Genetic variation and diversity in 199 *Melilotus* accessions based on a combination of 5 DNA sequences

Hongxiang Zhang¹*, Fan Wu¹*, Wenli Guo¹*, Rong Bai¹, Zhuanzhuan Yan¹, Blaise Pascal Muvunyi¹, Qi Yan¹, Yufei Zhang¹, Xianfeng Yi²*, Jiyu Zhang¹*

1 State Key Laboratory of Grassland Agro-ecosystems; Key Laboratory of Grassland Livestock Industry Innovation, Ministry of Agriculture; College of Pastoral Agriculture Science and Technology, Lanzhou University; Lanzhou, China, 2 Guangxi Institute of Animal Sciences, Nanning, China

*These authors contributed equally to this work.

1154128631@qq.com (XY); zhangjy@lzu.edu.cn (JZ)

Abstract

*Melilotus* is an important genus of legume plants and an herbaceous with excellent nitrogen fixation; it can tolerate extreme environmental conditions and possesses important medicinal value. However, there is limited genetic information about the genus; thus, we analysed four chloroplast loci (*rbc*L, *mat*K, *psb*A-*trn*H and *trn*L-*F*) and one nuclear region (ITS) to determine the genetic diversity of 199 *Melilotus* accessions from 18 *Melilotus* species. The *rbc*L and *mat*K sequences were highly conserved, whereas the *trn*L-*F* and ITS sequences contained variable and parsimony-informative sites. In our analyses of the single and combined regions, we calculated the pairwise distance, haplotype and nucleotide diversity and gaps and then constructed phylogenetic trees to assess the genetic diversity, and our results revealed significant variations among the different accessions. The genetic distance values were between zero and nine, and based on the combined regions, the highest frequency value was approximately four. *Melilotus* showed high haplotype and nucleotide diversity, particularly in the ITS sequences, with values of 0.86 and 0.0087, respectively. The single ITS sequence, *psb*A-*trn*H, and the combined *mat*K+*rbc*L+*trn*L-*F* (MRT) and *mat*K+*rbc*L+*psb*A-*trn*H+*trn*L-*F*+ITS (MRPTI) regions showed interspecific variation in the gap analysis. Phylogenetic trees calculated using ITS, *psb*A-*trn*H and MRPTI sequences indicated distinct genetic relationship in 18 species, and these species could be divided into two groups. By determining the genetic diversity of plants, we can evaluate the genetic relationships among species and accessions, providing a basis for preserving and utilizing the genetic resources of *Melilotus*.

Introduction

The genus *Melilotus* (sweet clover) consists of 19 annual and biennial species and belongs to the tribe Trifolieae of the legume family. Almost all species are native to North Africa or Eurasia, and many can be found in North of China and Central Asia [1,2]. *Melilotus* is an...
important forage crop, and certain species, such as *M. albus*, *M. officinalis* and *M. indicus* have been cultivated in many regions [3]. Compared with other forages, the members of *Melilotus* can tolerate extreme environmental conditions, such as drought, cold, and high salinity [1,4], and its nitrogen fixation rate is higher than that of other legumes, which can increase soil fertility [5]. Additionally, *Melilotus* is valuable because of its coumarin content [6], and thus represents a possible medicinal plant resource. Due to its affordability and abundance as well as its potential market value, *Melilotus* is worthy of further investigation [6].

Genetic diversity within a particular species helps plants adapt to various environmental conditions, such as fluctuating climate and soil conditions; thus, assessing the diversity of available plant genetic resources is necessary to identify the genes associated with useful biological functions that can then be rationally integrated to design new varieties [7]. Plant genetic diversity has gained increasing attention because of the increase in human population as evidenced by rapid urbanization and the conversion of cultivable lands. These are the critical factors contributing to food insecurity in the developing world [8]. Consequently, the Consultative Group for International Agricultural Researches has begun establishing research centres and gene banks to conserve the plant genetic resources of staple food crops around the world, such as maize from Mexico, rice from North China and potatoes from Peru (for more information, see [http://www.cigar.org/center/index.html](http://www.cigar.org/center/index.html)). The purpose of this organization is to maintain genetic diversity and to provide tools for population monitoring and assessment that can be used for conservation planning [9]. Forage crops also play an increasing role in farming system with the emphasis on development of sustainable agricultural production and the researches about genetic diversity on forage will assume greater importance for germplasm collections and breeding work [10,11]. Genetic diversity assessments are performed using morphological, biochemical and DNA marker analyses. Based on further studies on biological resources, improvements of molecular biology technology, the maturation of amplification and sequencing technologies, and decreases in costs, DNA markers have become the primary method of analysing genetic diversity. Currently, a wide range of DNA markers have been employed to assess genetic diversity, and these include random amplified polymorphic DNA (RAPD) has been adopted in bamboo [12], restriction fragment length polymorphisms (RFLPs) in rice [13], amplified fragment length polymorphisms (AFLPs) in walnuts [14], simple sequence repeats (SSRs) in potato [15], single nucleotide polymorphisms (SNPs) in wheat [16], and chloroplast DNA (cpDNA) in *Brassica napus* [17,18].

A previous study showed that the leaves, flower colour and structure, and pod and seed characteristics of *Melilotus* present extensive variation [19], and the agronomic and quality traits of 19 *Melilotus* species have been evaluated [20]. However, there is limited genetic information regarding *Melilotus* because previous research has concentrated on morphology, cultivation techniques and chemical composition, but SSR marker analyses have shown that *Melilotus* is highly diverse, which is indicative of high allelic richness in the accessions [21]. Understanding the genetic diversity of the plant will enable its genetic material to be preserved as a resources, such as in gene banks and DNA libraries [8]. Additionally, determining the genetic variability in crops provides useful information for breeders developing of new varieties. Cytoplasmic chloroplast genomes (cpDNA) have been widely employed as reference DNA to evaluate population-level genetic diversity [22]. Cytoplasmic chloroplast genomes isinherited highly conservatively in most angiosperms [23], and it has a simplex genome structure and shows vegetative segregation, intracellular selection, and reduced recombination [24,25]. The genetic diversities of *Brassica napus*, *Brassica rapa* and *Brassica oleracea* and their genetic relationships have been finely determined at the cpDNA level [17], and their cpDNA sequences have been used for determinations of genetic structure and population variability in genetic comparisons of Iranian *Asafoetida* (*Ferula assa-foetida* L.) populations [26]. Accordingly, adopting several regions that cover...
nuclear DNA and cpDNA might be an effective strategy for evaluating genetic diversity [27,28]. To further study the genetic diversity of and acquire molecular data on *Melilotus*, we used five sequences, including *rbcL*, *matK*, *psbA-trnH* and *trnL-F* and ITS, to assess the genetic diversity within 18 species.

