 varieties were different from the 15 others (Fig. 2a): Vernal and yellow-flowered Juurlu, two very dormant varieties from Canada and Estonia, respectively; Greenmed, a dormant turf-type variety with violet flowers; Franken Neu, an old German variety with many violet or variegated flowers and Luzelle, a French grazing type variety with variegated yellow flowers bred from accessions of both subsp. *falcata*
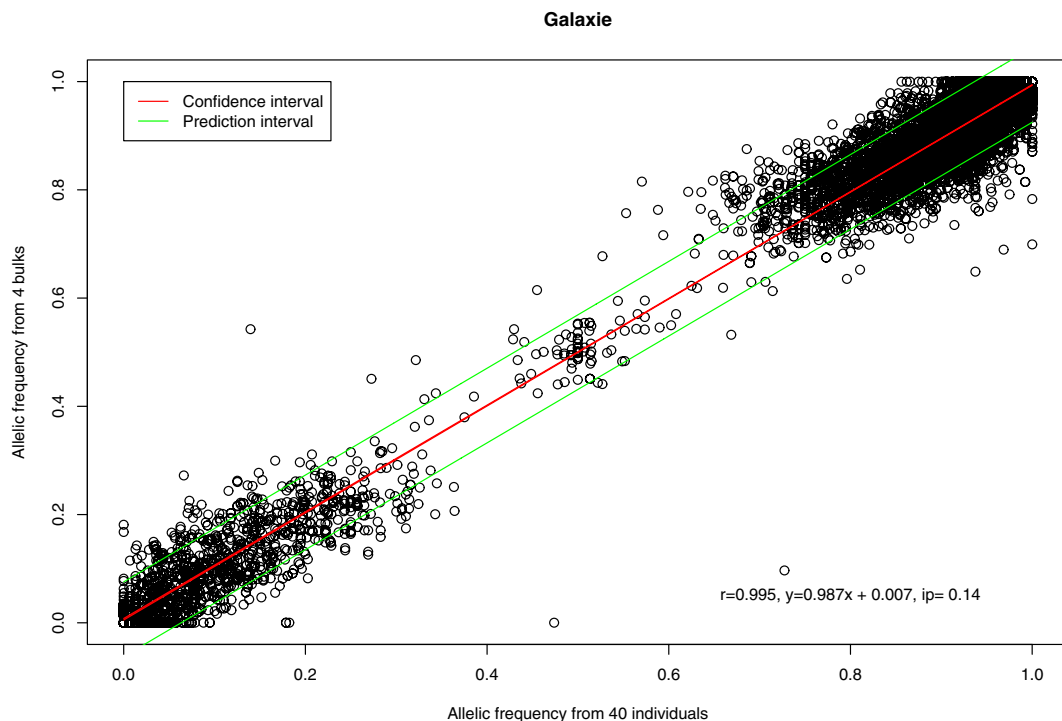
Galaxie

Fig. 1 Correlation between the average allele frequency determined on 4 bulks of 100 plants and 40 individual plants for the variety Galaxie. r is the coefficient of correlation, $y = ax + b$ is the equation of regression, Ip is the mean of the interval of prediction

