Canadian Journal of Plant Science
Revue canadienne de phytotechnie

Genetic diversity and relationship of sainfoin (*Onobrychis viciifolia* Scop.) germplasm as revealed by amplified fragment length polymorphism markers

Journal:	<i>Canadian Journal of Plant Science</i>
Manuscript ID	CJPS-2017-0094.R2
Manuscript Type:	Article
Date Submitted by the Author:	09-Sep-2017
Complete List of Authors:	Bhattarai, Surendra; University of Saskatchewan, Plant Sciences Coulman, Bruce; University of Saskatchewan, Department of Plant Sciences Fu, Yong-Bi; Agriculture and Agri-Food Canada Saskatoon Beattie, Aaron; University of Saskatchewan, Crop Development Centre Biligtetu, Bill; University of Saskatchewan, Department of Plant Science
Keywords:	Sainfoin, forage breeding, legume, genetic diversity

SCHOLARONE™
Manuscripts Only

Genetic diversity and relationship of sainfoin (*Onobrychis viciifolia* Scop.) germplasm as revealed by amplified fragment length polymorphism markers

Surendra Bhattarai¹, Bruce Coulman¹, Yong-Bi Fu², Aaron D. Beattie¹, Bill Biligetu^{1,3}

¹Department of Plant Sciences, University of Saskatchewan, 51 Campus Drive, Saskatoon, Saskatchewan, Canada S7N 5A8; ²Plant Gene Resources of Canada, Saskatoon Research and Development Centre, Agriculture and Agri-Food Canada, 107 Science Place, Saskatoon, Saskatchewan, Canada S7N 0X2.

³Corresponding Author (email: bill.biligetu@usask.ca)

Abstract

Sainfoin (*Onobrychis viciifolia* Scop.) is a perennial forage legume, widely distributed in the northern temperate regions of the world. Its genetic improvement has been relatively slow due, in part, to the lack of genetic information including molecular characterization of sainfoin germplasm. An attempt was made to evaluate genetic diversity and relationships among 38 sainfoin accessions collected from different regions of the world using amplified fragment length polymorphism (AFLP) markers. Five AFLP primer pairs were used to assess 367 individual plants, which produced 1042 polymorphic AFLP bands. The frequencies of the scored bands in all assayed individuals ranged from 0.003 to 0.973 with a mean value of 0.165. The analysis of molecular variance (AMOVA) revealed higher within-accession (84.3%) genetic variation than among accessions (15.7%). The genetic distance based on inter-accession distance matrices was significant for most of accessions, but it was not significant for accessions sharing similar parents. A dendrogram of the collected accessions showed two clusters at an inter-accession genetic distance coefficient of 0.36. The revealed information on genetic distance and genetic

diversity of the sainfoin accessions is useful for selecting genetically diverse germplasms for sainfoin genetic improvement efforts.

Keywords

Sainfoin, forage breeding, legume, genetic diversity

Sainfoin (*Onobrychis viciifolia* Scop.) belongs to the Fabaceae family and Hedysareae tribe, and is native to south central Asia (Burton and Curley 1968). The countries of Iran and Turkey are considered to be the main centers of diversity (Yildiz et al. 1999). It is a legume mostly cross-pollinated by honey bees (*Apis mellifera* L.) and leaf cutting bees (*Megachile rotundata* Fabricius.) (Goplen et al. 1991). Interestingly, sainfoin populations have tetraploid ($2n=4x=28$) or diploid ($2n=2x=14$) forms (Frame et al. 1998; Zarrabian et al. 2013). Sainfoin produces high quality, non-bloating forage for livestock production; but it yields approximately 5-20% less dry matter than alfalfa (*Medicago sativa* L.) in western Canada (Goplen et al. 1991). This plant adapted well to a wide range of climatic and soil conditions (Frame 2005). However, it displays several challenges for genetic improvement with poor regrowth, lack of persistence in forage mixtures, low tolerance to waterlogging and frost, and its poor competitive ability in the early stages of development (Hayot-Carbonero 2011; Bhattarai et al. 2016; Kempf et al. 2016).