### Materials and methods

#### Plant materials

Seeds for sampling were selected from 151 accessions from the National Gene Bank of Forage Germplasm (NGBFG, China) and 48 accessions from the National Plant Germplasm System (NPGS, USA) for a total of 199 accessions representing 18 *Melilotus* species (S1 Table). The NGBFG accessions were mainly distributed in North China and adjacent areas such as Russia, whereas the NPGS accessions were mainly distributed in other regions. Because of their hardiness, we rubbed the *Melilotus* seeds between two pieces of sand paper for 1 min, and the seeds were then germinated at 24˚C after incubation over a 16-h light/8-h dark cycle. Ten days later, approximately 20 seedlings of each accession were collected and maintained at -80˚C until assayed.

#### DNA extraction, amplification, and sequencing

Total genomic DNA was extracted from whole seedlings using the SDS (sodium dodecyl sulphate) method [29]. For each accession, two to three DNA samples were extracted, and we selected one sample that revealed a proper DNA concentration for sequencing to minimize errors. Five sequences (four chloroplast regions, *rbcL* [30], *matK* [31], *psbA-trnH* [32] and *trnL-F* [33] and one nuclear gene, ITS [34]) were amplified and sequenced (for the primer sequences, see S2 Table). Amplification was performed by polymerase chain reaction (PCR) in 25-μL mixtures containing 12.25 μL of 2× reaction mix, 0.25 μL of Golden DNA polymerase, 2 μL of each primer (1 μmol / mL), 2 μL of template genomic DNA (50 ng/mL) and 6.5 μL of deionized water as follows: 3 min at 94˚C, 35 cycles of 30 s at 94˚C, 30 s at 53˚C, and 50 s at 72˚C (the annealing temperature and extension time varied according to the sequences, see S2 Table), with a final extension of 7 min at 72˚C before holding at 4˚C. The amplified bands of the PCR products were validated by agarose gel electrophoresis and sequenced by Shanghai Shenggong Biotechnological, Ltd. (Shanghai, China).

#### Data analyses

We used the Contig Express module of Vector NTI Suite 6.0 (InforMax, Inc) to assemble and edit the contigs [35]. Sequences were aligned using DNAMAN 7.0 [36], and the nucleotide variations were then determined. To calculate the average distances and gaps among the 18 species, several sequences were combined, and the combinations of each accession were assembled such that all sequences were connected end to end in the same order. The genetic diversity of *Melilotus* was analysed using DnaSP software, and we estimated the genetic diversity of the species based on the five sequences by calculating the pairwise distances for each locus using MEGA 6.0 according to the number of differences model [37]. In addition, the Emboss Needle algorithm ([http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html](http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html)) was employed to analyse the dissimilarities in nucleotide deletions via a pairwise sequence alignment [38]. We used Bayesian method to construct phylogenetic trees and *Vicia sativa* was adopted as outgroup [36].
Results

DNA was extracted from 199 individuals representing 18 species of Melilotus. The results of the amplification were preliminarily validated by agarose gel electrophoresis (S1 Fig). We failed to obtain the PCR products of the psbA-trnH sequence in M. italicus. The PCR and sequencing success rates consistently exceeded 90% by optimizing the PCR amplification conditions.

Alignment and DNA sequence data

The sequence length, GC content, variable sites and genetic distances based on the five sequences in 18 species were analysed and summarized (Tables 1–5, S2 Table). The lengths of the ITS, matK, rbcL, psbA-trnH and trnL-F sequences were 691, 714, 756, 347 and 673 bp, respectively. Length variations occurred in each region, particularly in trnL-F, which ranged from 431 and 651 bp, but there was little length variation in the sequences of matK and rbcL. Taken together, M. albus, M. elegans, M. officinalis, M. polishus and M. suaveolens had similar sequence lengths. These species had different GC contents within the psbA-trnH and trnL-F sequences. The average GC content of psbA-trnH was 33.27%: M. siculus and M. spicatus had a higher than average GC contents, and M. indicus had the lowest GC content (31.66%). The average GC content of trnL-F was 33.28%. Melilotus indicus and M. albus had the highest and lowest GC contents, which were 34.60% and 31.56%, respectively.

The sequence diversity of the five loci in Melilotus was summarized, and the genetic diversity in each region of the 18 species was calculated (Tables 1–5, S2 Table). Analysis of the ITS sequences revealed 525 conserved sites and 121 variable sites, including 49 parsimony-informative sites and 72 single variable sites.

Table 1. Sequence characteristics, pairwise distances and analyses of genome genetic diversity based on ITS sequences.