PCA on 20 varieties

a

PCA on 15 varieties

b

PCA on 9 varieties

c
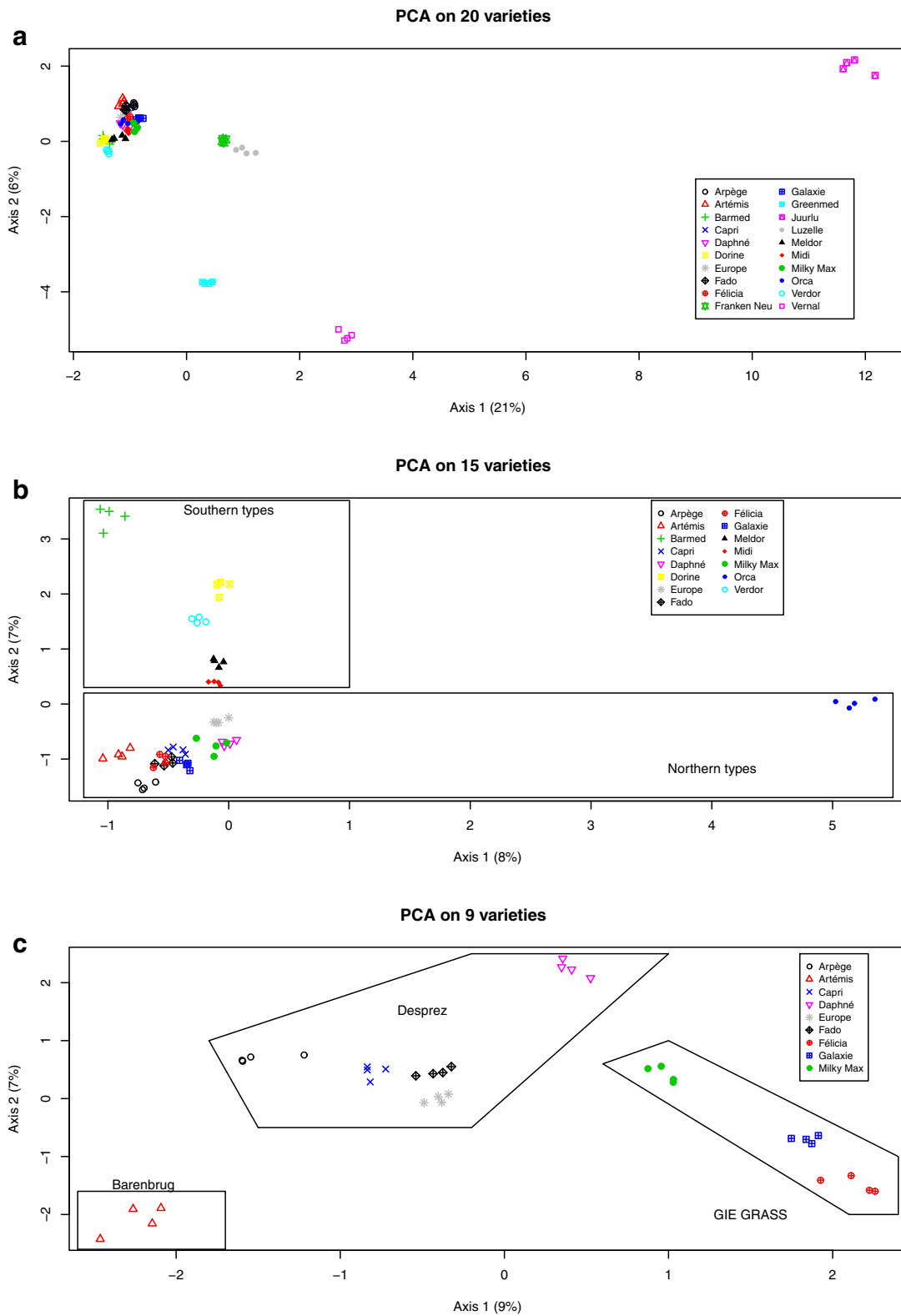
Fig. 2 For the varieties scored as 4 bulks of 100 plants, PCA with the 39,424 GBS markers: (A) with the 20 varieties, (B) with the 15 French varieties, (C) with 9 French varieties adapted to Northern France. The four dots with the same symbol and color corresponded to the four bulks of a variety