Although sainfoin was introduced to North America in 1786 (Hanna 1972), its forage production on this region largely relies on a few cultivars. In Canada, “Melrose” (released in 1969), “Nova” (released in 1980) and “AAC Mountainview” (released in 2013) are the only three cultivars of sainfoin. In recent years, there has been increasing demand for high yielding sainfoin cultivars with good stand persistence and rapid recovery under grazing in the temperate

regions of the North America. However, the progress in genetic improvement of sainfoin has been relatively slow due to lack of sufficient breeding efforts. A better understanding of genetic diversity and relationships of diverse germplasm would be useful for sainfoin breeding as it may provide additional genetic information to select germplasm based on forage yield components (Majidi et al. 2009). The maintenance of high genetic diversity within a breeding program is also crucial for sainfoin breeding, as it is susceptible to inbreeding depression (Rasouli et al. 2013).

Various morphological, anatomical and molecular characteristics techniques can be employed to study genetic diversity. The inference of genetic diversity using molecular markers should be more informative than morphological traits as the former is not affected by environmental factors (Awasthi et al. 2004). DNA-based fingerprinting technologies have been applied in genetic studies in a wide range of plant species. In sainfoin, random amplified polymorphic DNA (RAPD; Williams et al. 1990) and inter-simple sequence repeats (ISSR) markers have been utilized to analyze its genetic diversity (Nosrati et al. 2012; Rasouli et al. 2013; Zarrabian et al. 2013; Avci et al. 2014; Hejrankesh et al. 2014; Zarrabian and Majidi 2015). Efforts have also been made to develop amplified fragment length polymorphisms (AFLP) (Vos et al. 1995) and simple sequence repeat (SSR) markers with little success (Hayot-Carbonero 2011; Demdoun et al. 2012) until recently with the discovery of 101 informative SSR markers (Kempf et al. 2016). More recently, an RNA-seq analysis revealed 3786 potential SSRs and 77,000 putative single nucleotide polymorphisms (SNPs) (Mora-Ortiz et al. 2016), opening new opportunity to characterize and study sainfoin germplasm for its breeding and research.

In 2013, we initiated an AFLP analysis of sainfoin germplasm to assess the genetic diversity and genetic relationships among 38 sainfoin accessions collected from different regions of the world. The AFLP technique is a robust method of DNA fingerprinting, but it has limitations

(Williams et al. 1990; Vos et al. 1995). AFLP markers are dominant in nature, and cannot distinguish homozygous and heterozygous alleles, and their informativeness has been questioned in many aspects of genetic analysis (e.g., Mechanda et al. 2004; Koopman 2005). We nevertheless applied AFLP marker in this study with the following considerations. First, it does not require any prior sequence information and is suited for detecting genetic variation in cross-pollinated, polyploid crops with little genomic information. Second, it is more efficient for detecting polymorphisms and more informative than RAPD and ISSR (Powell et al. 1996; Loh et al. 2000). Third, AFLP-based genetic diversity analyses have been successfully demonstrated in many forage crops (Ferdinandez and Coulman 2002; Fu et al. 2004; Peng et al. 2008; Harris et al. 2010; Biligetu et al. 2013).

MATERIALS AND METHODS

Plant materials

The 38 sainfoin accessions used for this study are originated from 21 countries representing four continents (Table 1; Figure 1). Seedlings were grown in the greenhouse for eight weeks prior to being transferred to the field in July 2014 near Saskatoon (52° 07' N, 106° 38' W), SK, Canada. Thirty-two individual plants for each accession were space-planted in the field using a randomized complete block design with four replications consisting of eight individual plants. Row spacing was 1m between any two individual plants.

Tissue collection and DNA extraction

In August 2015, fresh leaves of 10 randomly selected plants were sampled from each of the 34 sainfoin accessions in the field plot. Due to a low winter survival rate, six individual plants were randomly selected for the accessions PI568208, CN45949 and LRC3519, and nine plants for

PI380949 (Table 1). Leaf samples were kept in a cooler on ice during transportation from the field to lab. The leaf samples were freeze-dried (in a Labconco Freeze Dry System for 48 h) and stored at -20°C. Approximately 10-15 mg of leaves of each plant were placed in a 1.5 mL micro-centrifuge tube with three, 2 mm glass beads. The leaf tissue was ground to a fine powder using a tissue lyser. The DNA was extracted using a DNeasyTM Plant Mini Kit (Qiagen Inc. Mississauga, ON) according to the manufacturer's instructions. Extracted DNA was quantified using a Nanodrop ND-8000 (Thermo Fisher Scientific, Wilmington, DE, USA), followed by dilution to 25ng μL^{-1} for AFLP analysis.

