Journal of Crop Science
ISSN: 0976-8920 & E-ISSN: 0976-8939, Vol. 2, Issue 1, 2011, PP-11-20
Available online at http://www.bioinfo.in/contents.php?id=65

A STRATEGY FOR DEVELOPMENT OF MICROSATELLITE MARKERS FOR LESS STUDIED ORGANISMS BASED ON EST-SSRS OF WELL STUDIED ORGANISMS WITHIN SAME FAMILY: CASE STUDY ON (*PROSOPIS JULIFLORA* L.) LEGUME PLANT

ADNAN A.S. AL-ASBAHI[1*] and FATIMA AHMED. ALHADI[2]
[1]Department of Biology (Biotechnology and Molecular Genetics), Faculty of Sciences, Sana'a University, P. O. Box 14686, Sana'a, Republic of Yemen
[2]Department of Biology (Plant Physiology-Ecology), Faculty of Sciences, Sana'a University, Sana'a, Republic of Yemen
*Corresponding author. E-mail: adnanasbahi@yahoo.com

Received November 20, 2010; Accepted December 14, 2010

Abstract-It is widely known that publicly available expressed sequence tags (ESTs) markers for a certain organism, are used for development of simple sequence repeats (SSRs) for that organism. SSRs have been used in a wide range of applications in genetic studies. However, many plants, animals and microbe species have only limited number of (ESTs) markers available in the databases such as, mesquite desert plant species (*Prosopis juliflora*). Available small amount of ESTs for *Prosopis juliflora* can be compared with those of well-studied legume species in both less and well-studied species for further analysis and usage as traditional SSR markers for genomic comparative studies with minimal amount of money and time required for genotyping accompanied with high level of transferability and polymorphism than what expected by using anonymous SSR or ESTs derived SSR randomly selected from closely related species or genera. *Prosopis juliflora* ESTs downloaded from Genbank were clustered into different unigenes, each of which represents ESTs of same and relevant function. Then, they were analyzed for the presence of various SSR classes in these EST unigenes knowm as (EST-SSRs) unigenes in order to assess sequence, length, repetitious folds, distribution and relative abundance of different SSR motifs of mono, di-, tri-, tetra-, and penta-nucleotide contained in these ESTs. Comparative analysis of (EST-SSRs) in *Prosopis juliflora* with those of well-studied legumes, such as *Medicago truncatula*, *Glycine* max and *Lotus japonicus* resulted in identification the SSR repeat motif categories existing in high frequency rates within those legume species. These SSR regarded as promising SSRs markers used as cross genera were tested in *P. juliflora* based on primer pairs of anonymous SSR of the well-studied species, *Medicago,* through PCR that have shown transferability of these markers into *P. juliflora* as cross-genera SSR markers. The developed anonymous SSR markers of loci across family would offer an opportunity in cheap and fast genotyping of less studied organisms such as *Prosopis* legume species that suffers from poor genome resources.
Keywords-*Prosopis juliflora*, Expressed Sequence Tag (EST), Simple Sequence Repeat (SSR), CAP3 software, Unigenes

General background

Simple sequence repeats (SSRs) markers, known as Microsatellites, exist both in coding and non-coding genomic DNA [1, 2,] in range of 1 to 6 bp repeat motifs and differ from genotype to genotype due to high mutation rate in these motifs associated with unequal crossing and replication slippage [3, 4]. Moreover, SSR reproducibility, multi-allelic nature, co-dominancy and abundance [5] together with their simple detection and high polymorphism [6] make them attractive as a powerful molecular approach of choice for diversity assessment [7] as well as for variety protection, variety characterization, molecular mapping, and marker assisted selection that links between phenotypic and genotypic variation [8].

In the past, SSRs were either developed from SSR-enriched libraries or identified by surveys of genomic sequence databases [9, 10]. However, development of SSR markers derived from genomic libraries is expensive as library construction and sequencing of clones is required [11, 12, 13]. Moreover, limited sizes of the available sequence data restricted development of SSRs [14], so that, status influencing plant breeding programs has not been solved for so far. Although, sources alternatives of microsatellites come mainly from data mining sequence information from the publicly available (EST) databases of important well studied organisms [15, 16, 17, 18], which are currently

known as the most widely sequenced nucleotide count [19]. Consequently, many recent studies have based upon publicly available (ESTs) of well-characterized taxa for the efficient development of large numbers of so-called EST-SSRs for that proved very useful for genetic studies in closely related taxa [20, 21, 22]. However, transferability of these EST-based SSRs developed based on conserved orthologous sets (COS) markers considered for use across genera [23]. However, the lack of polymerase chain reaction (PCR) suitable primers to be used for amplification of SSRs across species as they are frequently species-specific, meaning that markers developed in one taxon cannot be readily transferred to another [24, 25, 26, 27, 28]. These problems will be even more difficult to solve when intending development of SSRs for genotyping utilization of a less studied organisms based on bioinformatics publicly available (ESTs) data from other genera within a family, such as, the desert plant, *P. juliflora,* which is the main problem discussed in this paper.

In this study, we have examined non-redundant ESTs derived SSRs of *P. juliflora,* against their counterpart non-coding SSR markers in *M. truncatula* genomic which in turn have been previously tested in other legume *Medicago* species, in order to identify redundant SSRs among legume genera that also appeared in high frequency rates among other legumes. These non-coding SSR, known as traditional or anonymous SSRs transferability and polymorphisms across genera within legume family are possible and hence, the previous drawbacks associated with less polymorphism or less transferability based on utilization of conserved EST-SSRs or lack of (COS) within legume genera respectively.