Table 1 F_{ST} calculated between pairs of varieties. All F_{ST} were significant at $P < 0.05$

|	Arpège	Artémis	Barmed	Capri	Daphné	Dorine	Europe	Fado	Félicia	Franken Neu	Galaxie	Greenmed	Juurlu	Luzelle	Meldor	Midi	Milky Max	Orca	Verdor
Artémis	0.079																		
Barmed	0.093	0.093																	
Capri	0.078	0.079	0.090																
Daphné	0.081	0.084	0.094	0.080															
Dorine	0.090	0.089	0.094	0.086	0.087														
Europe	0.078	0.078	0.088	0.075	0.079	0.084													
Fado	0.078	0.080	0.092	0.078	0.080	0.089	0.077												
Félicia	0.084	0.084	0.095	0.081	0.084	0.090	0.079	0.082											
Franken Neu	0.090	0.091	0.100	0.087	0.090	0.096	0.085	0.090	0.091										
Galaxie	0.083	0.084	0.094	0.081	0.083	0.089	0.080	0.081	0.081	0.090									
Greenmed	0.116	0.116	0.120	0.114	0.114	0.116	0.113	0.115	0.117	0.115	0.116								
Juurlu	0.188	0.192	0.203	0.189	0.192	0.199	0.190	0.189	0.190	0.171	0.188	0.192							
Luzelle	0.107	0.106	0.115	0.103	0.108	0.111	0.102	0.105	0.107	0.106	0.108	0.125	0.175						
Meldor	0.088	0.089	0.094	0.086	0.087	0.092	0.084	0.087	0.089	0.094	0.088	0.117	0.197	0.110					
Midi	0.081	0.081	0.088	0.077	0.081	0.083	0.075	0.080	0.081	0.085	0.081	0.111	0.190	0.102	0.084				
Milky Max	0.079	0.081	0.092	0.079	0.078	0.085	0.077	0.079	0.080	0.087	0.078	0.111	0.188	0.104	0.085	0.078			
Orca	0.095	0.096	0.104	0.092	0.093	0.097	0.089	0.094	0.096	0.099	0.094	0.123	0.197	0.114	0.098	0.090	0.091		
Verdor	0.088	0.087	0.091	0.085	0.086	0.085	0.083	0.087	0.087	0.096	0.087	0.114	0.200	0.109	0.089	0.081	0.082	0.097	
Vernal	0.118	0.120	0.127	0.117	0.119	0.123	0.117	0.118	0.119	0.111	0.118	0.130	0.165	0.125	0.122	0.115	0.114	0.126	0.121

and *sativa*. A second PCA was carried out with the 15 remaining varieties, all registered in France. On the first two-axis plan, a separation was obtained between the 10 varieties adapted to the northern part of France and the 5 varieties adapted to the southern part of France (Fig. 2b). Among the north varieties, Orca, an old French variety with very dark purple flowers, was separated from the other ones. Finally, PCA was performed with the other nine north varieties only. Interestingly, the varieties were grouped according to the breeder, and particularly the varieties bred by GIE Grass were separated from the varieties bred by Florimond Desprez (Fig. 2c). As anticipated, the four bulks of each variety were grouped on the PCA.

The possibility to distinguish the varieties was tested by calculating AMOVA between pairs of varieties. All the tests were significant, with P values ranging from 0.025 to 0.033. F_{ST} between pairs of varieties (Table 1) ranged from 0.075 (between Europe and Capri, both bred by Florimont Desprez) to 0.203 (between Juurlu and Barmed, the two extreme varieties for autumn dormancy). All F_{ST} were significant at $P = 0.05$. Considering raw allele frequency data, it was noticeable that differences among varieties relied on small variation of allele frequency (not shown).

The fact that all varieties were distinct from one other can be attributed to the number of markers obtained in this study. We performed 100 samplings of 50, 100, 150, 500, 1000, 5000 or 10,000 markers for the 2 most similar varieties, Europe and Capri. The mean of the P values of the AMOVA and its standard deviation declined as the number of markers increased (Fig. 3). The mean F_{ST} value was not greatly affected by the number of markers, but, as for the P value of AMOVA, the standard deviation decreased as the number of markers increased. When the number of markers reached 1000, the mean of the AMOVA P value became very close to the value obtained with all markers ($P = 0.0288$) and the F_{ST} was significant ($P < 0.05$) in all subsamples.