AFLP analysis

AFLP analysis was performed using the AFLPTM Analysis System 1 (Life Technologies Burlington, ON) according to Vos et al. (1995) with some modifications. DNA was digested with enzymes *EcoRI* and *MseI* at 37°C for 2 h and inactivated at 70°C for 15 min. Immediately after restricted digestion of genomic DNA, the adapter was ligated to the restriction fragments with T4 DNA ligase at room temperature (20-25°C) for 2 h to create primary templates. The ligated DNA was diluted 10 fold in the 96 well plate with 0.1M TE (10mmol l^{-1} Tris-HCl pH 8.0, 0.1mmol l^{-1} EDTA, distilled water), then stored at -20°C. The ligated DNA was pre-amplified with pre-amplification primers. The amplified product was diluted 50 fold with 0.1M TE and used for selective amplification. Selective amplification reactions were performed using pre-amplified DNA and fluorescently labelled FAM *EcoRI* primers and unlabeled *MseI* primers. Ten *EcoRI*: *MseI* primer pairs were initially screened on six leaf samples and the five most informative primer combinations were selected and used for selective amplification. The product was diluted 40 fold with ultra-pure water. One microliter of diluted sample was combined with 8.75 μl of HI-DYE formamide and 0.25 μl of 600 LIZ and denatured for 3 minutes at 95°C and

analyzed on an Applied Biosystems 3100x1 genetic analyzer (Applied Biosystems Life Technologies, Burlington, ON). Data were collected using GeneMapper software version 4.1 (Applied Biosystems Inc.) with bin size set at 1bp and scored from 60 to 600 bp. The minimal acceptable peak height was 50 units. As a check, two randomly selected individuals were repeated on each of the plates to minimize technical and scoring errors.

Data analysis

Data from GeneMapper were exported to Microsoft Excel and converted into a binary matrix and scored as 1 (present) or 0 (absent) with missing values recorded as 9. The analysis of AFLP polymorphisms was carried out by counting the number of polymorphic bands and generating summary statistics. The polymorphism information content (PIC) for each primer was calculated to measure the informativeness of a marker using the following formula:

$$PIC_i = 2 f_i (1 - f_i) \text{ (Roldan-Ruiz et al. 2000),}$$

where ' i ' is the i th primer, PIC_i is the polymorphic information content of marker ' i ', f_i is the frequency of the amplified allele and $(1 - f_i)$ is the frequency of the null allele.

To assess AFLP variation, Arlequin software version 3.05 (Excoffier et al. 2005) was used to perform an analysis of molecular variance (AMOVA). Genetic associations among the 38 sainfoin accessions were analyzed using NTSYS-pc 2.1 (Rohlf 1997) based on the inter-accession distance matrices obtained from the AMOVA. A dendrogram of the 38 sainfoin accessions was also made in NTSYS-pc 2.1 using the unweighted pair group methods with arithmetic mean (UPGMA). The genetic association among the 367 individual plants was further assessed without restriction to accession origin. A neighbor-joining (NJ) analysis was made

using PAUP* (Swofford 1998) based on the original dataset of 1042 alleles, and a radiation tree was created using MEGA 3.01 (Kumar et al. 2004).

RESULTS

AFLP markers

The five *Eco*RI: *Mse*I-based primer pairs yielded clear and informative banding patterns useful for genetic analysis of sainfoin. The observed band frequencies ranged from 0.003 to 0.973 with a mean value of 0.165 (Table 2). Mean value, minimum and maximum band frequencies for each primer pair are shown in Table 2. The lowest mean frequency (0.126) was observed for the primer pair E+ACC/M+CTC while the highest mean frequency (0.196) was observed for primer pair E+ACA/M+CAT. The PIC value ranged from 0.129 to 0.186 with a mean value of 0.166 (Table 2). The presence of a large number of polymorphic bands detected by the five AFLP primer pairs showed that AFLP markers are informative for evaluating the genetic diversity of sainfoin germplasm.

Genetic variation

The 38 accessions were categorized into three groups: Group A (East Asia) Group B (Europe and West Asia), Group C (North America) (Table 1). As there is only one accession from the Africa, it was not included in AMOVA. The majority of genetic variability occurred within accessions (84.3%) as compared to among accessions (15.7%). In addition, genetic variation was higher within regions (99.3%) than among regions (0.7%) (Table 3). Although variation among the three groups made up a small fraction of the total genetic variation, it was significant ($P < 0.001$) based on the random permutation test. Within-accession variation analyzed by each group was 68%, 83% and 93%, for the accessions from Group A, Group B and Group C,

respectively (Table 3). Based on the percentage of polymorphic loci, the five most diverse accessions were WY-PX2-94, Lupinella, LRC3901, PI568208, and CN31800, respectively. The five least diverse accessions were CN38745, Marinella, LRC3519, CN45949, and PI636518, respectively (data not shown).

