Industrial Crops and Products 85 (2016) 84–92

Contents lists available at ScienceDirect

Industrial Crops and Products

journal homepage: www.elsevier.com/locate/indcrop

Analysis of genetic diversity and population structure in accessions of the genus *Melilotus*

Fan Wu [a], Daiyu Zhang [a], Jinxing Ma [b], Kai Luo [a], Hongyan Di [a], Zhipeng Liu [a], Jiyu Zhang [a,*], Yanrong Wang [a]

[a] State Key Laboratory of Grassland Agro-ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China
[b] National Quality Control & Inspection Centre for Grassland Industry Products, National Animal Husbandry Service, Ministry of Agriculture, Beijing 100125, China

ARTICLE INFO

Article history:
Received 12 December 2015
Received in revised form 22 February 2016
Accepted 23 February 2016
Available online 5 March 2016

Keywords:
Genetic diversity
Melilotus
Population structure
SSR
UPGMA

ABSTRACT

Melilotus is one of the most important legume plants, in part due to its production of coumarin, but the genetic diversity among the 18 species remains unclear. In the present study, the analysis of genetic diversity of the *Melilotus* species was performed with simple sequence repeat (SSR) markers. From the PCR amplification, we selected 18 out of 70 primers that were used in previous studies. Three hundred twenty-six sampled individuals were assayed to study the genetic diversity and polymorphisms based on the SSR markers. All analyzed markers were polymorphic, and 287 alleles were identified, with 15.94 alleles per locus detected. The polymorphism information content (*PIC*) values ranged from 0.71 to 0.93, with an average of 0.87, which indicates that the markers were highly informative. Based on the unweighted pair-group method with arithmetic mean (UPGMA) cluster analysis, we found that the 18 species were divided into two clusters. *M. italicus* and *M. speciosus* from cluster A and *M. indicus* and *M. segetalis* from cluster B were closely related. Population structure analysis suggested that the optimum number of groups was three. From the analysis of molecular variance, 17.79% of the variance was due to species differences, 31.61% of the variance was due to differences among populations within species and the remaining 50.60% was due to differences within populations. The results of the present study showed that these SSR markers will benefit the *Melilotus* research community for genetics and breeding. Furthermore, this study also established the foundation for future breeding programs, genetic improvement and coumarin production in the *Melilotus* species.

1. Introduction

Melilotus is an important Leguminosae crop and is comprised of 19 annual or biennial species. All species are native to Eurasia or North Africa (Aboel-Atta, 2009). Members of the *Melilotus* genus have high seed yields and are more tolerant to extremes in environmental conditions, such as drought, cold and high salinity, relative to most other forages (Rogers et al., 2008; Sherif, 2009). As forage legumes, they can also perform symbiotic nitrogen fixation with a number of bacterial species (Bromfield et al., 2010). The nitrogen fixation rate of *Melilotus* is higher than that of other legumes, making it beneficial for crop rotations (Stickler and Johnson, 1959). It is important for agriculture and animal husbandry, as it is a green manure crop for grass and can be used as crop fertilizer. In addition to *Melilotus* being an important forage crop, there has been increased interest in its medicinal value, given its variety of biological activities from its coumarin, flavones and saponin constituents (Cong et al., 2012).

Coumarins are an important group of natural compounds that are found in different species of plants in nature, such as *Dipteryx odorata* Willd. (Ehlers et al., 1995), *Angelica archangelica* L. (Hawryl et al., 2000), *Melittis melissophyllum* L. (de Vincenzi et al., 1997) and *Mikania glomerata* Spreng. (Celeghini et al., 2001). It has been reported that coumarin content ranges from 0.05 to 1.30%, as deter-

Abbreviations: AFLP, amplified fragment length polymorphism; AMOVA, analysis of molecular variance; h, Nei's (1973) gene diversity; H_E, expected heterozygosity; H_O, observed heterozygosity; I, Shannon's Information index; na, the observed number of alleles; ne, the effective number of alleles; NPGS, National Plant Germplasm System; NPL, the number of polymorphic loci; PCR, polymerase chain reaction; *PIC*, the polymorphism information content; PPL, the percentage of polymorphic loci; RFLP, restriction fragment length polymorphism; SDS, sodium dodecyl sulfate; SNP, single nucleotide polymorphism; SSR, simple sequence repeat; UPGMA, the unweighted pair-group method with arithmetic mean.

* Corresponding author.
E-mail address: zhangjy@lzu.edu.cn (J. Zhang).

http://dx.doi.org/10.1016/j.indcrop.2016.02.055
0926-6690/

mined from the coumarin content of 149 accessions in 15 *Melilotus* species. No coumarin was detected in an *M. segetalis* accession (Nair et al., 2010). In recent years, a preliminary study found that in addition to conventional carbohydrates, proteins, fats and oils, *Melilotus* contained a variety of chemicals, such as coumarin, flavones, phenolics, saponin and other varieties of chemical components with medicinal value. As one of the cheap abundant medicinal plant resources, *Melilotus* is worth further development for its vast market value (Cong et al., 2012).