| Species          | Length (bp) | G+C Content (%) | Intraspecific distance (mean) | Interspecific distance (mean) | Haplotype diversity (%) | Nucleotide diversity (%) |
|------------------|-------------|-----------------|------------------------------|------------------------------|-------------------------|--------------------------|
| M. albus         | 646         | 49.69           | 0.000                        | 0.5–3.48 (1.860)             | 0.418±0.073              | 0.0007±0.00074            |
| M. altissimus    | 650         | 49.33           | 0.000                        | 1.86–5.522 (2.698)           | 0.833±0.222              | 0.0038±0.00199            |
| M. dentatus      | 646         | 49.69           | 0.000                        | 0.5–3.48 (1.860)             | 0.778±0.091              | 0.0016±0.00095            |
| M. elegans       | 646         | 49.58           | 0–1.969 (1.217)              | 0.5–5.544 (2.188)            | 0.893±0.111              | 0.0034±0.00146            |
| M. hirsutus      | 646         | 49.36           | 0–0.971 (0.647)              | 0.5–4.343 (1.921)            | 0.667±0.314              | 0.0020±0.00146            |
| M. indicus       | 649         | 49.20           | 0–1.875 (1.234)              | 0.5–7.65 (3.558)             | 0.800±0.164              | 0.0033±0.00148            |
| M. infestus      | 649/658     | 49.31           | 0.000                        | 2.36–5.284 (4.446)           | 0.000                    | 0.000                     |
| M. italicus      | 648         | 50.00           | 0.000                        | 3.301–5.481 (3.694)          | 0.000                    | 0.000                     |
| M. officinalis   | 646         | 49.48           | 0–1.371 (0.713)              | 0.5–4.749 (2.156)            | 0.710±0.079              | 0.0018±0.00093            |
| M. polonicus     | 646         | 49.69           | 0.000                        | 0.5–3.48 (1.860)             | 0.000                    | 0.000                     |
| M. segetalis     | 649/652     | 49.36           | 4.765                        | 0.5–2.21 (3.694)             | 1.000±0.500              | 0.0294±0.0004567          |
| M. siculus       | 654/655     | 49.84           | 0.986–3.120 (2.358)          | 2.303–5.765 (4.620)          | 1.000±0.500              | 0.0015±0.00153            |
| M. speciosus     | 681         | 50.51           | 0.000                        | 4.654–6.147 (4.679)          | 0.000                    | 0.000                     |
| M. spicatus      | 647         | 49.40           | 0.000                        | 2.505–5.312 (2.947)          | 0.500±0.265              | 0.0007±0.00084            |
| M. suaveolens    | 646         | 49.63           | 0–1.641 (0.656)              | 0.942–5.594 (2.185)          | 0.529±0.117              | 0.0014±0.00110            |
| M. sulcatus      | 655         | 49.46           | 0.000                        | 2.303–5.266 (4.292)          | 0.000                    | 0.000                     |
| M. tauricus      | 648         | 49.69           | 0.000                        | 1.367–5.407 (2.355)          | 0.000                    | 0.000                     |
| M. wilgicus      | 646         | 49.94           | 0–0.994 (0.397)              | 0.970–5.441 (2.231)          | 0.55±0.165               | 0.0012±0.00099            |
| Average          | 691         | 49.62           |                              | 0.861±0.024                 | 0.0086±0.00255           |

ITS, internal transcribed spacer. The genetic diversity was analysed by a single sequence, and the same applies to the following tables.

https://doi.org/10.1371/journal.pone.0194172.t001
### Table 2. Sequence characteristics, pairwise distances and analyses of genome genetic diversity based on matK sequences.

| Species | Length (bp) | G+C Content (%) | Intraspecific distance (mean) | Interspecific distance (mean) | Haplotype diversity (%) | Nucleotide diversity (%) |
|---------|-------------|-----------------|-------------------------------|-------------------------------|-------------------------|-------------------------|
| M. ablus | 714         | 29.78           | 0.00                          | 0.00                          | 0.421±0.074            | 0.00074±0.00052         |
| M. altissimus | 714     | 29.69           | 0.00                          | 0.00                          | 0.833±0.222            | 0.00014±0.000108        |
| M. dentatus | 714       | 29.64           | 0.975–1.429 (1.133)           | 0.970–0.999 (2.008)           | 0.533±0.180            | 0.00177±0.00111         |
| M. elegans | 714        | 29.73           | 0.00                          | 0.00                          | 0.786±0.113            | 0.000170±0.00094        |
| M. hirsutus | 714       | 29.97           | 0.966–2.203 (1.716)           | 0.960–0.989 (1.989)           | 1.000±0.272            | 0.000467±0.00209        |
| M. indicus | 714        | 29.69           | 0.00                          | 0.00                          | 0.800±0.164            | 0.000224±0.00116        |
| M. infestus | 714       | 29.69           | 1.328–1.374 (1.348)           | 1.328–1.461 (1.348)           | 1.000±0.272            | 0.000280±0.00162        |
| M. italicus | 714        | 29.55           | 1.677–3.254 (2.048)           | 1.677–3.254 (2.048)           | 0.000±0.0000            | 0.000                  |
| M. officinalis | 714     | 29.64           | 0.00                          | 0.00                          | 0.518±0.122            | 0.000107±0.00076        |
| M. polonicus | 714       | 29.77           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. segetalis | 714       | 29.76           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. sclosus | 714        | 29.76           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. speciosus | 714      | 29.83           | 1.680                         | 1.680                         | 0.000±0000              | 0.000                  |
| M. spicatus | 714        | 29.83           | 1.680                         | 1.680                         | 0.000±0000              | 0.000                  |
| M. suaveolens | 714      | 29.69           | 0.971–2.936 (2.388)           | 0.971–2.936 (2.388)           | 0.667±0.314            | 0.000187±0.00132        |
| M. sulcatus | 714        | 29.90           | 1.000                         | 1.000                         | 1.000±0.500            | 0.000280±0.00198        |
| M. tauricus | 714        | 29.69           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. wilgicus | 714        | 29.83           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| Average   | 714        | 29.75           | 0.00                          | 0.00                          | 0.080±0.021            | 0.000294±0.00043        |

and nucleotide diversity. Melilotus siculus had high haplotype diversity but low nucleotide diversity. Analysis of the other four chloroplast regions showed that the haplotype and nucleotide diversities were lower than those of the ITS sequences. Melilotus infestus showed high

### Table 3. Sequence characteristics, pairwise distances and analyses of genome genetic diversity based on rbcL sequences.

| Species    | Length (bp) | G+C Content (%) | Intraspecific distance (mean) | Interspecific distance (mean) | Haplotype diversity (%) | Nucleotide diversity (%) |
|------------|-------------|-----------------|-------------------------------|-------------------------------|-------------------------|-------------------------|
| M. ablus   | 754         | 40.71           | 0.00                          | 0.00                          | 0.098±0.053             | 0.00013±0.00049         |
| M. altissimus | 754       | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. dentatus | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. elegans | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. hirsutus | 754       | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. indicus | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. infestus | 754       | 40.71           | 0.00                          | 0.00                          | 0.500±0.265             | 0.00332±0.00162         |
| M. italicus | 755        | 40.49           | 1.831–2.509 (1.936)           | 1.831–2.509 (1.936)           | 0.000±0000              | 0.000                  |
| M. officinalis | 755      | 40.71           | 0.00                          | 0.00                          | 0.228±0.102             | 0.00003±0.00036         |
| M. polonicus | 754       | 40.71           | 0.00                          | 0.00                          | 0.200±0.154             | 0.000027±0.00047        |
| M. segetalis | 754       | 40.71           | 1.705                         | 1.705                         | 1.000±0.500             | 0.000398±0.00230        |
| M. sclosus | 754        | 40.84           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. speciosus | 754       | 40.58           | 1.291–2.212 (1.641)           | 1.291–2.212 (1.641)           | 0.000±0000              | 0.000                  |
| M. spicatus | 754        | 40.45           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. suaveolens | 754      | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. sulcatus | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. tauricus | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| M. wilgicus | 754        | 40.71           | 0.00                          | 0.00                          | 0.000±0000              | 0.000                  |
| Average   | 756        | 40.69           | 0.316±0.046                   | 0.316±0.046                   | 0.00117±0.00093         |

https://doi.org/10.1371/journal.pone.0194172.t002

https://doi.org/10.1371/journal.pone.0194172.t003
Table 4. Sequence characteristics, pairwise distances and analyses of genome genetic diversity based on *psbA-trnH* sequences.