Genetic association

Inter-accession genetic distances were statistically significant between any two accessions except for 14 pairs of accessions (data not shown). There are three clusters: Cluster I (PI636518 from China), cluster II (36 other accessions) and cluster III (PI568208 from Turkey) (Figure 2). Cluster II consisted of many sub-clusters, which were aligned together by genetic background and origin. The cultivars or breeding lines derived from United States (WY-PX2-94, Shoshone and SF-Laramie-73) were grouped with accessions from Ukraine (PI258773 and PI273791). This sub-cluster was loosely grouped with the other two cultivars from the United States (Renumex and Remont). Another distinct sub-cluster in dendrogram includes the oldest Canadian sainfoin cultivar Melrose aligned closely with Canadian accession (SL895) and accessions from Armenia (PI313056 and PI110404), and England (Giant). Another sub-cluster consists of Canadian cultivar Nova aligned together with Canadian accession (SCO38401), accessions from Kazakhstan (PI440577), France (CN31800) and Russia (PI502554). There was no distinct clustering of accessions from each of the three original groups. However, many accessions from Europe grouped together in a separate sub-cluster (PI319062, Dnepropetrovsk, Hampshire Common, and PI313066).

Phylogenetic analysis of the 367 individual plants using the NJ model with the phi-distance method is shown in Figure 3. Individuals from North America and Europe were distributed relatively uniformly around the NJ tree while individual plants from East or West Asia showed a

tendency to cluster together. Principal coordinates analysis revealed a distinctive cluster for accession PI636518 and PI568208, however the individual plants from other accessions showed an overlapping distribution (data not shown).

DISCUSSION

As sainfoin is a relatively less developed forage legume with limited genetic information available, this study generated a set of useful genetic information for its breeding. Five AFLP primer combinations produced clear polymorphic marker profiling of sainfoin accessions. Significant genetic variation existed for within-accession, among accessions and among groups. The accessions PI636518 and PI568208 were found to be the most genetically diverse accessions. The majority of accessions were grouped together in sub-clusters based on cultivar pedigree. This information is useful for an initial selection for genetically diverse parent populations for crossing and creating genetically diverse populations.

Our results showed high within-accession genetic variation (84%), which was similar to the high within-accession variation observed for wild germplasm of sainfoin (Nosrati et al. 2012) and regional (Iran) accessions of sainfoin (Rasouli et al. 2013; Zarrabian et al. 2013; Hejrankesh et al. 2014). High levels of within-accession variations based on AFLPs was also reported for other out-crossing perennial forage species such as perennial ryegrass (Guthridge et al. 2001), white clover [*Trifolium repens* L.] (Kolliker et al. 2001), fringed brome [*Bromus ciliates* L.] (Fu et al. 2005), plains rough fescue [*Festuca hallii* (Vasey) Piper.] (Qui et al. 2007), and side-oat grama (Biliget et al. 2013). Hamrick and Godt (1989) showed that out-crossing plant species tend to exhibit high within-accession (80-90%) genetic variation. Sainfoin is an insect pollinated perennial forage legume, which is distributed worldwide from sub-tropical to temperate regions of the world. Our results are aligned with the expectation for a high percentage of outcrossing in

this species and consistent with the previous finding of self-fertilization rate of 0.98% and 5.1% under plant isolation and manual pollination conditions, respectively (Thomson 1938). Interestingly, Kempf et al. (2015) reported a self-fertilization rate of 3.9% under non-directed pollination, but up to 65% self-fertilization in an artificially directed pollination population. Also, this study revealed a higher within-accession genetic variation for the accessions from the North America (93%) and Europe and West Asia (83%) than the accessions from East Asia (68%). Further studies with large sample sizes are needed to clarify the discrepancy.