To examine this idea, the main objectives undertaken in the present paper are; (a) Mining the all available ESTs collections of *P. juliflora* (less studied species) for non-redundant EST-SSRs, and their motif sequences frequency and distribution across less and well-studied legume genera involving, in *P. juliflora* and *M. truncatula* (b) Motifs with high frequency rate among various genera, will be the reference for the selected genomic microsatellite markers (SSRs) of *M. truncatula* to screen *P. juliflora*, (c) validate selected SSRs through PCR amplifications experiments for transferability into *P. juliflora.*

Materials and Methods

Part I; Bioinformatics Mining

All publicly available *P. juliflora* ESTs that are (1,470), were downloaded from national center of biotechnology institute databases (NCBI) dbEST division at (http:// www.ncbi.nlm.nih.gov/dbEST/index.html) in October of 2010 in FASTA file format. Cluster analysis was employed to define groups of overlapping EST sequences representing the same native transcripts as putative non-redundant (unigenes or singletons) that are also aligned to remove Poly Adenine tails formed post-transcription of these genes by using free license CAP3 [29] available at http://mobyle.pasteur.fr. Consigns of *P. juliflora* are screened for the internal SSRs detection, length, abundance, and distribution by using Perl Script based software available at http://www.shrimp.ufscar.br/cid/. They are searched for all possible combination of mono-, di-, tri-, tetra-, penta- and hexa-nucleotide repeat motifs whose minimum repeat units are defined as ten, six, for mono and di-nucleotide motifs respectively and five for the others. The most frequent mers identified in ESTs derived SSRs are compared with their non-coding SSR counterparts of the well-studied plant species belonging to different genera within legume family including, *M. truncatula*, Soybean and *Lotus japonicus* in order to identify the motif category that relatively showed high frequency rates in *P. juliflora* genome and among other legume species SSRs.

Part II; Plant growth, DNA isolation, PCR amplification, and anonymous SSR marker screenin

Plant seed were treated with 10% sulfuric acid for one hour to enhance its growth on autoclaved soil in which seeds lasts for 30 days before emerging enough tissues for DNA extraction. Genomic DNA was extracted from leaf tissues using CTAB protocol previously described by [30, 31] approximately 100 mg of leaf tissue, frozen in liquid nitrogen, was ground to a final powder with mortar and pestle. Leaf tissue was frozen in liquid nitrogen and completely ground in 2ml eppendorf tube. Preheated CTAB buffer lyses was added and mixed well until homogenized and then incubated in water bath at 65°C for 1 hour. 700µL of chloroform isoamylalcohol (24:1) was added and then mixed well and centrifuged at 12000rpm for 10. Then 400µL of upper layer was transferred and was added with 400µL isopropanol. Then DNA was washed in 70% ethanol. The DNA was dried and dissolved in Tris-EDTA (TE) buffer, and the concentration was quantified in a spectrophotometer (NanoDrop Technologies, Wilmington, USA) at 260 nm, and kept at) -80 Celsius for storage.

Forty-six SSR primer pairs were selected based on frequency of SSRs motifs categories corresponding to those non-coding SSR markers previously reported as cross species markers within a legume reference, *M. truncatula* species (table. 1) were used for screening of *P. juliflora* genomic DNA by PCR technique for verification its transferability and hence utility as across genera markers that can save time and money paid for identification of SSR in less studied species having a small amount of

genome information data. Each PCR reaction was conducted in 20 µl volume containing 5 ng of Maguire's genomic DNA template, final concentrations of 0.4 pM of both forward and reverse primers, 2 mM of $MgCl_2$, 0.1 mM of dNTPs, 1X Tris-acetate-EDTA (TAE) buffer, 40 mM Tris acetate, pH approx. 8.3, containing 1 mM EDTA, and 0.1 U of AmpliTaq polymerase enzymes, and diluted to volume with double-distilled water. The cycling conditions for PCR on a GeneAmp® PCR System 9700 (PE-Applied Biosystems) thermal cycler were optimized to an initial denaturation of 5 min at 94 C°, followed by 35 cycles of 94 C° for 1min, annealing temperatures, ranging from 50 to 55 C° for 1 min, and extension at 72 C° for 1 min. This was followed by final extension of 10 min at 72 C°. PCR products were separated electrophoretically on high resolution 4% agarose gels, visualized by ethedium bromide staining and scored manually.

Results

The results have shown 296 non-redundant ESTs representing putative unigenes of a redundancy of 54. 2 %. Trim-EST program of EMBOSS package within CAP3 used to exclude poly adenine tails in these unigenes dataset left with a total base pair of 2159430 bp, was then used as a source for remarkable SSR detection and frequency identification. Out of 296 putative unigenes, only 75 have internal SSR, known as (EST- SSRs) with total sequence size of 143022 bp. Number of SSR containing sequences are 65, while number of sequences containing more than 1 SSR are 8 and the SSRs present in compound formation are 5. The frequency SSR distribution is 1SSR/1.907 kb. Identified SSRs are clustered into 13 different types of motifs constructed for mono, di-, and tri-nucleotide SSRs that are; (A/T), (AG/CT), (AT/AT), (CA/TG), (AC/GT), (GA/TC), (TA/TA), (AAG/CTT), (AAT/ATT), (TAA/TTA), (AGG/CCT), (GAG/CTC), (TCC/GGT), (AAC/TTG), (ACG/TGC) as it is shown in (table 2) and (figure 1).

Each of the eight ESTs derived SSRs categories found in *Prosopis* was compared with its corresponding SSR category in other legume generas such as, *Medicago truncatula*, *Glycine max* and *Lotus japonicus*, that have shown various similarities in of SSR frequency rates among legumes genera as it is shown in (table 2) and (figure 2). 46 primer pairs of *Medicago truncatula* cross species anonymous SSR markers (table 1) that also involve previous SSR categories, have been tested in *P. juliflora* genomic DNA, through PCR reaction. There is only three markers have shown amplification patterns illustrated in (figure 4) showing the amplified DNA bands run on 4% agarose gel that include mtgsp_003F07 and mtgsp_001G08 out of 5 markers belong to AT category as well as mtgsp_001c10 out of 4 markers belong to AG category.