At present, several molecular marker techniques are employed to assess genetic diversity in various crops, specially in legume plants. For example, single nucleotide polymorphism (SNP) in *Miscanthus* (de Cesare et al., 2010), restriction fragment length polymorphism (RFLP) in rice (Sun et al., 2001), amplified fragment length polymorphism (AFLP) in sesame (Uzun et al., 2003), random amplified polymorphic DNA (RAPD) in bean (Basheer-Salimia et al., 2013) and simple sequence repeats (SSR) (Wang et al., 2011) in switchgrass have all been used on a large scale, and there are different marker systems for studying genetic diversity in different contexts (Tam et al., 2005). In previous studies, genetic diversity of alfalfa (*Medicago sativa* L.) revealed by AFLP analysis and it suggested that 26 Iranian cultivated populations grouped into four main clusters with no correlation between genetic and geographical diversity (Keivani et al., 2010). The assessment of 58 faba bean (*Vicia faba* L.) genotypes using SRAP markers indicated the high genetic diversity and broad genetic basis due to the rich polymorphism rate (100%) and low genetic similarity (0.21) (Alghamdi et al., 2012). The analysis of genetic structure and diversity in 100 vegetable soybean accessions using simple sequence repeat (SSR) markers showed the vegetable soybean germplasms in China were relatively close and great consistency with the germplasm origins, seed coat colors or their pedigrees (Dong et al., 2014). SSR markers are the markers of choice because of locus specificity, hypervariability, co-dominance and high reproducibility (Powell et al., 1996; Varshney et al., 2009; Varshney et al., 2005) and have been proven to be promising for various genomic applications (Gupta and Varshney, 2000). However, current studies in the *Melilotus* genus mainly concentrate on morphology, cultivation techniques and chemical ingredients, and few studies on mutants, root nodules or genetic diversity have been reported.

Our previous results showed that the interspecific relationships within the *Melilotus* genus based on the phylogenetic tree are clearly monophyletic in the legume family (Di et al., 2015). A preliminary evaluation of agronomy and quality traits of 19 *Melilotus* accessions showed that coumarin content could vary from 0.16–1.02% (Luo et al., 2014). In this study, the SSR markers were used to study the genetic diversity and population structure among 50 accessions of 18 *Melilotus* species, totalling 326 plant samples that were collected. The SSR markers founded in this study will be a useful resource for genetic study and germplasm evaluation for coumarin production in the *Melilotus* genus.

2. Materials and methods

2.1. Plant materials and DNA extraction

Seeds from fifty accessions of eighteen *Melilotus* species were obtained from the National Plant Germplasm System (NPGS, America) as summarized in Table 1 and Fig. 1. The accessions were numbered from 1 to 50, a numbering that was maintained throughout the study. Seeds of all accessions were germinated, and the seedlings were cultivated for DNA extraction in a greenhouse at Lanzhou University in Yuzhong of Gansu Province, China.

Three to twelve individuals of each accession (Table 1), totalling 326 individuals, were sampled and used for polymorphism investigations of the selected SSR markers. Genomic DNA was extracted from the young leaves according to an SDS (sodium dodecyl sulfate) method (Shan et al., 2011). DNA samples were diluted to 50 ng/μL and stored at −20 °C prior to polymerase chain reaction (PCR) amplification.

2.2. PCR amplification

Seventy SSR primer pairs were selected (Supplemental Table 1); these primer sequences were published in previous studies (Winton et al., 2007; Zhou et al., 2014). Primers were excluded from the study if they failed to amplify consistently in PCR amplification. The 18 SSR primer pairs used in the final analysis are presented in Table 2.

The PCR was performed in a total reaction volume of 10 μL containing 1.0 μL DNA, 4.9 μL 2 × Reaction Mix (500 μM dNTP each, 20 mM Tris–HCl, 100 mM KCl, 3 mM $MgCl_2$), 0.1 μL 2.5 U/μL Golden DNA Polymerase, 1.0 μL each primer and 2.0 μL double distilled water. PCR cycling conditions were 3 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at the annealing temperature (Supplemental Table 1), 30 s at 72 °C and a final extension step of 7 min at 72 °C.

2.3. Data analysis

The amplified bands were scored as present (1) or absent (0), and only reproducible bands were considered. The indexes of expected heterozygosity (H_E), observed heterozygosity (H_O) and polymorphic information content (*PIC*) were calculated as previously described (Chung et al., 2013). To evaluate the genetic diversity within species and populations, the following parameters were calculated: the number of polymorphic loci (NPL), the percentage of polymorphic loci (PPL), the observed number of alleles (na), the effective number of alleles (ne), Nei's (1973) gene diversity (h) and Shannon's Information index (I). The program POPGENE 32 (Yeh and Boyle, 1997) was used to calculate NPL, PPL, na, ne, h and I values. The analysis of molecular variance (AMOVA) was used to partition the total genetic variation among species, among populations within species and within populations by the method of AMOVA Version 1.55 (Excoffier, 1993). A cluster analysis was performed to generate a dendrogram using the unweighted pair-group method with arithmetic mean (UPGMA) and Nei's unbiased genetic distance with the help of the SAHN-Clustering by NTSYS-pc.V.2.1 (Rohlf, 2000) software package. A model-based approach implemented in the software program STRUCTURE 2.3 was used to subdivide the individuals into different subgroups (Falush et al., 2007; Pritchard et al., 2000). Due to the estimated 'log probability of data' [LnP(D)] of STRUCTURE overestimating the number of subgroups (Pritchard and Wen, 2003), we used the ad hoc measure ΔK (Evanno et al., 2005) to estimate the number of groups. The membership of each genotype was tested for the range of genetic clusters from $K = 1$ to $K = 8$ (each with 10 independent runs) with the admixture model.

3. Results

3.1. The polymorphism of SSR markers

In this study, 10 DNA samples from different *Melilotus* species were randomly chosen to conduct the primers screening. The primers that successfully amplified and produced clear and stable bands of the expected size by PCR amplification were selected. After the preliminary screening, 18 of 70 pairs of SSR primers with the obvious polymorphism were selected and used in the following analysis.

The eighteen SSR primers used in this study amplified 287 alleles with an average of 15.94 alleles per locus. The majority of the