PI636518 from China and PI568208 from Turkey were the most different lines based on UPGMA analysis. This finding is consistent with the phenotypic characteristics observed in our field study (data not shown). Plants from accession PI636518 (cluster I) had short stems, small seeds, and low forage biomass production. Accession PI568208 (cluster III) had glossy leaves and a long spiked seed pod. Rest of the 36 accessions (cluster II) had erect growth habit (except PI494667), green leaves and higher biomass yield. In addition, the composition of many sub-clusters in the dendrogram could be explained, in part, by their pedigrees and origins. For example, the majority of the accessions from the United States clustered closely together reflecting their common ancestry. For example, cultivar Renumex was developed from Remont (USDA 2015) and WY-PX2-94 is the second generation breeder class seed of Shoshone (Gray 2004). Similarly, accessions LRC3519 and LRC3901, which are sister breeding lines from Lethbridge, Canada showed a non-significant difference in genetic distance. In contrast, the accessions from England (Giant and Hampshire Common) were clustered distinctly apart. This was likely because the cultivar Giant was developed from accessions from the Middle East, while Hampshire Common originated from central Europe and clustered closely with accessions from Bulgaria (PI313066) (Hayot-Carbonero et al. 2011). The Canadian accessions (Nova and

SCO38401) were clustered together with an accession from Kazakhstan (PI440577) supporting the fact that Nova originated from Kazakhstan (Hanna 1980). The Canadian sainfoin cultivar Melrose was developed from Russian lines (Hanna et al. 1970) and Melrose clustered with accessions from former USSR (Armenia) (PI110397 and PI110404). Thus, the research results on genetic relationships are useful for the development of synthetic sainfoin cultivars. Future sainfoin breeding could potentially achieve more heterosis by crossing genetically distinct accessions such as PI636518, PI568208, CN38745, Marinella and PI494667.

Acknowledgments

The authors would like to thank Ms. Carolee Horbach, Mr. Gregory Peterson and Ms. Vicky Roslinsky at the Agriculture and Agri-Food Canada (AAFC) Saskatoon Research and Development Centre for their technical assistance in AFLP analysis; and Dr. Fangqin Zeng and Mr. Kiran Baral for their useful comments on the earlier version of this manuscript. The research was funded by the Saskatchewan Forage Network, the Saskatchewan Ministry of Agriculture and the Canada-Saskatchewan Growing Forward 2 Bilateral Agreement.

References

- Avci, S., Ilhan, E., Erayman, M., and Sancak, C. 2014. Analysis of *Onobrychis* Genetic diversity using SSR markers from related legume species. J. Anim. Plant Sci. **24**: 556–566.
- Awasthi, A.K., Nagaraja, G.M., Naik, G.V., Kanginakudru, S., Thangavelu, K., and Nagaraju, J. 2004. Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. BMC Genet. **5**: 1–9.
- Bhattarai, S., Coulman, B., and Biliget, B. 2016. Sainfoin (*Onobrychis viciifolia* Scop.): Renewed interest as a forage legume for western Canada. Can. J. Plant Sci. **96**: 748–756.
- Biliget, B., Schellenberg, M.P., and Fu, Y.B. 2013. Genetic diversity of side-oats grama [*Bouteloua curtipendula* (Michx.) Torr.] populations as revealed by amplified fragment length polymorphism markers. Can. J. Plant Sci. **93**: 1105–1114.
- Burton, J.C., and Curley, R.L. 1968. Nodulation and nitrogen fixation in sainfoin (*Onobrychis sativa* LAM.) as influenced by strains of rhizobia. In: Cooper, C.S. and Carleton, A.E. eds. Sainfoin Symposium. Mont., Agric. Exp. Stn. Bull. **627**: 3–5.
- Demdoun, S., Munoz, F., Delgado, I., Valderrabano, J., and Wunsch, A. 2012. EST-SSR cross-amplification and genetic similarity in *Onobrychis* genus. Genet. Resour. Crop Ev. **59**: 253–260.
- Excoffier, L., Laval, G., and Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol. Bioinform. **1**: 47–50.

Ferdinandez, Y.S.N., and Coulman, B.E. 2002. Evaluating genetic variation and relationships among two brome grass species and their hybrid using RAPD and AFLP markers. *Euphytica*. **125**: 281–291.

Frame, J. 2005. Forage legumes for temperate grasslands. FAO. Science Publishers, Inc., Enfield, NH, USA. 309pp.

Frame, J., Charlton, J.F.L., and Laidlaw, A.S. 1998. Temperate forage legumes. CAB International. Wallingford, UK. 273–289.

Fu, Y.B., Coulman, B.E., Ferdinandez, Y.S.N., Cayouette, J., and Peterson, P.M. 2005. Genetic diversity of fringed brome (*Bromus ciliatus*) as determined by amplified fragment length polymorphism. *Can. J. Bot.* **83**: 1322–1328.

Fu, Y.B., Phan, A.T., Coulman, B., and Richards, K.W. 2004. Genetic diversity in natural populations and corresponding seed collections of little Bluestem as revealed by AFLP markers. *Crop Sci.* **44**: 2254–2260.