Discussion

Since, costs and time are limiting factors determining the effectiveness and usefulness of a particular approach to be considered in development of SSR markers. Accordingly, EST derived SSR markers (EST-SSRs) have gained popularity among many researchers as a result of their capability in reducing the amount of time and money needed for genotyping. However, they have shown lower level of polymorphism associated with higher transferability in number of taxa within a single species [32, 33, 34] when compared to anonymous SSRs [20, 35]. Moreover, their utilization in comparative studies of self-pollinated species included in legume family is associated with several limitations. For example, conserved sequences in is significantly will reduce polymorphism of EST-SSRs versus SSRs [36]. In addition, as ESTs are derived from coding regions, large introns as well as unrecognized intron splice sites could disrupt priming sites, resulting in null alleles similar to a problem for their anonymous counterparts SSRs [22, 37,38]. On the other hand, anonymous SSRs have low cross species transferability caused by frequent mutation and replication slippage [15, 39] as well as increases a nul allele occurrence [40] restricting their utilization and effectiveness power in genotyping.

The present study provides a new strategy, that able to combine advantages and of both, EST-SSRs and anonymous SSRs (derived from non-coding DNA regions), markers respectively. The created marker will also exclude the disadvantages of both. This strategy rely on; firstly, identification of SSR categories (motif sequence) having relatively high frequency rate in ESTs nucleotide sequences of the less studied *P. juliflora* to be compared with the corresponding frequency rates of well-studied legumes genera such as reference model *M. truncatula, Glycine max* and *Lotus japonicus* legume genera in as shown in (table 2 and figure 2 and 3). Secondly, SSR motif sequence with frequency rate located within less and well-studied legumes' average rate such as, markers having the repeat motifs (AG/CT) and (AT/AT) are the most possible promising markers to be tested by PCR in tested *P. juliflora* to confirm the previous idea. The results gained in this study have supported the previous idea. Because, results shown high transferability rate of *M. truncatula* anonymous SSRs markers' having those repeat motifs shared by less and well-studied legumes' genera (figure 4). However, there is no transferable markers belong to the other SSRs with class motif (SSRs categories) that exist in *P. juliflora* either with low frequency rate or less than corresponding average rate in other legumes (figure 3). However, transferability of markers having repeat motifs didn't exist in *P.*

juliflora EST-SSRs such as AAC/TTG is completely absent (figure 1, 2 &3). Similarly, markers with low frequency repeat motifs no longer transferable to in *P. juliflora* in spite of their high frequency rate in other legumes. The adopted strategy in this study based on EST-SSRs motif frequency comparisons between less and well-studied individuals from different genera within a single family, can be regarded as predictive indicators for transferable anonymous SSRs markers selected from rich sequence data genera to screen another genera with poor sequence data. They can also serve as guidelines for motif type and starting number of SSR markers to be used for genotyping at family level with great efficiency and less time and money requirements when compared with another approaches based on EST-SSR or traditional SSR markers alones that are associated with low polymorphism and transferability either within and among unrelated species. This is due to that, SSR motifs are highly variable among, loci and among alleles [41], so that, they will serve as polymorphic markers across species, and genera. In addition, these kinds of markers would have a potential utilization for assessment of syntenic relationships among genera within legume family. Moreover, development of SSR markers by following this strategy is straightforward solution for less studied species based on ESTs databases of well-studied species within a family.

Conclusion

Available public ESTs data of the less studied legumes, *P. juliflora,* analyzed for SSR content and distribution, and then compared with available data of three other legumes including *M. truncatula* resulted in identification of di-nucleotide repeats as the shared class of repeats among and within within legumes. The SSR markers of the model legume, *M. truncatula*, known as cross species that are also containing repeat motifs, (AT/TA and AG/CT) that are shared by many legumes' genera, were further used to predict of cross genera markers based on their primer pairs used through PCR for verification of possible conservation of these markers as cross genera within legume family. The approved transferable SSRs then, will represent potentially valuable markers for population genetic analyses at the level of species, genera and family. They will also offer a number of clear benefits, including rapid and inexpensive development, without the known drawbacks of EST-SSRs and coding SSR markers. Therefore, the strategy we used for development of theses markers will represent the best way forward for the analysis of species for which only limited resources are available.

Acknowledgement

The authors thanks Georg Jander and his lab members at Boyce Thompson Institute for Plant Researches (BTI), Ithaca, NY, USA for their contributions and support funding for lab supplies from National Science Foundation grant IOS-0718733. The first author also thanks the Center for International education of Scholars, USA for fellowship support at Cornell University and BTI.

References

[1] Ellegren H. (2004) *Nature Reviews Genetics*, 5(6), 435–445.
[2] Natalya S. (2008) *Genome*, 51(2), 79-90.
[3] Richards R.I., and Sutherland G.R. (1992) *Nature Genetics*, 1(1), 7–9.
[4] Weber J.L. (1990) *Genomics*, 7(4), 524–530.
[5] Wang H., Wei Y., Yan Z. and Zheng Y. (2007) *Journal of Applied Genetics*, 48(1), 35–42.
[6] Powell W., Machray G. and Provan J. (1996). *Trends in Plant Science*, 1(7), 215–222.
[7] Bernadette J., Sandrine F., Philippe B., Gaelle C., Sylvain S., Thierry H., and Christian H., (2003) *BMC Plant Biology*, 3 (2-3), 1-19.
[8] Varshney R., Sigmund R., Borner A., Korzun V., Stein N., Sorrells M., Landridge P. and Graner A. (2005) *Plant Science*, 168 (x), 195–202. doi:10.1016/j.plantsci.2004.08.001
[9] Akkaya M.S., Bhagwat A.A. and Cregan P.B. (1992) *Genetics*, 132(4), 1131–1139.
[10] Bell C.J. and Ecker J.R. (1994) *Genomics,* 19(1), 137–144.
[11] Brouwer D.J. and Osborn T.C., (1999) *Theoretical and Applied Genetics*, 99(7-8), 1194-1200.
[12] Wu K.K., Burnquist, W., Sorrells M.E., Tew T.L., Moore P.H. and Tanksley S.D. (1992) *Theoretical and Applied Genetics*, 83 (3), 294-300.
[13] Morgante M. and Olivieri A.M. (1993) *Plant journal*, 3 (1), 175–182
[14] Anderson J. A., Churchill G.A., Autrique J.E., Tanksley S.C. and Sorrells M.E. (1993). *Genome*, 36(1), 181–186.
[15] Kantety R.V., La Rota M., Matthews D.E. and Sorrells M.E. (2002) *Plant Molecular Biology*, 48(5-6), 501-510.
[16] Thiel T., Michalek W., Varshney R.K., and Graner A. (2003) *Theoretical and Applied Genetics*, 106 (3), 411–422.
[17] Saha M.C., Mian M.A., Eujayl I., Zwonitzer J.C., Wang L., May G.D. (2004) *Theoretical and Applied Genetics*, 109 (4), 7 83–791.
[18] Asp T., Frei U.K., Didion T., Nielsen K., and Lübberstedt T. (2007) *BMC Plant Biology*, 7 (36). doi, 10.1186/1471-2229-7-36.

