COMPARATIVE ANALYSIS OF THE COMPLETE CHLOROPLAST GENOME OF THREE PROSOPIS SPECIES IN JORDAN

Hamad Adel Alkhatatbeh* Khaleel I Jawasreh ** Maher J. Tadros***
Researcher Prof. Prof.
*** Dept. Plant Prod., Faculty of Agric., Jordan Univ. of Sci. and Techn.**
**** Dept. of Animal Prod., Faculty of Agric., Jordan Univ. of Sci. and Techn.**
***** Dept. of Natural Res. and Environ and Dept. of Plant Prod./ Faculty of Agric, Jordan Univ. of Sci. and Techn.**
E-mail: mtadros@just.edu.jo, Corresponding Author.

ABSTRACT

Worldwide, *Prosopis* Genus is widely spread and well known to be of high tolerance to harsh conditions. The *P. juliflora* and *P. cineraria* species were introduced to the Mediterranean, while *P. farcta* is a native one. The genomic structure of the chloroplast of *P. juliflora, P. cineraria* and *P. farcta* were targeted in this study. The chloroplast DNA samples were sequenced by genetic analyzer sequencer "Ion S5™ System. The results indicated the size of the genome to be ranged between 162900 bp in *P. farcta* and 163667 bp in *P. cineraria*. The full chloroplast of *P. juliflora* and *P. cineraria* genome were reported for the first time nationally in Jordan, while globally *P. farcta* was the first to be analyzed genetically. The present study offers an important portfolio of *Prosopis* species chloroplast genome analyses, this could help with identification of species and speed up biological and genetic diversity researches.

Key words: DNA sequence, Tolerance, Ion S5, variable sec.

مقارنة التحليل الكامل للمادة الوراثية في البلاستيدات الخضراء لثلاثة أنواع من السلم (*Prosopis juliflora, Prosopis cineraria and Prosopis farcta*) في الأردن

حمد عادل الخطاطبه خليل جواسره ماهر تادرس
باحث أستاذ أستاذ
قسم الإنتاج النباتي، قسم الإنتاج الحيواني وقسم الموارد الطبيعية والبيئة
كلية الزراعة – جامعة العلوم والتكنولوجيا الأردنية

المستخلص

ينتشر جنس السلم على نطاق واسع في العديد من البلدان ومن المعروف جيدًا أنه يتمتع بدرجة عالية من التحمل والتكيف مع الظروف القاسية. أظهر وجود مثل هذه الأنواع في المناطق المحلية والإقليمية سلوكًا غازيًا كأنواع مدخلة تؤثر على التنوع البيولوجي للنظام الإيكولوجي للأنواع المحلية. تم إدخال النوعين *P. juliflora* و*P. cineraria* للأردن، بينما يعتبر *P. farcta* موطنه الأصلي في منطقة البحر الأبيض المتوسط. أجريت هذه الدراسة لمعرفة التركيب الجيني للبلاستيدات الخضراء *(P. juliflora, P. cineraria, P. farcta)* . جينومات البلاستيدات الخضراء (cpDNA) دائرية ومحفوظة نسبيًا بين نباتات الأرض من حيث الحجم والبنية والمحتوى الجيني. يمكن أن يوفر الجينوم الكامل لنباتات البروسوبيس دقة أفضل حول نشأة وتطور الأصناف في المناطق الجغرافية. لذلك، تم تحليل تسلسل جينوم البلاستيدات الخضراء للأنواع الثلاثة باستعمال آلة التسلسل IonS5. أشارت النتائج إلى أن حجم الجينوم يتراوح بين 162900 زوج قاعدي في *P. farcta* و163667زوج قاعدي في *P. cineraria*. أظهر التحليل الوراثي وجود تشابه كبير بين *P. juliflora* و *P. cineraria* مقارنة مع *P. farcta*. أشارت نتائجنا لأول مرة إلى الجينوم الكامل للبلاستيدات الخضراء لـ *P. juliflora* و *P. cineraria* في الأردن، بينما عالميًا، كان *P. farcta* أول مرة يتم تحليله وراثيا. سيفتح الاختلاف الجيني الجزيئي بين وداخل أنواع *Prosopis* الباب للباحثين لإجراء دراسات ارتباط الجينوم على نطاق واسع لتحسين بعض السمات المحددة المفيدة لأخصائي المراعي والثروة الحيوانية أيضًا.