Goplen, B.P., Richards, K.W., and Moyer, J.R. 1991. Sainfoin for western Canada. Agriculture Canada Publication. 1470/E.

Gray, F.A. 2004. Released notice of ‘Shoshone’ sainfoin. Wyoming Agricultural Experiment Station. [Online] Available:
<http://plantsciences.montana.edu/foundationseed/documents/varietyrelease/2005/Shoshone.pdf>
[24 Nov. 2016].

- Guthridge, K.M., Dupal, M.P., Kolliker, R., Jones, E.S., Smith, K.F., and Forster, J.W. 2001. AFLP analysis of genetic diversity within and between populations of perennial ryegrass (*Lolium perenne* L.). *Euphytica*. **122**: 191–201.
- Hamrick, J.L., and Godt, M.J.W. 1989. Allozyme diversity in plant species. Pages 43–63. *In* A.H.D. Brown, M.T. Clegg, A.L. Kaher, and B.S. Weir, eds. Plant population genetics, breeding and genetic resources. Sinauer Associates, Sunderland, MA.
- Hanna, M.R. 1972. Sainfoin for western Canada. Dept. Agr. Pub. 1470. 1–18.
- Hanna, M.R. 1980. Nova sainfoin. *Can. J. Plant Sci.* **60**: 1481–1483.
- Hanna, M.R., Cooke, D.A., and Goplen, B.P. 1970. Melrose Sainfoin. *Can. J. Plant Sci.* **50**: 750–751.
- Harris, K., Anderson, W., and Malik, R. 2010. Genetic relationships among napiergrass (*Pennisetum purpureum* Schum.) nursery accessions using AFLP markers. *Plant Genet. Resour.* **8**: 63–70.
- Hayot-Carbonero, C. 2011. Sainfoin (*Onobrychis viciifolia*), a forage legume with great potential for sustainable agriculture, an insight on its morphological, agronomical, cytological and genetic characterization. Ph.D. Dissertation, University of Manchester, Manchester, U.K. 85pp.
- Hayot-Carbonero, C., Mueller-Harvey, I., Brown, T.A., and Smith, L. 2011. Sainfoin (*Onobrychis viciifolia*): a beneficial forage legume. *Plant Genet. Resour.-C.* **9**: 70–85.

- Hejrankesh, N., Ali-Mousavizadeh, S., Haghghi, A.R., and Rashidi, V. 2014. Evaluation of genetic diversity of sainfoin (*Onobrychis viciifolia* Scop.) landraces using RAPD markers. *J. Curr. Res. Sci.* **2**: 739–748.
- Koopman, W.J.M. 2005. Phylogenetic signal in AFLP data sets. *Syst. Biol.* **54**:197–217.
- Kempf, K., Grieder, G., Walter, A., Widmer, F., Reinhard, S., and Koelliker, R. 2015. Evidence and consequences of self-fertilisation in the predominantly outbreeding forage legume *Onobrychis viciifolia*. *BMC Genet.* **16**:117.
- Kempf, K., Mora-Ortiz, M., Smith, L.M.J., Kolliker, R., and Skot, L. 2016. Characterization of novel SSR markers in diverse sainfoin (*Onobrychis viciifolia*) germplasm. *BMC Genet.* **17**:124.
- Kolliker, R., Jones, E.S., Jahufer, M.Z.Z., and Forster, J.W. 2001. Bulkied AFLP analysis for the assessment of genetic diversity in white clover (*Trifolium repens* L.). *Euphytica*. **121**: 305–315.
- Kumar, S., Tamura, K., and Nei, M. 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief. Bioinform.* **5**: 150–163.
- Loh, J.P., Kiew, R., Set, O., Gan, L.H., and Gan, Y.Y. 2000. Amplified fragment length polymorphism fingerprinting of 16 banana cultivars (*Musa* cvs.). *Mol. Phylogenet. Evol.* **17**: 360–366.
- Majidi, M.M., Mirlohi, A., and Amini, F. 2009. Genetic variation, heritability and correlations of agro-morphological traits in tall fescue (*Festuca arundinacea* Schreb.). *Euphytica*. **167**: 323–331.