| Species         | Length (bp) | G+C Conten (%) | Intraspecific distance (mean) | Interspecific distance (mean) | Haplotype diversity (%) | Nucleotide diversity (%) |
|-----------------|-------------|----------------|--------------------------------|--------------------------------|-------------------------|--------------------------|
| *M. albata*     | 318         | 33.64          | 0.000                          | 0.250                          | 0.000                   | 0.000                    |
| *M. altissima*  | 317/318     | 33.05          | 0.000                          | 0.250                          | 0.000                   | 0.000                    |
| *M. dentatus*   | 318/321     | 32.91          | 0.000                          | 0.250                          | 0.000                   | 0.000                    |
| *M. elegans*    | 318         | 33.64          | 0.000                          | 0.250                          | 0.000                   | 0.000                    |
| *M. hirsutus*   | 318/319     | 32.88          | 0.000                          | 0.250                          | 0.000                   | 0.000                    |
| *M. indicus*    | 319         | 31.66          | 1.660                          | 4.590                          | 0.000                   | 0.000                    |
| *M. infestus*   | 317/318     | 33.68          | 0.929                          | 2.342                          | 0.000                   | 0.000                    |
| *M. officinalis*| 318         | 33.54          | 0.464                          | 1.624                          | 0.000                   | 0.000                    |
| *M. polonicus*  | 318         | 33.64          | 1.389                          | 4.644                          | 0.000                   | 0.000                    |
| *M. tauricus*   | 318         | 33.33          | 0.982                          | 4.518                          | 0.000                   | 0.000                    |
| *M. volgicus*   | 322/334     | 32.72          | 0.000                          | 0.000                          | 0.000                   | 0.000                    |
| Average         | 347         | 33.27          | 0.507±0.042                    | 0.00643±0.00305                | 0.000                   | 0.000                    |

Table 5. Sequence characteristics, pairwise distances and analyses of genome genetic diversity based on *trnL-F* sequences.

| Species         | Length (bp) | G+C Conten (%) | Intraspecific distance (mean) | Interspecific distance (mean) | Haplotype diversity (%) | Nucleotide diversity (%) |
|-----------------|-------------|----------------|--------------------------------|--------------------------------|-------------------------|--------------------------|
| *M. albata*     | 432         | 31.56          | 0.965                          | 8.074                          | 0.667±0.163             | 0.02299±0.00409          |
| *M. altissima*  | 643         | 33.13          | 0.940                          | 8.060                          | 0.000                   | 0.000                    |
| *M. dentatus*   | 639         | 33.02          | 1.389                          | 4.644                          | 0.000                   | 0.000                    |
| *M. elegans*    | 431         | 32.10          | 1.300                          | 4.282                          | 0.000                   | 0.000                    |
| *M. hirsutus*   | 657         | 33.49          | 0.981                          | 8.095                          | 0.000                   | 0.000                    |
| *M. indicus*    | 442/612     | 34.60          | 0.187                          | 8.107                          | 0.000                   | 0.000                    |
| *M. infestus*   | 442/646     | 33.13          | 0.943                          | 8.147                          | 0.000                   | 0.000                    |
| *M. italicus*   | 617         | 34.36          | 0.943                          | 8.147                          | 0.000                   | 0.000                    |
| *M. officinalis*| 451         | 33.48          | 0.742                          | 0.000                          | 0.000                   | 0.000                    |
| *M. polonicus*  | 452/453     | 33.26          | 0.745                          | 0.000                          | 0.000                   | 0.000                    |
| *M. segretalis* | 612/655     | 33.97          | 0.943                          | 8.147                          | 0.000                   | 0.000                    |
| *M. siculus*    | 653/654     | 33.21          | 0.943                          | 8.147                          | 0.000                   | 0.000                    |
| *M. speciosus*  | 641         | 32.23          | 0.943                          | 8.147                          | 0.000                   | 0.000                    |
| *M. spicatus*   | 452/607     | 34.39          | 0.742                          | 0.000                          | 0.000                   | 0.000                    |
| *M. suaveolens* | 432         | 32.60          | 0.935                          | 8.144                          | 0.000                   | 0.000                    |
| *M. sulcatus*   | 647/656     | 33.85          | 2.062                          | 0.000                          | 0.000                   | 0.000                    |
| *M. tauricus*   | 612/646     | 33.48          | 0.804                          | 0.000                          | 0.000                   | 0.000                    |
| *M. wilgicus*   | 638         | 33.23          | 0.806                          | 0.000                          | 0.000                   | 0.000                    |
| Average         | 673         | 33.28          | 0.965                          | 9.044                          | 0.000                   | 0.12924±0.01834          |

https://doi.org/10.1371/journal.pone.0194172.t004
https://doi.org/10.1371/journal.pone.0194172.t005
Genetic variation and diversity in 199 Melilotus accessions based on a combination of 5 DNA sequences

Genetic distance and similarity
To determine the average distances among the 18 species, we combined several sequences to calculate the pairwise distances (Fig 1). These combinations included matK+trnL-F (MT),
matK+rbcl+trnL-F (MRT), matK+rbcl+psbA-trnH+trnL-F (MRPT), matK+rbcl+ITS+trnL-F (MRIT), and matK+rbcl+psbA-trnH+trnL-F+ITS (MRPTI). Our results revealed that the highest frequency was approximately four in all combinations. The genetic distance value ranged from zero to nine according to the different accessions. For the five single regions, more
than half of the species exhibited zero intraspecific distances within the ITS and matK sequences. Several species, including *M. albus*, *M. italicus*, *M. officinalis*, *M. polonicus*, *M. speciosus* and *M. tauricus*, showed small intraspecific distances. *M. indicus*, *M. infestus*, *M. italicus*, *M. siculus*, *M. speciosus* and *M. sulcatus* had similar interspecific distances because of the large distances between these species calculated using four loci (except *trnL*-F, which presented high interspecific distances in all species).