[19] Natalya S. (2008) Genome, 51(2), 79-90.
[20] Gupta P., Rustgi S., Sharma S., Singh R., Kumar N. and Balyan H. (2003) *Molecular Genetics and Genome*, 270 (4), 315–323.
[21] Bhat P.R., Krishnakumar V., Hendre P.S., Rajendrakumar P., Varshney R.K. and Aggarwal R.K. (2005) *Molecular Ecology Notes*, 5, 80–83. doi: 10.1111/j.1471-8286
[22] Varshney R.K., Graner A. and Sorrells M.E. (2005) *Trends in Biotechnology*, 23 (1), 48–55. doi, 10.1016/j.tibtech.2004.11.005.
[23] Parida S.K., Anand R., Kumar K., Dalal V., Singh N.K. and Mohapatra T. (2006) *Theoretical and Applied Genetics*, 112 (5), 808-817.
[24] Temnykh S., Declerck G., Lukashova A., Lipovich L., Cartinhour S., and McCouch S. (2001) Genome Research, 11(8), 1441-1452.
[25] Fulton T.M., van der Hoeven R., Eannetta N.T., and Tanksley S.D. (2002). *Plant Cell*, 14 (7), 1457-1467.
[26] Decroocq V., Favé M., Hagen L., Bordenave L. and Decroocq S. (2003) *Theoretical and Applied Genetics*, 106 (5), 912-922.
[27] Squirrell J., Hollingsworth P.M., Woodhead M., Russell J., Lowe A.J. and Gibby M. (2003) *Molecular Ecology*, 12(6), 1339–1348.
[28] Zhang L.Y., Bernard M., Leroy P., Feuillet C. and Sourdille P. (2005) *Theoretical and Applied Genetics*, 111 (4), 677-687.
[29] Huang X., and Madan A. (1999) *Genome Research*, 9 (9), 868-877.
[30] Doyle J. and Doyle J. (1987) *Rapid DNA isolation procedure for small quantities of fresh leaf tissue. Photochemistry Bulletin*, 19(1), 11-15.
[31] Cullings K. (1992) *Molecular Ecology*, 1 (4), 233-240.
[32] Cho Y., Ishii T., Temnykh S., Chen, X., Lipovich L., McCouch S., Park W., Ayres N. and Cartinhour S. (2000) *Theoretical and Applied Genetics*, 100(5), 713–722.
[33] Gutierrez M.V., Patto M.C., Huguet T., Cubero J.I., Moreno M.T. and Torres A.M. (2005). *Theoretical and Applied Genetics*, 110 (7), 1210–1217.
[34] Pashley C.H., Ellis J.R., McCauley D. E. and Burke J.M. (2006) *Journal of Heredity*, 97(4), 381–388.
[35] Liewlaksaneeyanawin C., Ritland C.E., El-Kassaby Y.A. and Ritland K. (2004) *Theoretical and Applied Genetics*, 109 (2), 361–369.
[36] Ellis J. R., Pashley C.H., Burke J.M. and McCauley D.E. (2006) *Molecular Ecology*, 15 (9), 2345–2355.
[37] Lehman T. (1996) *Genetics*, 144 (3), 1155–1163
[38] Callen D.F., Thompson A.D., Shen Y., Phillips H.A., Richards R.I., Mulley J.C. and Sutherland G.R. (1999) *American Journal of Human Genetics*, 52 (5), 922–927.
[39] Park Y., Lee I. and Kim N. (2009) *Molecules* 14(x) , 4546-4569, doi:10.3390/molecules1411454
[40] Chu H., Hu Y., Wang H.C., and Li J. (2009) *Molecular Ecology Resources*, 10(1), 150–155.
[41] Ellegren H. (2000) *Trends in Genetics*, 16 (2), 551–558.
[42] Diwan N., Bhagwat A.A., Bauchan G.B. and Cregan P.B. (1997) *Genome*, 40 (6), 887-895
[43] Diwan N., Bouton J.H., Kochert G. and Cregan (2000) *Theoretical and Applied Genetics*, 101(1-2), 165-172.
[44] Baquerizo-Audiot E., Desplanque B., Prosperi J.M., and Santoni S. (2001) *Molecular Ecology Notes*, 1(1-2), 1-3.
[45] Julier B., Flajoulot S., Barre P., Cardinet G., Santoni S., Huguet T. and Huyghe C. (2003) *BMC Plant Biology*, 3(9), 1471-1490.
[46] Eujayl I., Sledge M.K., Wang L., May, G.D., Chekhovskiy K., Zwonitzer J.C. and Mian M.A. (2004) *Theoretical and Applied Genetics*, 108(3), 414-422.
[47] Badri M., Zitoun A., Ilahi H., Huguet T. and Aouani M.E. (2008) *Journal of Genetics*, 87 (3), 241-255.
[48] Badri M., Zitoun A., Soula S., Houcine I., Huguet T. and Aouani M.E. (2008) *Conservation genetics,* 9(6), 1509-1520.
[49] Mun J., Kim D., Choi H., Gish J., Debelle F., Mudge J., Denny R., Endre G., Saurat O., Dudez A., Kiss G., Roe B., Young N. and Cook D. (2006) *Genetics*, 172 (4): 2541–2555.

.