Correlation between molecular distinctness and phenotypic distinctness

The two phenotypic distances, COYD and GAIA, were significantly correlated ($r = 0.520$, P value of Mantel test = 0.005, Online Resource 5). For some pairs of varieties that had low COYD values (0 or only 1 quantitative trait that was significantly different), the GAIA value was high. This means that the qualitative traits included in GAIA
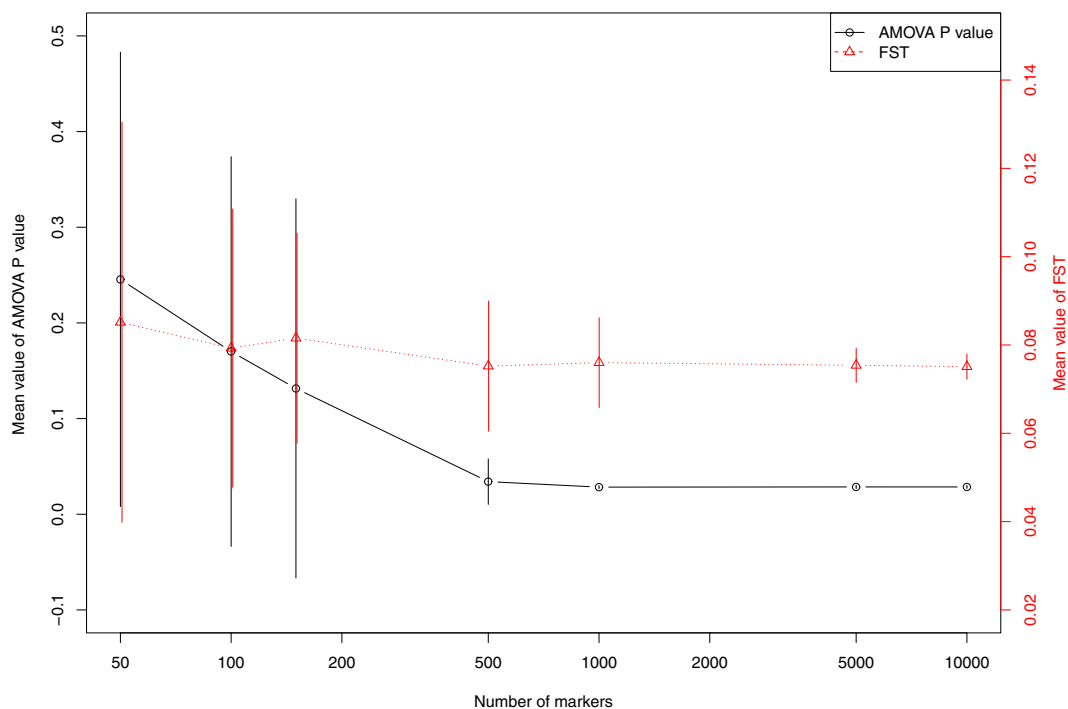
Fig. 3 Effect of the number of markers (from 50 to 10,000) on mean AMOVA P value (left axis), mean F_{ST} value (right axis) and standard deviation (vertical bars) between the varieties Europe and Capri

increased the ability to differentiate the varieties over COYD. The two genetic distances, F_{ST} index and Nei distance, were highly correlated ($r = 0.990$, P value of Mantel test <0.001, Online Resource 5).