الكلمات المفتاحية: تسلسل DNA، التحمل، الأيون أس 5، المتغيرات.

Received:11/5/2022, Accepted:14/8/2022

INTRODUCTION

Prosopis (Mesquite) is considered one of the worldwide flowering plants belonging to the Fabaceae family which is found in arid or semi-arid climatic zones. This genus includes 44 species distributed in the Near East, Mexico, North Africa, Argentina, America and the Northeast of Brazil, and the Caribbean (40). *Prosopis* genus varies widely in their productivity and uses (40). Due to their multi-purpose nature, and their ability to produce many products, these type of trees and shrubs have a very important role among natural resources in semiarid and arid regions, especially desert areas such as animal feed, wood fuel and nitrogen fixation (40,17), furthermore, these shrubs and trees have the ability to survive efficiently in the poorest dry soils (1). Biodiversity is an important factor for the ecosystem sustainability and forests are among the important components of this system, which is a pillar for the maintenance of vegetation as the loss of these ecosystems poses a threat to the global ecosystem, and thus affects humans, plants, animals, and other living organisms. Therefore, there is national, regional and a global interest for preserving the spread of trees worldwide, especially those adapted to climate change, irregular rainfall, changing temperatures, salinity, and drought. Among such trees species is *Prosopis*, since this type of trees are distributed in two forms in the world. *Prosopis* might be native or invasive presented in many countries of the world, including Asia, Africa and America (9). For the continuity of the ecosystem and for preserving the native or invasive trees found in the ecosystem, it is necessary to understand the genetic makeup of such species (genetic variation among species) as genetic resources contribute substantially to the adaptation process of the trees during fluctuations in the climate conditions. The utilization of genetic variation between and within species is one of the important alternatives used for conserving different species found in any ecosystem. This could be adding a lot to the traditional approach that adopted previously by plant breeders through identifying the most tolerant species concluded by phenotypic variation. Avoiding the long generation intervals calculated for such species. The genetic information used early for selection purposes to produce a high tolerant species to harsh conditions such as drought, salinity, and others. For this reason, when trees are exposed to environmental stresses, the ability of trees to adapt depends on the genetic makeup of any studied trees. Therefore, the genetic map of each species must be studied separately in order to know the genetic variances as the phenotypic manifestation is basically based on genetic variance (36). In Jordan, although Tadros et al., (49) reported three *Prosopis* species grown in the Jordan valley and the Dead Sea areas in terms of morphological characteristics, It was confirmed that genetic analysis using the entire genomic information and protein's coded genes can provide more accurate information about the origin and evolution of varieties in geographic regions, as well as how gene flows among common genotypes (22). Understanding the genetic origins of the species and undertaking analysis and comparison with other species genomes, enabled the strategic development for preserving the species which, contribute a lot in understanding the evolutionary relationships among the genomes and how they adapt. Although of the rarely and expensive cost of the genetic studies (25), it was necessary to proceed with the detection of genes in *Prosopis* species to investigate adaptations at the molecular and morphological level. Chloroplast genomes (CPG) have evolved into a worthy new resource for identification of many species, genetic engineering, population genetics, and plant phylogenetics, chloroplast's organelles perform photosynthesis, fatty acids, pigments, and the biosynthesis of amino such as Alanine, Cysteine, Aspartic Acid, Glutamic Acid, Phenylalanine and Glycine (39,43). The CPG is generally conserved in many families, including angiosperms, such as Fabaceae, Geraniase, Campaniolacia, and Oleaceae (18). The investigation of the chloroplast genome will open the area for the researchers to study the detected genes, mutations and other molecular characteristics and their possible association with economical traits, the ability of including them in selection process of those three species (*P. juliflora, Prosopis cineraria* and *P. farcta*). The objectives of this study