Table 1- Cross-species Medicago truncatula anonymous SSR (microsatellites) markers used for PCR screening of Mesquite plant (Prosopis juliflora) genomic DNA

No	Marker Name	Repeat Motifs / folds	SSR Forward Primer (5'-3')	SSR Reverse Primer (5'-3')	Transferability (+/-)	No. of PCR Bands	Band Size (bp)
AT repeat motif							
1	mtgsp_004G08	AT-11	GTGGAAGTTTGATCATTTCG	CAAATCTCACCCACTACA	-	-	-
2	mtgsp_003F07	AT-11	CAGGAGGACAAACACAACTC	TCCACGCTTTGAGATAAAATGA	+	1	~190
3	mtgsp_003C03	AT-16	CCGGACACCCCACTTATTTA	CCTCACAGTGCATTGGAAAA	-	-	-
4	mtgsp_001G08	AT-12	CGGACAAAACAGATTGTCCTT	GAAGGTGCGTTTAGCAACA	+	1	~180
5	mtgsp_001C04	AT-19	TGGCAAAGTGATGAGAGGGT	ATATACCACCACAGCCGGAG	-	-	-
TA repeat motif							
6	mtgsp_001B05	TA-32	TTGTCTTAAATCGGATGGCT	AGCGTTACAGGGTGTTCCTG	-	-	-
7	mtgsp_005B08	TA-10	GCGTTAGCATGGGTTAATGG	GCAAACAATGGTGTGTCGAG	-	-	-
8	mtgsp_001E07	TA-8	GCCCTAAGGACTGCATTTTG	CCCCTCCTAAACCCTCAATC	-	-	-
9	mtgsp_004B09	TA-16	TATCAGTCATAAGGCAAAAG	CTATTTAAGAGTATAATTGGGA	-	-	-
10	mtgsp_005h11	TA-24	GTTGCCCGTTTGTCTGCTAT	AAGGTGACCAAGGAATGTGG	-	-	-
AG/CT repeat motif							
11	mtgsp_004d03	AG-14	GAGAGATCACGACGGGACAT	GCTTGTTGGGTTGGTCTGTT	-	-	-
12	mtgsp_001b05	AG-17	TGGGTTTACGGAAGGATGAA	GCAAGCACTTCCCATAAAGC	-	-	-
13	mtgsp_001c10	AG-18	CAAATTTGCCCCACACAAAT	TTGCCTTTCTGCCACTCTTT	+	1	170
14	mtgsp_005g08	AG-21	GCTTTCCGTAGGTTTGGTTG	TTCTTTCCCTCCTCCTGGTT	-	-	-
GA repeat motif							
15	mtgsp_001f03	GA-20	TGCAACAACAATGACCCACT	GTTTGCGGTGGAAACAAAGT	-	-	-
16	mtgsp_002F04	GA-19	AATCGCGTTGAGGTAACTGG	CTGCTTCTCTTCCCATTTCG	-	-	-
17	mtgsp_001e09	GA-16	CGGTGATTTTGAGCAAGAAA	ATGCCCTCCGGATTACATTA	-	-	-
18	mtgsp_002f04		AATCGCGTTGAGGTAACTGG	GTCCTTTCTCTGCTGCTGCT	-	-	-
19	mtgsp_003h04	GA- 28	ATGGTTTATATCCGCTGCGA	GGGGCATTTTCAGAATTTCA	-	-	-
AC, CA, TG repeat motifs							
20	mtgsp_005D09	AC-20	TGTTTCGATGGATTTGATTTTG	GCTTGCCTTCAACCTTCAAC	-	-	-
21	mtgsp_002c04	AC-29	GATAGCAGCTCAGGGCAGTT	GAGGCAACCTCACAAGGAAA	-	-	-
22	mtgsp_002c05	CA-16	CATGGTGCATTCAGAGCAAT	GCGGGTGTCTTCTAGCAAAT	-	-	-
23	mtgsp_003h02	CA-14	CCTGGTGTCAACGGATCCTA	AATCTGTGGTGCAAAAAGCA	-	-	-
24	mtgsp_004a06	TG-17	TCAATGGATGAGATTTGGTGTC	TGTGGTAAAATGTGGTCCCTAA	-	-	-
25	mtgsp_002h12	TG-14	TCCAACCCCTGAACCTTACA	CCTGAACCTTTGGTGTGGTT	-	-	-
TTC, TCT, GTT, GGT repeat motifs							
26	mtgsp_001g03	TTC-.30	TCTTCACCTCCAAGGTCCAC	CCCCACACTTAGGAGAAGCA	-	-	-
27	mtgsp_003a08b	TTC-15	TCCAACCCAATAGCCCATTA	GGGAACGAGCAACCAAATAA	-	-	-
28	mtgsp_001G05	TCT-12	AAGAATGACGAAGAGGCGAA	TCAGAAATTCCCTCCCATTG	-	-	-
29	mtgsp_005A11	GTT-6	AGAACGCTATGGCGTCATTT	GGGTTTCCCATTGTCGTATG	-	-	-
30	mtgsp_001e10	GGT-7	CATGTTGGTCTCGGTTTGTG	GCCGCTAACAAACATCCAGT	-	-	-
AGC, CAC, AAG, AGA, GAA repeat motifs							
31	mtgsp_004g06	AGC-8	TTGGCAGCATTGTCATCATT	GCTCCAAGCAATGCTACTCC	-	-	-
32	mtgsp_004f08	CAC-7	CATGGATTCTGTTGCAGGTG	GGACCAGAGGCTACAGCTTG	-	-	-
33	mtgsp_005f09	AAG-12	GCGTTTTCCATGGTTGAGTT	CTCCCTCCAATTCGTTTCAA	-	-	-