The correlation between COYD and Nei distance or F_{ST} index was positive and reached 0.626 and 0.686, respectively (Fig. 4a and b, Online Resource 6), and the Mantel tests were highly significant ($P < 0.001$). Moreover, the GAIA distance was correlated to Nei distance and F_{ST} index, with $r = 0.653$ (Mantel test, $P < 0.001$) and $r = 0.705$ (Mantel test, $P < 0.001$), respectively (Fig. 4c and d, Online Resource 6). Several pairs of varieties had a low genetic distance but a high phenotypic distance while the opposite was infrequent. The graph between F_{ST} and COYD revealed an outlier pair of varieties (Greenmed and Juurlu) that had a high F_{ST} (0.19) and a low COYD (1). However, as their GAIA value was high (85), this pair was plotted with the other pairs of varieties in the graph between F_{ST} and GAIA. Indeed, Greenmed and Juurlu greatly differed for a single quantitative trait (date of beginning of flowering, $P = 0.001$), but were highly different when introducing the qualitative traits (flower colors; Fig. 4).

Juurlu was the variety that generated F_{ST} higher than 0.16 and Nei distances higher than 0.006. When the correlations between F_{ST} or Nei distances and COYD or GAIA were calculated without the Juurlu variety, the correlations decreased by remained highly significant (Online Resource 7). As a main result, the two methods of differentiation, either based on phenotypic traits or molecular markers, gave consistent distances between pairs of varieties.

Discussion

Accuracy of GBS in alfalfa

In this study, the genotyping of individuals or bulks of individuals of tetraploid lucerne with GBS has been successful, as already demonstrated by other groups (Annicchiarico et al. 2016; Li et al. 2014b; Rocher et al. 2015). However, we have also obtained two new important results. Firstly, the possibility to score allele dosage at the individual level, by using thresholds on allele frequencies in order to define classes of genotypes, was shown. This means that the heterozygous plants can now be separated as simplex, duplex or triplex genotypes (A000, AA00 or AAA0, respectively).
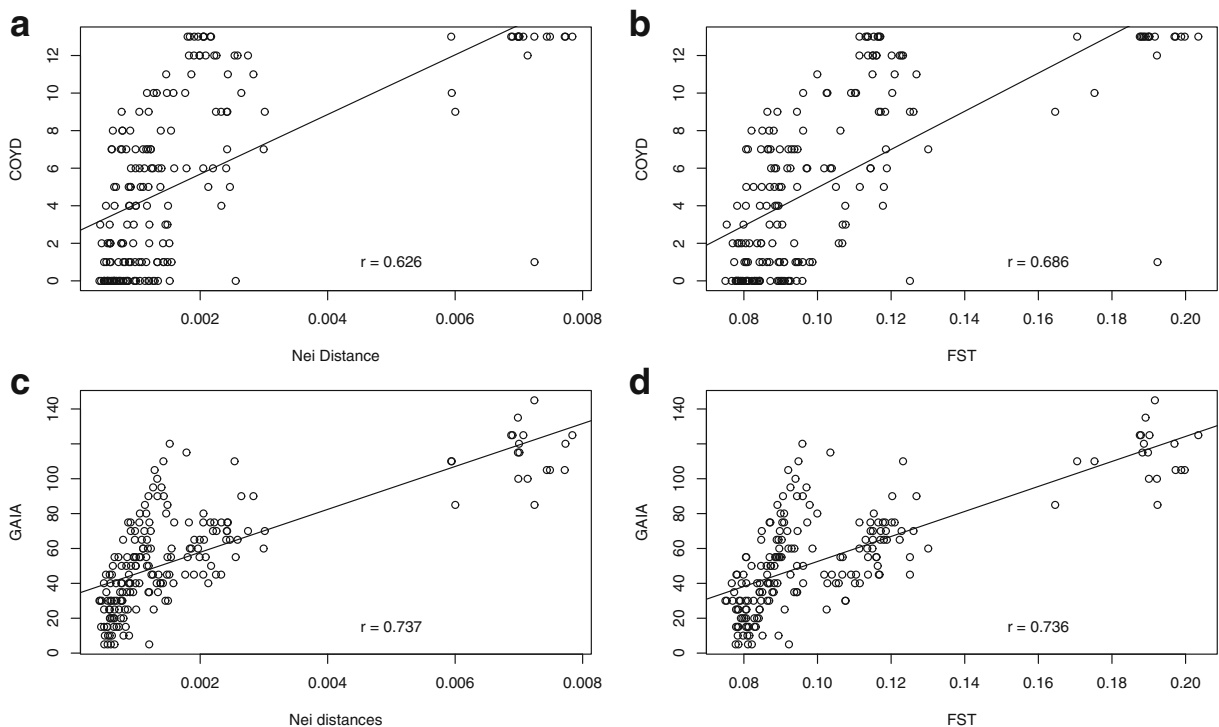
Fig. 4 Correlation between genetic distances (Nei distance or FST index) and phenotypic distances (COYD and GAIA). Mantel tests are presented in Online Resource 5