were to detect the sequence of the complete CPG sequence of three *Prosopis* species grown in Jordan (*P. juliflora, Prosopis cineraria* and *P. farcta*) and to compare with previously sequenced data related to a similar family, the three chloroplast *Prosopis* species genomes will provide the main structure of the genomic sequence including all information related to the genes, exons, introns, and the possible function of each gene.

MATERIAIS AND METHODS

Sample collection

Forty plant samples of *P. juliflora*, *P. farcta* and *Prosopis cineraria* fresh leaves were collected from different regions in Jordan; the boundary area of Al-Omari (Al-Hazeem) (31°53′37.62″ N,37°07′51.60E), AlAghwar (32°20′47.75″ N,35°62′15.94E) and the Dead Sea (Sweimah) (31°77′40.11″ N, 35°59′45.94″ E). The coordinated were identified by using Global Position Systems (GPS).

DNA extraction

The DNA extraction of chloroplasts, nucleic acid quality control, quantification, and gel electrophoresis were conducted to prepare the library for quantification using a bio-analyzer (37). DNA was extracted from plant tissues by using a modified protocol (45) used at the Biotechnology Omics Lab and the University of Nizwa, Oman. The genetic comparison was conducted among the three types, *Prosopis juliflora, Prosopis farcta*, and *Prosopis cineraria,* based on the chloroplasts analysis due to their importance in plants being responsible for photosynthesis, food synthesis and many chemical processes in plants (31).

Chloroplast genome sequencing

To produce genomic libraries, instructions from manufacturer (Life Technologies USA) were followed. Ion Shear™ Plus Reagents package was utilized to enzymatically portion cpDNA into fractions of 400bp and library kit for Ion Xpress™ Plus gDNA fragment. The quantification and qualification libraries are based on Qubit 3.0 fluorometer and Bioanalyzers (Agilent 2100 Bioanalyzer Systems, Life Technologies USA). After the library preparation, Ion One Touch™ 2, was used to amplify the template, enhanced by Ion 530 and 520 OT2 Reagents. The Ion One Touch™ ES enrichment method.

Genome assembly using chloroplast references

For P. juliflora, P. farcta, and Prosopis cineraria, raw reads totaled 540522846, 67657939, and 288522703, respectively. The produced cp genome transcripts were aligned to previously available cp genomes of P. juliflora (MN104889.1), Prosopis cineraria (MN104890.1), and Prosopis glandulosa (KJ468101), respectively, using Bowtie2 (v.2.2.3) (30) as reference genomes in the software Geneious Pro (v.10.2.3) (24). The mean exposure for the P. juliflora and P. farcta assemblies was 213X, 112X, and 175X, respectively. The inverted was identified using the obtain repeats plugin in Geneious Pro (v.10.2.3).

Prosopis species genome annotation

The software of tRNA scan-SE 1.21 (38) was utilized in tRNA gene detections, and the annotation of *P. farcta*, *P. juliflora*, *and P. cineraria* (CPG) Dual Organellar Genome Annotator (DOGMA) (54) was used to evaluate tRNA, ribosomal RNA, and protein gene coding. Additionally, tRNA scan-SE (38) and Geneious Pro (v.10.2.3). Kearse et al., 2012 (24) were used to manually change the genomes by adjusting the introns limit codons (stop and start) in addition to their comparison that made to *P. glandulosa* genome. OGDRAW (32) was also used to demonstrate the structural characteristics of the Prosopis species cp genome. Furthermore, mVISTA (19) in Shuffle-LAGAN mode was used to examine cp genome divergence among these organisms, with *P. juliflora* serving as the reference genome.