Fig 3. Average interspecific gene gap value for single and combined regions based on the analysis of hundreds of sequences by pairwise sequence alignment.

https://doi.org/10.1371/journal.pone.0194172.g003
Furthermore, we calculated the similarity, gaps and scores through a pairwise sequence alignment to analyse the intraspecific and interspecific divergence (Figs 2 and 3). We observed few nucleotide deletions in matK and rbcL, zero gaps, and approximately 100% similarity. A diversity of gaps was observed in the other three single sequences and combined regions. The intraspecific gap value of *M. infestus* and *M. segetalis* was large in the ITS sequence and close to zero in the other sequences. Approximately half of the species showed differences in interspecific gaps. *M. speciosus* had the highest gap value of 0.069. According to the alignment of *psbA*-
trnH, almost all species showed variations in interspecific gaps. The four species, *M. dentatus*, *M. segetalis*, *M. spicatus* and *M. wolgicus*, showed high intraspecific diversity in gap values. For the trnL-F sequence, the intraspecific gap value was higher in *M. altissimus*, *M. indicus*, *M. infestus* and *M. officinalis*, *M. polonicus*, *M. segetalis*, *M. siculus*, *M. speciosus*, *M. spicatus*, *M. suaveolens*, *M. sulcatus*, *M. tauricus*, and *M. wolgicus*.

[Fig 5. Bayesian tree with branch length, based on ITS sequences.](https://doi.org/10.1371/journal.pone.0194172.g005) The abbreviations represent 18 species: Ma—*M. albus*, Mm—*M. altissimus*, Md—*M. dentatus*, Me—*M. elegans*, Mh—*M. hirsutus*, Mi—*M. indicus*, Mf—*M. infestus*, Mr—*M. italicus*, Mo—*M. officinalis*, Mp—*M. polonicus*, Mg—*M. segetalis*, Ms—*M. siculus*, Mu—*M. speciosus*, Mn—*M. suaveolens*, Mi—*M. sulcatus*, Mt—*M. tauricus*, and Mw—*M. wolgicus*.
Cluster analysis

Based on genetic distance and similarity analysis, the ITS and psbA-trnH showed high discriminating in Melilotus. And phylogenetic trees were constructed using the Bayesian method based on the ITS, psbA-trnH respectively, and a combined region of the five sequences (Figs 4 and 5, S2 Fig). According to the Bayesian trees, all species showed distinct diversity with little mixing and we could divided 18 species into two groups on the basic of clades and branch length in Figs 4 and 5. The first group included M. albus, M. altissimus, M. dentatus, M. elegans, M. hirsutus, M. officinalis, M. polonicus, M. segetalis, M. suaveolens, M. tauricus and M. wolgicus. The second group contained the others. Similar to the analysis of genetic diversity, M. albus, M. officinalis, M. tauricus and M. polonicus showed small intraspecific distances since there are few clades among them. What’s more, the species in the first group have a closer genetic relationship due to similar branch length. The variation was smaller when we combined five sequences (S2 Fig). But we also could find the differences of accessions from different regions. Five accessions of M. albus from NPGS (USA) didn’t get together with other accessions from NGBFG (China). And accessions from the same area tend to be together, such as Mo_CF... and Mo_PI... in M. officinalis.

Discussion

Melilotus is an important forage and green manure crop with high protein content and the ability to fix nitrogen [39], and it also plays an important role in soil improvement, is drought resistant and moderately winter hardy, and has good dry matter production [40,41]. However, there is limited information regarding its morphology, cultivation techniques and chemical composition; thus, genetic diversity analyses based on five sequences were performed in this study to provide a reference for the conservation of genetic resources.

As shown in Tables 1–5 and S2 Table, the trnL-F and ITS sequences provided more variation sites and informative characteristics than the other three sequences. Many studies have reported that ITS sequences have more variable and informative sites than cpDNA [32,42]: White (1990) found that the ITS sequence conserved its length and presented high nucleotide variability [43]. Chloroplast DNA, which is conservatively inherited, has a simple genome structure [17] and presents differences among populations and individuals in many species [44,45]. Due to differences in length, the trnL-F region contained more variation sites and greater genetic distances than the other regions in this study. The rbcL and matK sequences were highly conserved in Melilotus, and the similarity among accessions was greater than 99.20% and 98.60%, respectively. The analysis of the five loci showed that the ITS and psbA-trnH sequences had more genetic diversity in Melilotus, whereas rbcL and matK could be used in combination with other sequences. In addition, the intraspecific and interspecific distances calculated using the rbcL and matK sequences were smaller than those of the other sequences. Our results were similar to the findings obtained by Chen, Yao et al [46], who found that the order of the sequences according to the interspecific distance value (from large to small) was ITS, psbA-trnH, matK and rbcL. Previous researches showed that the rbcL and matK sequences were suitable for analyzing relationship at higher taxonomic levels [47,48], and when assessing genetic diversity at the species or genus level, the ITS and psbA-trnH sequences worked better and the combination of several sequences yielded the best result. Moreover, the sequence characteristics in Melilotus could be used in further genetic diversity studies.

Our previous study showed that the SSR marker analysis revealed highly significant differences in genetic differentiation among accessions within Melilotus species accessions [21]. According to the haplotype and nucleotide diversity calculated using the five sequences, Melilotus showed higher diversity, particularly within the ITS sequences, than Euphrasia and
Rhododendron [49,50], which had a haplotype and nucleotide diversity of 0.86 and 0.0087, respectively. The analysis of gaps in the 18 species revealed a high degree of variation based on the ITS, psbA-trnH and combined MRT and MRPI sequences. Additionally, the distance in Melilotus was significantly higher than that in species of Rhodiola [32] calculated using ITS, rbcL, matK and psbA-trnH regions. Compared with Medicago sativa, Melilotus has not been cultivated commercially and shows higher diversity [21,51]. In addition, cluster analysis divided 18 species into two groups, which is similar to the results of previous studies [21,34]. Traditional identification depends on the shape of the torus and flower colour [52], but plant morphology might vary greatly within a single plant [53]. This research divided Melilotus into two groups according to flower colour. The white group contains four species, M. albus, M. tauricus, M. wolgicus and M. speciosus, and the other species composed the yellow group [53]. However, due to differences in the result of these molecular studies, flower colour has no obvious link with the phylogenetic classification [34] and might not be a reliable basis for classification. Compared with the results in previous studies [21,34], M. albus, M. altissimus, M. elegans, M. officinalis, M. polonicus, M. suaveolens, M. wolgicus were included in the first group and M. italicus, M. speciosus, M. siciulus, M. indicus, M. sulcatus, M. infestus, M. spicatus were included in the second group. These Melilotus species clustering within the same group may have closer genetic relationships.