This provides a more accurate genotyping that should increase the precision in any genetic study, be it either a genetic distance calculation or a test of marker effects on trait variation, such as in association studies. Secondly, we have shown that allele frequencies scored on bulks of individuals are reproducible and well correlated to the allele frequencies calculated as an average of individual genotyping in an autotetraploid species. The use of bulks of genotypes is time- and cost-effective in allogamous species for which populations, including varieties, are composed of heterozygous individuals.

We have obtained several thousands of markers with a low frequency of missing data, thus enabling all types of genetic studies such as diversity analysis, genetic mapping, genome-wide association study or genomic selection on a species for which the investment in genomics has been relatively low. A new development could be to fine-tune the choice of restriction enzyme and common adaptor [with the possible addition of specific base(s) for increased reduction of genome complexity] to obtain the needed marker numbers for a specific experiment with a minimum sequencing effort.

Variety distinction

On this set of 20 varieties that already passed the DUS tests for variety registration, all pairs of varieties were significantly different, even when they had close genetic origins. This result confirms the power of GBS markers for lucerne variety distinctness (Annicchiarico et al. 2016), mainly because of the high number of markers that can be obtained. Indeed, when more than 1000 markers were used, the capacity of variety distinction was reached. Such numbers of markers have never been reached when using AFLP or simple sequence repeats (SSR). Similarly, other studies report a significant distinctness among accessions with molecular markers: in a set of 13 inbred lines of maize (*Zea mays*) genotyped with 33,200 SNP markers (Jambrovic et al. 2014), in a set 27 accessions of ryegrass (*Lolium* sp.) genotyped at the individual level with 369 SNP markers (Wang et al. 2014) and in a set of 8 *L. perenne* accessions genotyped at the population level with GBS markers (Byrne et al. 2013). The power of GBS markers for lucerne variety distinctness has to be confirmed on a larger variety set, especially in the situation where varieties have similar phenotypes and close genetic origins. In addition, the threshold for judging the distinctness has to be elaborated, for example, by comparing different seed lots of the same variety.

The genetic Nei distance and F_{ST} index were significantly correlated to phenotypic distances calculated from GEVES DUS tests (COYD and GAIA). The value of this correlation was affected by the number and the identity of the varieties under study so more varieties are needed to definitively state on the correlation between genetic and phenotypic distance. In contrast, the correlation was not significant in a study carried out on Italian lucerne varieties (Annicchiarico et al. 2016). This discrepancy could originate from the lower number of markers and the lower number of morphological traits obtained on the Italian varieties. In the latter study, some pairs of varieties could not be differentiated with GBS markers. The narrow genetic basis (Italian varieties only) cannot be conjured up because some pairs of varieties were also very close in our study but nevertheless distinguished, and the number of markers was higher than the threshold of 1000 we calculated. The statistics we used (Nei genetic distance and F_{ST} index) could be more efficient than those used by Annicchiarico et al. (2016; allele frequency difference, Least Significant Difference (LSD) for individual principal component axes and discriminant analysis). In most studies where a low number of markers was used, a poor correlation between genetic and phenotypic distances was found (Gunjaca et al. 2008) or a triangular correlation was observed (Burstin and Charcosset 1997; Dillmann et al. 1997). In the latter studies, in contrast to ours, high genetic distances could be observed for pairs of varieties with low phenotypic distances. More recently, on barley (*Hordeum vulgare*), a high correlation was also found between phenotypic and genetic distances, on 431 varieties phenotyped in DUS tests and genotyped with 3072 SNP markers (Jones et al. 2013). In this latter study, the correlation was not improved beyond a certain number of markers.