34	mtgsp_004a05	AGA-9	AGGGTTCAAGCTTGGTTTGAT	AAAAGCCCGTGGTTTCTTGT	-	-	-
35	mtgsp_001d07	GAA-29	CCATCTTCACCTCCAAGGT	CCCCACACTTAGGAGAAGCA	-	-	-
36	mtgsp_005e05	GAA-8	AGTGAGAAGGAGCGCATGAT	GGGTTTTCAGATTTGCTTGTG	-	-	-
TAT, ATA, TAA, AAT, ATT, TTA repeat motifs							
37	mtgsp_004d06	TAT-.22	AAAGGTTTCTTTTGACAAATCCA	ACCAAACAGGCAGTTCTTGG	-	-	-
38	mtgsp_001c12	ATA-17	TCTTAATACCCGTGGGAGCA	AGCATGTACGGTTTTGGTGG	-	-	-
39	mtgsp_005b09	ATA-13	TCAAAGCCACAAGCAATCAG	GTGGCTCACTTTGTCCAGGT	-	-	-
40	mtgsp_005e04	TAA-9	ACTCGCATCACATTGTGGAA	GACTTTGTCGGGTGGTCCTA	-	-	-
41	mtgsp_004f01	AAT-7	CTCCTTTCCACGTGTCGTTT	GGTCGCCGACAGAGATTAAG	-	-	-
42	mtgsp_002f12	AAT-11	TGAGAGAGCTACGGGAGGAA	CGATGTAGCAAAATTCGAGGA	-	-	-
43	mtgsp_003f07	ATT-13	AGGGGAAGATTTTTGGGAAG	GACTCAGGTCGCAAGAAAA	-	-	-
44	mtgsp_001A05	ATT-24	GCACGACTGCCAAGTCTTCT	CGAGGAAGTATTTCATGCCA	-	-	-
45	mtgsp_001C05	TTA-8	TACTGGGTTCACGCACAAAA	TTCAACCGTACCGCTCTTCT	-	-	-
46	mtgsp_005a07	TTA-16	TCCGCTAAAACCAGGATACG	TCTCAGGCTACTGGGATTTTG	-	-	-

Microsatellites markers were tested previously in different taxa within *Medicago* model plant species and shown high level of transferability as it is reported by: **[**42, 43, 44; 45, 46, 47, 48**]**. SSR markers having the numbers (mtgsp_003F07, mtgsp_001G08 and mtgsp_001c10) have shown positive PCR amplification in Mesquite genomic DNA tested in this paper, which contain AT and AG repeat motifs. Abbreviations: SSR, simple-sequence repeat; +, Present of PCR product; - Absent of PCR product; na, not available. Observed PCR product size is base number (bp).

Table 2- Frequency rates of various microsatellite categories (motifs) per million base pairs in genomic and EST sequences of some legume plant genera

Repeat Motif	**Frequency rates in EST Sequences**			
	Prosopis juliflora	***Medicago truncatula***	***Glycine max***	***Lotus japonicus***
A/T	36	-	-	-
C/G	-	0.7	0.8	0
AT/AT /TA/TA	8	2.4	7	1.7
AG/CT/ GA/TC	16	13.6	12.5	18.5
AC/GT/ CA/TG	12	0.6	0.6	0.1
AAT/ATT/ TAA/TTA	4	1.8	1.7	0
AAG/AGA/GAA/CTT/TTC/TCT	2	6.4	2.9	14
AGG/CCT/ TCC/GGA /GAG/CTC	6	0.3	0.4	5.1
AAC/ACA/CAA/GTT/TTG/TGT	0	0.9	2.3	3.9
ACG/CGA/GAC/CGT/GTC/TCG	0	0	0.3	0.4
Reference	The Current Study	**[**49**]**	**[**49**]**	**[**49**]**

ESTs sequences of the mentioned legumes; *Prosopis juliflora, Medicago truncatula*, *Glycine max* and *L. japonicus*, can be described generally as: 1) (A/T) Mononucleotide repeats for EST sequences are the more abundant, 2) (AG) dinucleotide repeats are more abundant, 3) (AT) repeats frequency rate in *Prosopis juliflora* is closely similar to those rates of the three other legumes.