The significant differences were revealed among Melilotus species, and several, such as M. infestus and M. segetalis, showed higher diversity than the others. To reduce the loss of Melilotus genetic resources, it is necessary to strengthen the collection and protection of wild germplasm resources. Current situation of the conservation of genetic resources in Melilotus is too many individuals of M. albus and M. officinalis were collected but the germplasm collections of other species were insufficient. Imbalance of these species is the main problem we are facing in germplasm collection and conservation. Surveys should occur worldwide and more individuals of the species with high diversity revealed by Tables 1–5 might need to be collected. The growing population pressure and urbanization of agricultural lands as well as the rapid modernization of every aspect of our day-to-day activities have caused biodiversity decreased in directly and indirectly, and the large-scale cultivation of genetically homogenous varieties also reduces species diversity and genetic variation [54]. What’s more, this study identified additional Melilotus genetic resources for breeding purpose. The loss of genetic diversity has been recognized as the result of a genetic bottleneck imposed on crop plants during domestication and by modern plant-breeding practices [55]; thus, this research could provide a reference for the conservation of genetic resources that currently exist for the future breeding work.

Our results identified the characteristics of five sequences in Melilotus and indicated that analyses of these regions represent a valuable method for assessing genetic diversity. The analysis of the five loci provided important genetic information that will assist in germplasm collection and conservation of Melilotus.

Supporting information

S1 Fig. Electrophoretogram of the PCR amplifications based on five sequences. From one to eighteen: M. albus, M. altissimus, M. dentatus, M. elegans, M. hirsutus, M. indicus, M. infestus, M. italicus, M. officinalis, M. polonicus, M. segetalis, M. siciulus, M. speciosus, M. spicatus, M. suaveolens, M. sulcatus, M. tauricus, and M. wolgicus.

(TIF)

S2 Fig. Bayesian tree based on the combined regions of five sequences. The abbreviations represent 18 species: Ma-M. albus, Mm-M. altissimus, Md-M. dentatus, Me-M. elegans, Mh-M. hirsutus, Mi-M. indicus, Mf-M. infestus, Mr-M. italicus, Mo-M. officinalis, Mp-M. polonicus,
Mg.-M. segetalis, Mc.-M. siciulus, Ms.-M. speciosus, Mu.-M. spicatus, Mn.-M. suaveolens, Mi.-M. sulcatus, Mt.-M. tauricus, Mw.-M. wolgicus. See S1 Table for the accession number.

S1 Table. Information on the 199 Melilotus accessions included in this study.

S2 Table. Information on the primers, amplification conditions and sequence statistics for the five sequences.

S3 Table. Variable sites of the five sequences in the 18 species. The four colours represent the four canonical bases: green for adenine (A), blue for cytosine (C), purple for guanine (G) and orange-red for thymine (T). The symbol “▲” represents the sequence deletions from 1 to 188.

Acknowledgments
This work was supported by the National Basic Research Program (973) of China (2014CB138704), the National Natural Science Foundation of China (31572453), the Program for Changjiang Scholars and the Innovative Research Team in Chinese Universities (IRT_17R50), the Open Project Program of the State Key Laboratory of Grassland Agro-ecosystems (SKLGAE201702), and the 111 project (B12002). Additionally, we thank the NPGS and NGBFG for providing the experimental materials used in our study.

Author Contributions
Conceptualization: Fan Wu, Jiyu Zhang.
Data curation: Hongxiang Zhang.
Funding acquisition: Xianfeng Yi, Jiyu Zhang.
Investigation: Rong Bai, Qi Yan, Xianfeng Yi.
Methodology: Wenli Guo, Zhuanzhuan Yan, Blaise Pascal Muvunyi.
Resources: Xianfeng Yi.
Software: Wenli Guo, Zhuanzhuan Yan, Blaise Pascal Muvunyi.
Supervision: Jiyu Zhang.
Validation: Wenli Guo, Rong Bai, Qi Yan, Yufei Zhang, Jiyu Zhang.
Writing – original draft: Hongxiang Zhang.
Writing – review & editing: Fan Wu.

References
1. Al Sherif EA. Melilotus indicus (L.) All., a salt-tolerant wild leguminous herb with high potential for use as a forage crop in salt-affected soils. Flora. 2009; 204: 737–746. https://doi.org/10.1016/j.flora.2008.10.004.
2. Pavlova D, Tosheva A. Karyological study of Melilotus alba Med. (Fabaceae) populations in Bulgaria. Caryologia. 2002; 55: 105–110. http://dx.doi.org/10.1080/00087114.2002.10589264.
3. Brenner DM. Sweetclover descriptors for GRIN. Newsletter. 1970; 141: 51–55.
4. Rogers M, Colmer T, Frost K, Henry D, Cornwall D, Hulm E, et al. Diversity in the genus Melilotus for tolerance to salinity and waterlogging. Plant Soil. 2008; 304: 89–101. https://doi.org/10.1007/s11104-007-9523-y

5. Stickler FC, Johnson I. Dry matter and nitrogen production of legumes and legume associations in the fall of the seeding year. Agron J. 1959; 51: 135–137. https://doi.org/10.2134/agronj1959.002196200510030004x

6. Cong J, Chen F, Sun C. Study on comprehensive development of Melilotus suaveolens L. Journal of Anhui Agricultural Sciences. 2012; 5: 2962–2963.

7. Onda Y, Mochida K. Exploring Genetic Diversity in Plants Using High-Throughput Sequencing Techniques. Curr Genomic. 2016; 17: 358–367.

8. Govindaraj M, Vetrieventhal M, Srivinasa M. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genet reseash international. Genet Res Int. 2015; 2015:431487. http://dx.doi.org/10.1155/2015/431487.