The use of genetic distances for DUS testing has been questioned many times. At the present time, genetic distances are used in a number of species to restrict the number of varieties to be tested in phenotypic trials (UPOV 2013). A similar use of genetic distances calculated from GBS markers could be applied to lucerne DUS testing. This procedure would limit the size of the field experiments and thus, because of less within-trial environmental variation, increase the precision of the trials. Another solution could be to modify the UPOV regulations and accept a marker-based distinctness as a basic method because a genetic distance calculated from a high number of markers effectively reflects genetic similarity or dissimilarity between varieties. Phenotypic

traits would then still be needed to describe and characterize the varieties. Such evolution would require further studies to choose the best genotyping and statistical methods and to define the thresholds to be used in order to differentiate distinct varieties, similar varieties and essentially derived varieties. Taking into account the increasing size of reference collection that is very sharp with the increase of UPOV members and breeding activity (UPOV 2005b), the system based on phenotypic traits alone is becoming more and more challenging. The UPOV working group on biochemical and molecular techniques can examine the results of this study as an input to decide on the use of molecular markers in DUS tests.

Compliance with ethical standards

Conflict of interest The authors declared no conflict of interest in the authorship and publication of this document.

References

Annicchiarico P, Nazzicari N, Ananta A, Carelli M, Brummer EC (2016) Assessment of cultivar distinctness in alfalfa: a comparison of genotyping-by-sequencing, simple-sequence repeat marker, and morphophysiological observations. Plant Genome 9:1–12

Ariss JJ, Vandemark GJ (2007) Assessment of genetic diversity among nondormant and semidormant alfalfa populations using sequence-related amplified polymorphisms. Crop Sci 47:2274–2284

Bhandari HS, Al Lawati AH, Ray IM (2011) Amplified fragment length polymorphism marker diversity among high yielding perennial alfalfa (*Medicago sativa* L.) core collection accessions. Crop Sci 51:1110–1121

Burstin J, Charcosset A (1997) Relationship between phenotypic and marker distances: theoretical and experimental investigations. Heredity 79:477–483

Byrne S, Czaban A, Studer B, Panitz F, Bendixen C, Asp T (2013) Genome Wide Allele Frequency Fingerprints (GWAFFs) of populations via genotyping by sequencing. Plos One 8

Crochemore ML, Huyghe C, Kerlan MC, Durand F, Julier B (1996) Partitioning and distribution of RAPD variation in a set of populations of the *Medicago sativa* complex. Agronomie 16:421–432

Dillmann C, BarHen A, Guerin D, Charcosset A, Murigneux A (1997) Comparison of RFLP and morphological distances between maize Zea mays L inbred lines. Consequences for germplasm protection purposes. Theor Appl Genet 95:92–102

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. Plos One 6

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

Gunjaca J, Buhinicek I, Jukic M, Sarcevic H, Vragolovic A, Kozic Z, Jambrovic A, Pejic I (2008) Discriminating maize inbred lines using molecular and DUS data. Euphytica 161:165–172

Herrmann D, Flajoulot S, Barre P, Huyghe C, Ronfort J, Julier B (2018) Comparison of morphological traits and molecular markers to analyse diversity and structure of alfalfa (Medicago sativa L.) cultivars. Genet Resour Crop Evol 65: 527–540

Herrmann D, Flajoulot S, Julier B (2010) Sample size for diversity studies in tetraploid alfalfa (*Medicago sativa*) based on codominantly coded SSR markers. Euphytica 171:441–446