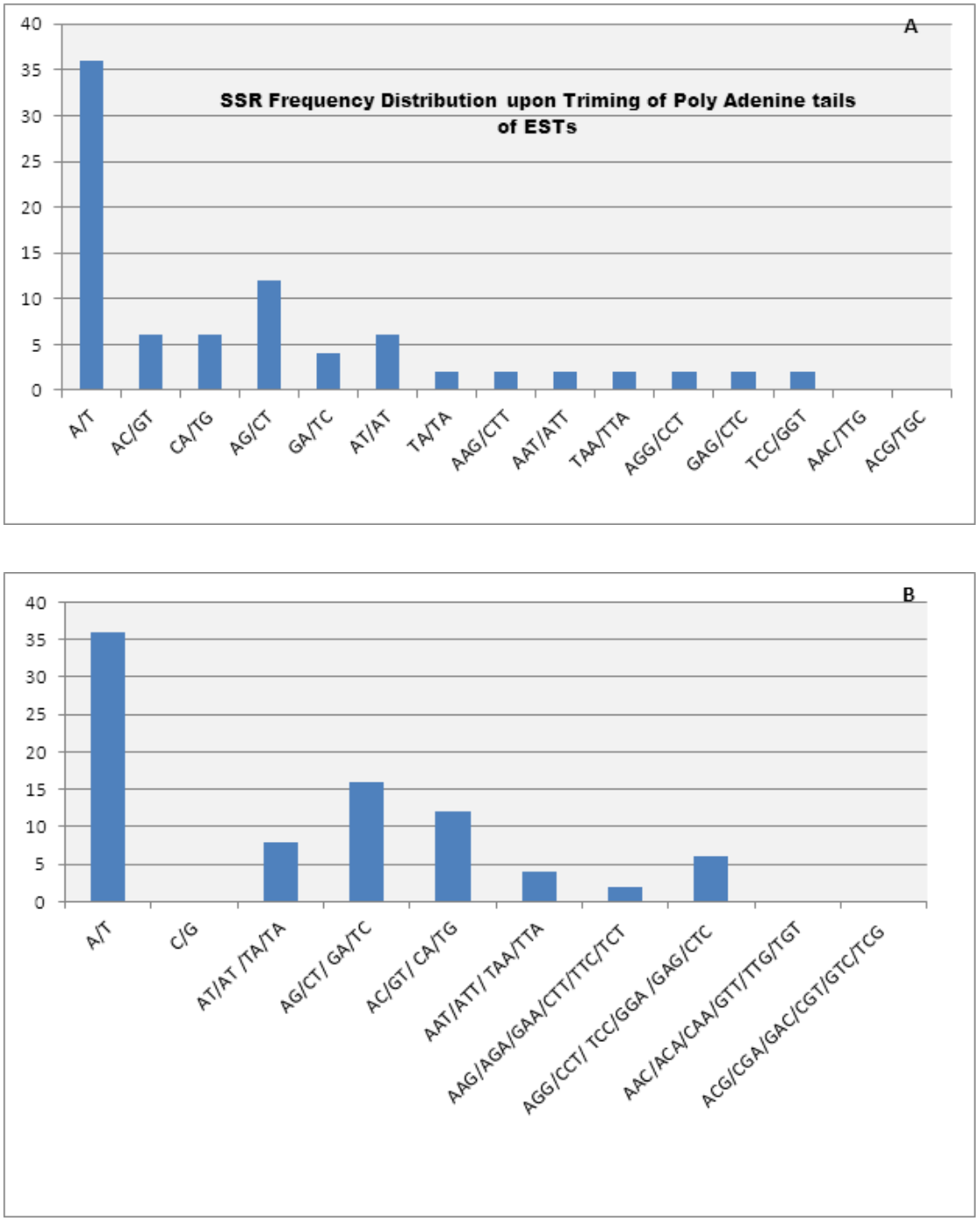

Fig. 1-AB-*Prosopis juliflora* SSR motif repeats distribution upon trimming of Poly Adenine tails from the ESTs sequences. In A; Frequency rates of SSR repeat motifs are categorized as, high; (A/T), intermediate; (AG/CT), (AT/AT), (CA/TG), (AC/GT) low; (GA/TC), (TA/TA), (AAG/CTT), (AAT/ATT), (TAA/TTA), (AGG/CCT), (GAG/CTC), TCC/GGT) and absent: (AAC/TTG), (ACG/TGC), while B; SSR motifs are organized according to motif repeats shared with other legumes SSR categories .

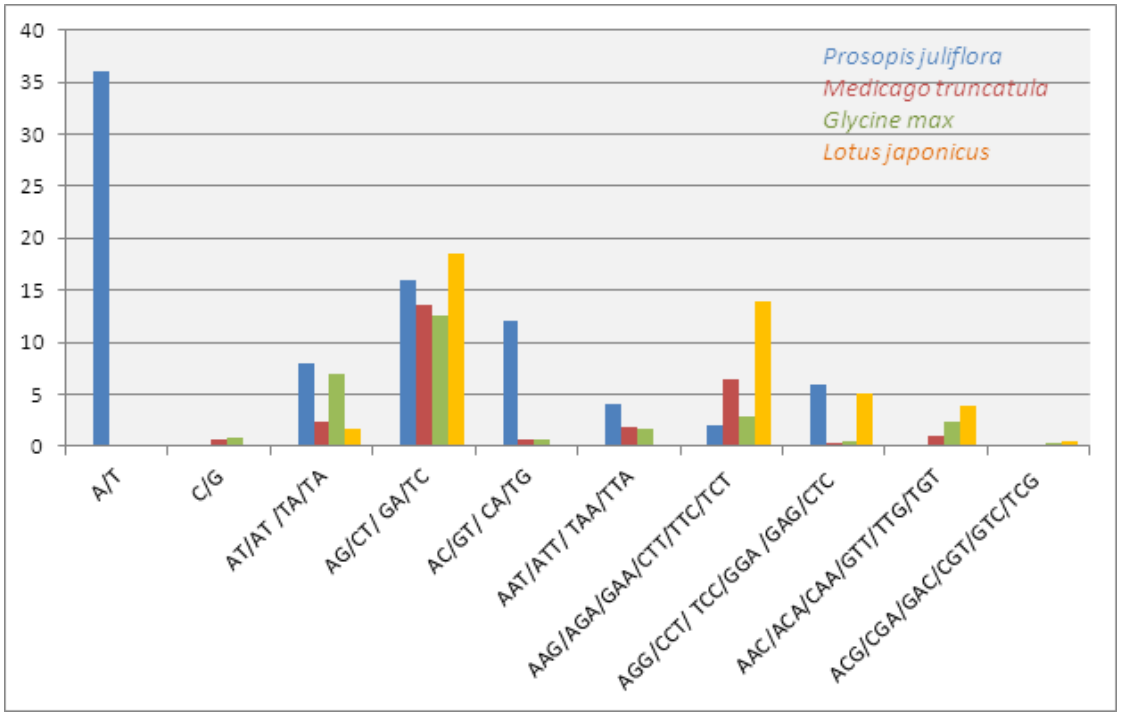

Fig. 2-SSR repeat motifs distribution of *Prosopis juliflora* in comparison with corresponding SSR repeat motifs of the legumes; *Medicago truncatula*, *Glycine max* and *Lotus japonicus.* Frequency rates that are closely related among the four legumes are; (AG/CT/GA/TC), (AT/AT/TA/TA).