9. Smale M, Mar I, Jarvis DI. The economics of conserving agricultural biodiversity on-farm: research methods developed from IPGRI's global project 'strengthening the scientific basis of in situ conservation of agricultural biodiversity'. International Plant Genetic Resources Institute. 2002; Rome.

10. Robertson LD, Singh KB, Erskine W, Amaem. Useful genetic diversity in germplasm collections of food and forage legumes from West Asia and North Africa. Genetic Resour Crop Evol. 1996; 43: 447–460.

11. Karaca M. Genetic diversity among forage bermudagrass (Cynodon spp.): Evidence from chloroplast and nuclear DNA fingerprinting. Crop Sci. 2002; 42: 2118–2127. https://doi.org/10.2135/cropsci2002.2118

12. Baishya S, Rathi S, Sarma A. Genetic Variation in Bamboo Species of North East India through RAPD. Indian Journal of Agricultural Biochemistry. 2016; 29: 36–40. https://doi.org/10.5998/0974-4479.2016.00006.X

13. Sun C, Wang X, Li Z, Yoshimura A, Iwata N. Comparison of the genetic diversity of common wild rice (Oryza rufipogon Griff.) and cultivated rice (O. sativa L.) using RFLP markers. Theor Appl Genet. 2001; 102: 157–162. https://doi.org/10.1007/s001220051631

14. Li AM, Zubair SJ, Abbass AM, Jubaqal JM. Genetic diversity among Walnuts (Juglans regia) population in Kurdistan Region–Iraq using AFLP-PCR. ZANCO Journal of Pure and Applied Sciences. 2016; 28: 50–55. http://dx.doi.org/10.21271/ZJPAS.28.4.8.

15. Ghebreslassie BM, Gitihiri SM, Mekarii RW, Ghislaq M, Magembe E. Genetic diversity assessment of farmers’ and improved potato (Solanum tuberosum) cultivars from Eritrea using simple sequence repeat (SSR) markers. Afr J Biotechnol. 2016; 15:1883–1891. https://doi.org/10.5897/AJB2016.15237

16. Spanic V, Korzun V, Ebmeyer E. Assessing genetic diversity of wheat genotypes from different origins by SNP markers. Cereal Res Commm. 2016; 44: 361–369. http://dx.doi.org/10.1556/0806.44.2016.012.

17. Qiao J, Cai M, Yan G, Wang N, Li F, Chen B, et al. High-throughput multiplex cpDNA resequencing clarifies the genetic diversity and genetic relationships among Brassica napus, Brassica rapa and Brassica oleracea. Plant Biotechnol J. 2016; 14: 409–418. https://doi.org/10.1111/pbi.12395 PMID: 26031705

18. Lee JH, Lee DH, Choi BH. Phylogeography and genetic diversity of East Asian Neolitsea sericea (Lauraceae) based on variations in chloroplast DNA sequences. J Plant Res. 2013; 126: 193–202. https://doi.org/10.1007/s10265-012-0519-1 PMID: 22990429

19. Moussavi S. Species of Melilotus in Iran (key to the species, descriptions and their distributions). Rostanahia 2. 2001.

20. Luo K, Di H, Zhang J, Wang Y, Li Z. Preliminary evaluation of agronomy and quality traits of nineteen Melilotus accessions. Pratacultural Sci. 2014; 31: 2125–2134.

21. Wu F, Zhang DY, Ma JX, Luo K, Di HY, Liu ZP, et al. Analysis of genetic diversity and population structure in accessions of the genus Melilotus. Ind Crops Prod. 2016; 85: 84–92. https://doi.org/10.1016/j.indcrop.2016.02.055.

22. Petit RJ, Abdelhamid EM, Pons OM. Identifying populations for conservation on the basis of genetic markers. Conserv Biol. 1998; 12(4): 844–855. https://doi.org/10.1111/j.1523-1739.1998.96489.x

23. Dumolin S, Demesure B, Petit RJ. Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. Theor Appl Genet. 1995; 91: 1253–1256. https://doi.org/10.1007/BF00220937. PMID: 24170504

24. Birky CW. The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. Annu Rev Genet. 2001; 35: 125–148. https://doi.org/10.1146/annurev.genet.35.102401.090231. PMID: 11700280

25. Qiao J, Cai M, Yan G, Wang N, Li F, Chen B, et al. High-throughput multiplex cpDNA resequencing clarifies the genetic diversity and genetic relationships among Brassica napus, Brassica rapa and Brassica oleracea. Plant Biotechnol J. 2016; 14: 409–418. https://doi.org/10.1111/pbi.12395 PMID: 26031705

26. Lee JH, Lee DH, Choi BH. Phylogeography and genetic diversity of East Asian Neolitsea sericea (Lauraceae) based on variations in chloroplast DNA sequences. J Plant Res. 2013; 126: 193–202. https://doi.org/10.1007/s10265-012-0519-1 PMID: 22990429
25. Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. DNA Research. 1999; 6: 283–290. https://doi.org/10.1093/dnares/6.5.283. PMID: 10574454

26. Khederzadeh S, Samiei M, Mobarakia A, Ezeddinloo L, Haghi HA. Genetic Comparison of Iranian Asafetida (*Ferula assa-foetida*) Populations Based on cpDNA Ribosomal Protein L16 Intron. IJAIR. 2017; 5: 2319–1473.

27. Fjellheim S, Tanhuanpää P, Marum P, Manninen O, Rognli OA. Phenotypic or molecular diversity screening for conservation of genetic resources? An example from a genebank collection of the temperate forage grass timothy. Crop Sci. 2015; 55: 1646–1659.

28. Dong W, Liu J, Yu J, Wang L, Zhou S. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. PLoS ONE. 2016; 7: e35071. https://doi.org/10.1371/journal.pone.0035071.

29. Shan Z, Wu HL, Li CL, Chen H, Wu Q. Improved SDS method for general plant genomic DNA extraction. Guangdong Agricultural Sciences. 2011; 8: 113–115.