Ilhan D, Li XH, Brummer EC, Sakiroglu M (2016) Genetic diversity and population structure of tetraploid accessions of the medicago sativa-falcata complex. Crop Sci 56:1146–1156

ISO 5725-6:1994 (E) (1994) Accuracy (trueness and precision) of measurement methods and results - Part 6: use in practice of accuracy values. In: results ITS-Mma (ed). International Organization for Standardization, Geneve

Jambrovic A, Mazur M, Radan Z, Zdunic Z, Ledencan T, Brkic A, Brkic J, Brkic I, Simic D (2014) Array-based genotyping and genetic dissimilarity analysis of a set of maize inbred lines belonging to different heterotic groups. Genetika-Belgrade 46:343–352

Jombart TCCSP, Ahmed I, Calboli F, Cori A (2016) Package 'adegenet'. Version 2.1.0 edn. CRAN

Jones H, Norris C, Smith D, Cockram J, Lee D, O'Sullivan DM, Mackay I (2013) Evaluation of the use of high-density SNP genotyping to implement UPOV model 2 for DUS testing in barley. Theor Appl Genet 126:901–911

Li XH, Han YH, Wei YL, Acharya A, Farmer AD, Ho J, Monteros MJ, Brummer EC (2014a) Development of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium. Plos One 9

Li XH, Wei YL, Acharya A, Jiang QZ, Kang JM, Brummer EC (2014b) A saturated genetic linkage map of autotetraploid alfalfa (*Medicago sativa* L.) developed using genotyping-by-sequencing is highly syntenous with the *Medicago truncatula* genome. G3-Genes Genomes Genetique 4: 1971–1979

Maureira I, Ortega F, Campos H, Osborn T (2004) Population structure and combining ability of diverse *Medicago sativa* germplasms. Theor Appl Genet 109:775–782

Pembleton LW, Cogan NOI, Forster JW (2013) StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol Ecol Resour 13: 946–952

Qiang H, Chen Z, Zhang Z, Wang X, Gao H, Wang Z (2015) Molecular diversity and population structure of a worldwide collection of cultivated tetraploid alfalfa (*Medicago sativa* subsp. *sativa* L.) germplasm as revealed by microsatellite markers. PLoS One 10:e0124592

Raineri E, Ferretti L, Esteve-Codina A, Nevado B, Heath S, Perez-Enciso M (2012) SNP calling by sequencing pooled samples. BMC Bioinf 13

Rocher S, Jean M, Castonguay Y, Belzile F (2015) Validation of genotyping-by-sequencing analysis in populations of tetraploid alfalfa by 454 sequencing. Plos One 10

The R Development Core Team (2004) The R reference manual, base package, volume 1. Network Theory Limited, Bristol

UPOV (2002) General introduction to the examination of distinctness, uniformity and stability and the development of harmonized descriptions of new varieties of plante. TG/1/3

UPOV (2005a) Lucerne, guidelines for the conduct of tests for distinctness, uniformity and stability. In: UPOV (ed) TG6/5, Geneva, p 129

UPOV (2005b) UPOV report on the impact of plant variety protection. UPOV, Geneva

UPOV (2008) Examining distinctness. In: UPOV (ed) TGP/9. UPOV, Geneva, p 36

UPOV (2013) Guidance of the use of biochemical and molecular markers in the examination of distinctness, uniformity and stability (DUS). DOCUMENT TGP/15, Geneva, pp 1–10

UPOV (2016) Trial design and techniques used in the examination of Distinctness, Uniformity and Stability. In: UPOV (ed) TGP/8. UPOV, Geneva

Wang J, Pembleton LW, Baillie RC, Drayton MC, Hand ML, Bain M, Sawbridge TI, Spangenberg GC, Forster JW, Cogan NOI (2014) Development and implementation of a multiplexed single nucleotide polymorphism genotyping tool for differentiation of ryegrass species and cultivars. Mol Breed 33: 435–451