- Mechanda, S.M., Baum, B.R., Johnson, D.A., and Arnason, J.T. 2004. Sequence assessment of comigrating AFLP bands in Echinacea—implications for comparative studies. *Genome* **47**:15–25.
- Mora-Ortiz, M., Swain, M.T., Vickers, M.J., Hegarty, M.J., Kelly, R., Smith, L.M.J., and Skot, L. 2016. De-novo transcriptome assembly for gene identification, analysis, annotation, and molecular marker discovery in *Onobrychis viciifolia*. *BMC Genomics*. **17**:756.
- Nosrati, H., Feizi, M.A.H., Tarrah, S.S., and Haghighi, A.R. 2012. Population genetic variation in sainfoin (*fabaceae*) revealed by RAPD markers. *Tom XIX Issue:1*, 11–16. [Online] Available: <http://www.bioresearch.ro/bioresearch/revistaen.html> [3 July 2015].
- Peng, Y., Zhang, X., Deng, Y., and Ma, X. 2008. Evaluation of genetic diversity in wild orchardgrass (*Dactylis glomerata* L.) based on AFLP markers. *Hereditas*. **145**: 174–181.
- Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., and Rafalski, A. 1996. The utility of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* **2**: 225–238.
- Qui, J., Fu, Y.B., Bai, Y., and Wilmshurst, J.F. 2007. Patterns of amplified restriction fragment polymorphism in natural populations and corresponding seed collections of plains rough fescue (*Festuca hallii*). *Can. J. Bot.* **85**: 484–492.
- Rasouli, M., Jafari, A.A., Tabaei-Aghdaei, S.R., Shanjani, P.S., and Darvish, F. 2013. Assessment of genetic variability of 36 populations of sainfoin (*Onobrychis viciifolia* Scop.) based on RAPD markers. *Int. J. Biosci.* **3**: 15–26.

- Rohlf, F. J. 1997. NTSYS-pc 2.1: Numerical taxonomy and multivariate analysis system. Exeter Software, New York, NY.
- Roldan-Ruiz, I., Dendauw, J., Van Bockstaele, E., Depicker, A., and De Loose, M. 2000. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). *Mol. Breed.* **6**: 125–134.
- Swofford, D. L. 1998. PAUP*. Phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland, MA.
- Thomson, J.R. 1938. Cross- and self-fertility in sainfoin. *Ann. Appl. Biol.* **25**:695–704.
- USDA NRCS. 2015. The PLANTS Database. National Plant Data Center, Baton Rouge, LA 70874–4490 USA. [Online] Available: <http://plants.usda.gov> [11 July 2015].
- Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van De Lee, T., Homes, M., Frijters, A., Pot, J., Peleman, J., and Kuiper, M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407–4414.
- Williams, J.G.K, Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.* **18**: 6521–6535.
- Yildiz, B., Ciplak, B., and Aktoklu, E. 1999. Fruit morphology of sections of the genus *Onobrychis* Miller (Fabaceae) and its phylogenetic implications. *Isr. J. Plant Sci.* **47**: 269–282.
- Zarrabian, M., and Majidi, M.M. 2015. Genetic diversity and relationships within and among *Onobrychis* species using molecular markers. *Turk. J. Bot.* **39**: 681–692.

Zarrabian, M., Majidi, M.M., and Ehtemam, M.H. 2013. Genetic Diversity in a Worldwide Collection of Sainfoin Using Morphological, Anatomical, and Molecular Markers. *Crop Sci.* **53**: 2483–2496.

For Review Only

Table 1. Information for the 38 sainfoin accessions studied by AFLP analysis.

Accessions	Location	Continent	Group ^a	Code	NP ^b	Type of germplasm
PI440577	Astana, Kazakhstan	East Asia	A	20	10	Wild type
CN45949	Xinjiang, China	East Asia	A	7	6	Landrace
PI577670	Xinjiang, China	East Asia	A	31	10	Landrace
PI636518	Xinjiang, China	East Asia	A	1	10	Wild type
PI110397	Armenia	West Asia	B	24	10	-
PI110404	Armenia	West Asia	B	32	10	-
PI568208	Afyon, Turkey	West Asia	B	22	6	Wild type
CN38745	Ordu, Turkey	West Asia	B	17	10	Landrace
PI380949	Zanjan, Iran	West Asia	B	13	9	Landrace
Lupinella	Bologna, Italy	Europe	B	11	10	Cultivar
PI313066	Bulgaria	Europe	B	26	10	-
PI319062	Burgos, Spain	Europe	B	5	10	-
Visnovsky	Czech Republic	Europe	B	4	10	Cultivar
Giant	England	Europe	B	29	10	Cultivar
Hampshire						
Common	England	Europe	B	23	10	Cultivar
CN31800	France	Europe	B	34	10	Wild type
PI502554	Kalmykia, Russia	Europe	B	30	10	Landrace
PI313056	Norway	Europe	B	28	10	-
PI313049	Poland	Europe	B	8	10	-
Marinella	Slovakia	Europe	B	3	10	Cultivar
Svkpiestt	Slovakia	Europe	B	25	10	Cultivar
PI494667	Suceava, Romania	Europe	B	38	10	Wild type
PI318606	Switzerland	Europe	B	12	10	Landrace
Dnepropetrovsk	Ukraine	Europe	B	6	10	Landrace
PI258773						
(Miatiletka)	Ukraine	Europe	B	2	10	Cultivar
PI273791	Ukraine	Europe	B	35	10	-
LRC 3519	Lethbridge, Canada	North America	C	14	6	Breeding line
LRC 3901	Lethbridge, Canada	North America	C	15	10	Breeding line
Nova	Lethbridge, Canada	North America	C	33	10	Cultivar
Melrose	Saskatoon, Canada	North America	C	21	10	Cultivar
SL895	Saskatoon, Canada	North America	C	27	10	Breeding line
SCO38401	Swift Current, Canada	North America	C	36	10	Breeding line
Remont	Montana, United States	North America	C	19	10	Cultivar
Renumex	New Mexico, United	North America	C	16	10	Cultivar