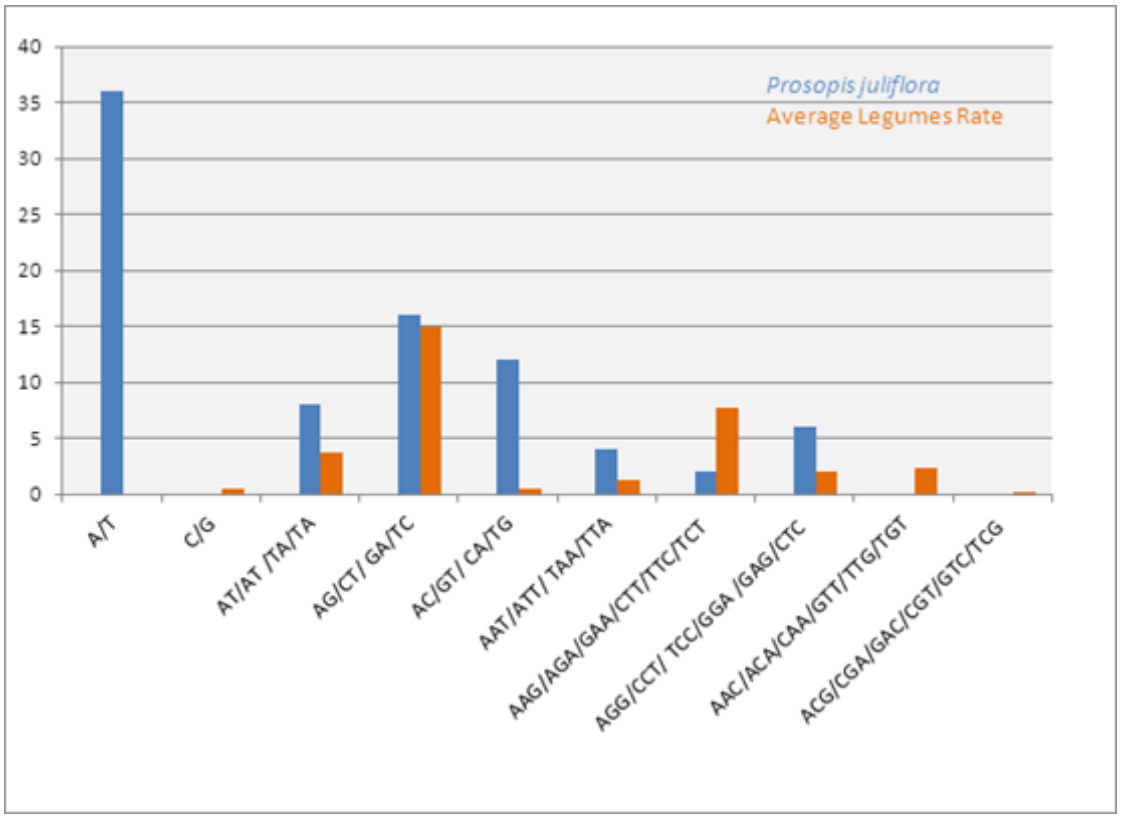

Fig. 3-SSR repeat motifs distribution of *Prosopis juliflora* in comparison with average SSR motifs rates of the legumes; *Medicago truncatula*, *Glycine max* and *Lotus japonicus.* Frequency rates of *Prosopis juliflora* that are nearly similar to the average rate of other legumes are; (AG/CT/GA/TC), (AT/AT/TA/TA)

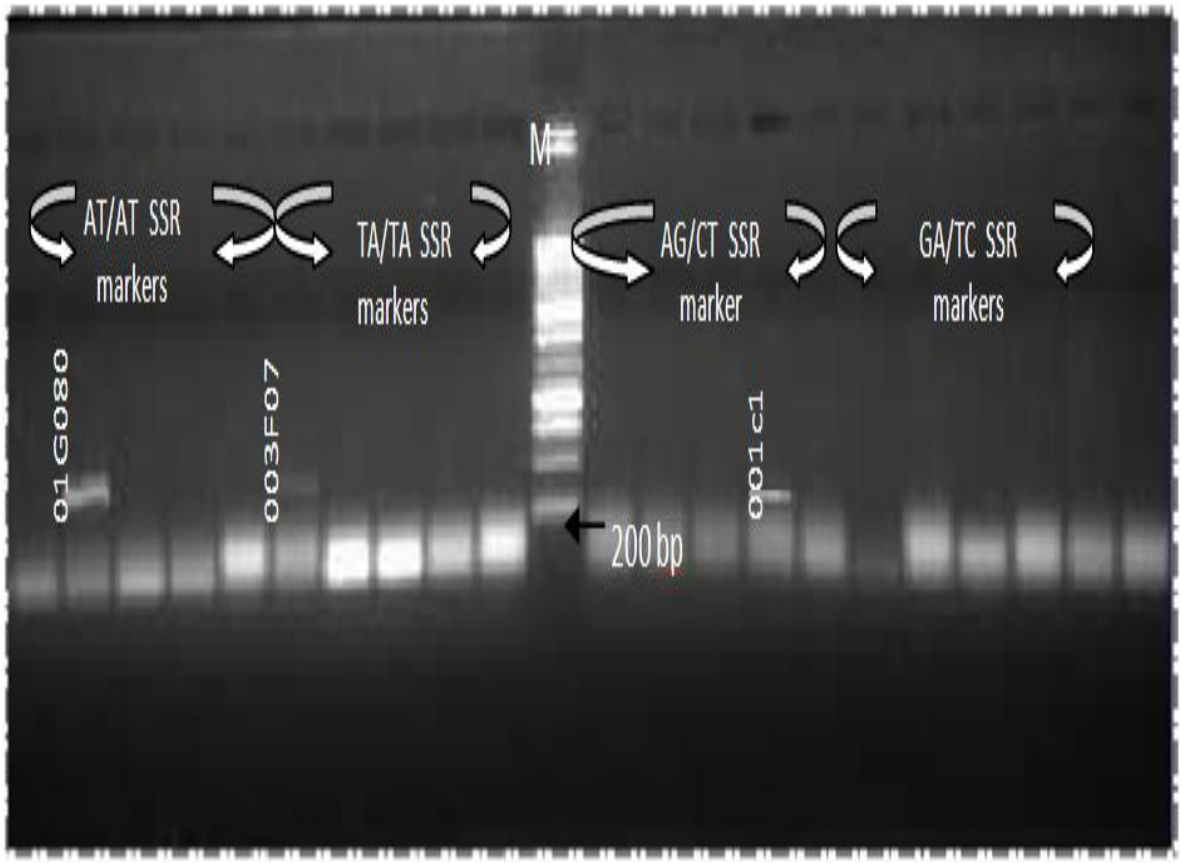

Fig. 4-Schematic appearance of PCR amplification of some SSR markers having AT and AG repeat motifs sequences. The gel image also shows the negative amplification of some markers having medium, low and rare frequency rate of SSR repeat motifs in *Prosopis juliflora* Abbreviation: M, 1kilo base pair DNA ladder sequences.