30. Dong W, Cheng T, Li C, Xu C, Long P, Chen C, et al. Discriminating plants using the DNA barcode *rbcLb*: an appraisal based on a large data set. Mol Ecol Resour. 2014; 14: 336–343. doi: 10.1111/1755-0998.12185. PMID: 24119263

31. Yu J, Xue JH, Zhou SL. New universal matK primers for DNA barcoding angiosperms. J Syst Evol. 2011; 49: 176–181. https://doi.org/10.1111/j.1759-6831.2011.00134.x

32. Zhang JQ, Meng SY, Wen J, Rao GY. DNA barcoding of *Rhodiola* (Crassulaceae): a case study on a group of recently diversified medicinal plants from the Qinghai-Tibetan Plateau. PLoS ONE. 2015; 10: e0119921. https://doi.org/10.1371/journal.pone.0119921. https://doi.org/10.1371/journal.pone.0119921 PMID: 25774915

33. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol. 1991; 17: 1105–1109. https://doi.org/10.1007/BF00037152 PMID: 1932684

34. Di HY, Duan Z, Luo K, Zhang DY, Wu F, Zhang JY, et al. Interspecific phylogenetic relationships within genus *Melilotus* based on nuclear and chloroplast DNA. PLoS ONE. 2015; 10: e0132596. https://doi.org/10.1371/journal.pone.0132596 PMID: 26167689

35. Lu G, Moriyama EN. Vector NTI, a balanced all-in-one sequence analysis suite. Brief Bioinform. 2004; 5: 378–388. https://doi.org/10.1093/bib/5.4.378. PMID: 15606974

36. Brady SG, Schultz TR, Fisher BL, Ward PS. Evaluating alternative hypotheses for the early evolution and diversification of ants. PNAS. 2006; 103: 18172. https://doi.org/10.1073/pnas.0605858103 PMID: 17079492

37. Peralta A, Robles C, Martínez A, álvaro L, Valera A, Calamante G, et al. Identification and molecular characterization of Orf virus in Argentina. Virus Genes. 2015; 50: 381–388. https://doi.org/10.1007/s11262-015-1189-6. https://doi.org/10.1007/s11262-015-1189-6 PMID: 25796398

38. Lucas C, Thangaradjou T, Papenbrock J. Development of a DNA barcoding system for seagrasses: successful but not simple. PLoS ONE. 2012; 7: e29987. https://doi.org/10.1371/journal.pone.0029987. https://doi.org/10.1371/journal.pone.0029987 PMID: 22253849

39. Zhang YM, Ma HL, Calderón-Urrea A, Tian CX, Bai XM, Wei JM. Anatomical changes to protect organelle integrity account for tolerance to alkali and salt stresses in *Melilotus officinalis*. Plant Soil. 2016; 406: 327–340. https://doi.org/10.1007/s11104-016-2875-4

40. Rogers M, Colmer T, Nichols P, Hughes S, Frost K, Cornwall D, et al. Salt tolerance and avoidance mechanisms at germination of annual pasture legumes: importance for adaptation to saline environments. Plant Soil. 2009; 315: 241. https://doi.org/10.1007/s11104-008-9747-5

41. Nichols P, Malik A, Stockdale M, Colmer T. Salt tolerance and avoidance mechanisms at germination of annual pasture legumes: importance for adaptation to saline environments. Plant Soil. 2009; 315: 241. https://doi.org/10.1007/s11104-008-9747-5

42. Liu JQ, Wang YJ, Wang AL, Hideaki O, Abbott RJ. Radiation and diversification within the *Ligularia–Cremanthodium–Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. Mol Phylogenet Evol. 2006; 38: 31–49. https://doi.org/10.1016/j.ympev.2005.09.010. https://doi.org/10.1016/j.ympev.2005.09.010 PMID: 16290033

43. White T. Analysis of phylogenetic relationships by amplification and direct sequencing of ribosomal RNA genes. PCR Protocols: a guide to methods and applications. 1990.

44. Bai WN, Liao WJ, Zhang DY. Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytol. 2010; 188: 892. https://doi.org/10.1111/j.1469-8137.2010.03407.x PMID: 20723077
45. Pauwels M, Vekemans X, Godé C, Frérot H, Castric V, Saumitou-Laprade P. Nuclear and chloroplast DNA phylogeography reveals vicariance among European populations of the model species for the study of metal tolerance, Arabidopsis halleri (Brassicaceae). New Phytol. 2012; 193: 916–928. https://doi.org/10.1111/j.1469-8137.2011.04003.x PMID: 22225532

46. Chen S, Yao H, Han J, Liu C, Song J, Shi L. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE. 2010; 5: e8613. https://doi.org/10.1371/journal.pone.0008613. PMID: 20662805

47. Johnson LA, Soltis DE. matK DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. Syst Bot. 1994; 19: 143–156. https://doi.org/10.2307/2419718

48. Chase MW, Soltis DE, Olmstead RG, Morgan D, Les DH, Mishler BD. Phylogenetics of Seed Plants: An Analysis of Nucleotide Sequences from the Plastid Gene rbcL. Ann Mo Bot. Gard. 1993; 80: 528–580. https://doi.org/10.2307/2399846

49. Chung JD, Lin TP, Chen YL, Cheng YP, Hwang SY. Phylogeographic study reveals the origin and evolutionary history of a Rhododendron species complex in Taiwan. Mol Phylogenet Evol. 2007; 42: 14–24. https://doi.org/10.1016/j.ympev.2006.06.027. PMID: 17070712

50. Wu M.J, Huang S.F, Huang T.C, Lee P.F, Lin T.P. Evolution of the Euphrasia transmorrisonensis complex (Orobanchaceae) in alpine areas of Taiwan. Journal of biogeography. 2005; 32: 1921–1929. https://doi.org/10.1111/j.1365-2699.2005.01327.x

51. Wang Z, Yan H, Fu X, Li X, Gao H. Development of simple sequence repeat markers and diversity analysis in alfalfa (Medicago sativa L.). Mol Biol Rep. 2013; 40: 3291–3298. https://doi.org/10.1007/s11033-012-2404-3. PMID: 23275197

52. Flora of China: Science Press. 2004.

53. Stevenson GA. An agronomic and taxonomic review of the genus Melilotus Mill. Canadian J Plant Sci. 1969; 49: 1–20.

54. Álvarez-Buylla E, Aragón F, Hilbeck A, Dusen EV, Whalon ME, Wilkes G. Maize and Biodiversity: The Effects of Transgenic Maize in Mexico. Nation. 2004; 291: 30.

55. Lopes MS, El-Basyoni I, Baenziger PS, Singh S, Royo C, Ozbek K, et al. Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. J Exp Bot. 2015; 66: 3477–3486. https://doi.org/10.1093/jxb/erv122. PMID: 25821073