States						
SF-Laramie-73 Shoshone	Wyoming, United States	North America	C	37	10	Breeding line
	Wyoming, United States	North America	C	10	10	Cultivar
WY-PX2-94 PI338651	Wyoming, United States	North America	C	9	10	Breeding line
	Morocco	Africa	-	18	10	-

^aAccessions were separated into groups according to their locations.

^bNumber of plant assayed.

For Review Only

Table 2. AFLP polymorphisms generated among 367 sainfoin plants of 38 accessions by five AFLP primer pairs.

Primer pairs	AFLP bands scored	PIC ^a	Frequency of scored bands		
			Mean	Maximum	Minimum
E+AAG/M+CAC	222	0.186	0.186	0.978	0.003
E+ACA/M+CAG	261	0.167	0.165	0.970	0.003
E+ACA/M+CAT	244	0.180	0.196	0.973	0.003
E+ACC/M+CTC	166	0.129	0.126	0.967	0.003
E+AGC/M+CAC	149	0.169	0.151	0.976	0.003
All	1042	0.166	0.165	0.973	0.003

^aPolymorphism information content.**Table 3.** Analysis of molecular variance (AMOVA) based on AFLP data generated among 38 sainfoin accessions collected from 21 different countries belonging to three groups.

Source of variation		Degree of freedom	Sum of square	Variance component	Distribution of variation (%)	Probability
All	Among Accessions	37	7420.9	13.35	15.71	<0.001
	Within Accessions	329	23569.7	71.64	84.29	
	Total	366	30990.6	84.99		
Group ^a	Among Regions	2	588.9	1.87	2.19	<0.001
	Within Regions	354	29483.6	83.29	97.81	
	Total	356	30072.5	85.16		
Group A ^b	Among Accessions	3	913.9	27.63	31.89	<0.001
	Within Accessions	32	188.1	59.00	68.11	
	Total	35	2802.0	86.63		
Group B	Among Accessions	22	4700.4	14.57	17.00	<0.001
	Within Accessions	202	14369.0	71.13	83.00	
	Total	224	19069.4	85.70		
Group C	Among Accessions	10	1299.4	5.50	6.67	<0.001
	Within Accessions	95	7312.6	76.97	93.30	
	Total	105	8612.0	82.47		

^aThe accession PI338651 (Morocco, Africa) was excluded in the group analysis.^bGroup A: East Asia; Group B: West Asia and Europe; Group C: North America.

Figure Captions

Figure 1 Geographic location of collected 38 accessions representing 21 different countries from four groups (North America, green; Europe and West Asia, sky blue; Africa, blue; East Asia, red).

Figure 2 AFLP-based genetic relationships among the 38 sainfoin accessions developed by UPGMA cluster analysis using the phi-statistics as inter-accession genetic distance.

Figure 3 Genetic relationships among 367 sainfoin individuals of 38 accessions representing four continents, as revealed by AFLP-based Neighbor-joining (NJ) tree. Individuals from Group A (red), Group B (black), Group C (green), and Africa (blue) are highlighted with different colors. Each individual plant is labeled with accession code (1-38) followed by individual plant number (1-10).

Figure 1

44x20mm (300 x 300 DPI)

52x29mm (300 x 300 DPI)

Figure 3

89x94mm (300 x 300 DPI)