Repeat identification

REPuter program (28) was used to identify forward and reverse repeats. A minimal requirement was a 15-bp sequence with 90 percent identity. In addition, SSRs were determined using MISA software (8) with the following look for parameters: ≥ ten units for repeated a single bp, ≥ eight units of repetition for 2 bp repetitions, ≥ four units of repetition for bp repeats of three and four, and ≥ three units of repetition for bp repeats of five and six. By utilizing the default parameters, Tandem Repeats Finder version 4.07 (27) searches for tandem duplicates.

Divergence of chloroplast genomes and their phylogenetic relationship

The complete genome sequence variation and shared genes between related species of Prosopis were specified. To classify the unclear and incomplete gene annotation, after comparing genome sequence and multiple sequence alignment, a comparative analysis method was applied. The entire cp genomes were aligned using MAFFT version 7.222 (23) with default parameters, and pairwise sequence divergence was assessed using Kimura's two-parameter (K2P) model Küster & Williams, 1964 (29). 23 CPG sequences were collected from the NCBI database a phylogenetic location's inference of *P. f, P. c* and *P. j* within the sub family of Caesalpinioideae (Leguminosae). Alignments of the entire chloroplast genomes were created using conserved gene ordering and the cp genomes' layouts (52), and three different methods—Bayesian inference (BI), implemented in MrBayes 3.1.2 (44); maximum parsimony (MP), using PAUP 4.0 (48); and maximum likelihood (ML), utilizing MEGA six (27), by using previously described settings—were used to infer a phylogenetic tree (6,7). Model Test version v2.1.02 (41) was used to evaluate the best replacement model GTRG using the Bayesian posterior probabilities (PP) in BI tests using the Akaike information criterion (AIC). Starting with random trees and sampling one out of every 100 generations, the Markov Chain Monte Carlo (MCMC) process was used to simulate 1,000,000 generations using four incrementally heated chains. The values of first 30% of trees were discarded as burn-in. To approximate the posterior probabilities, the maximum parsimony run used an experimental search of thousand additions randomly of the replicated found in the sequence the Tree-bisection-reconnection (TBR) branch-swapping tree detection criteria. Similar to this, The Kimura 2-parameter model with invariant sites and gamma-distributed rate heterogeneity were used to generate the ML analysis thereafter the starting tree of 1000 start booting repeats was obtained by BIOJ tree software (20).

RESUITS AND DISCUSSION

Chloroplast genome structure and general features of prosopis

In this study forty samples of fresh leaves were collected from each of *Prosopis juliflora (P. j), Prosopis farcta (P. f)* and *Prosopis cineraria (P. c)*. The chloroplast DNA samples were sequenced by genetic analyzer sequencer "Ion S5™ System". The resulted sequences summary that showed the features of the genome of the studied three *Prosopis* species are illustrated in (Table 1). The three sequenced *Prosopis* species in the CPG of the (*Prosopis juliflora (P. j)* (MZ073640), *Prosopis farcta (P. f)* (MZ073639) and *Prosopis cineraria (P. c)* (MZ073638)) were observed to be circular molecules liked typical in angiosperm CPG forming a conformation resemble too quadripartite. The CPG sizes of *P. juliflora, P. cineraria* and *P. farcta* are 163236, 163667, and 162900 base pairs, respectively (Figure 1, Table 1). The analyzed and compared of *P. juliflora, P. cineraria* and *P. farcta* CPG with five associated CPG, ranging in size from 161240 bp (*Dichrostachys cinerea*) (NC_035346.1) to 164692 bp (*Leucaena trichandra*) (KT428297.1) (Table 1). Among *Prosopis* species, *P. cineraria* (163667bp) and *P. juliflora* (163236 bp) (CPG) were larger than *P. farcta* (162900 bp). The *P. j, P. c* and *P. f* CPG typically round in shape and consist of four parts: (a)Large single copy (LSC) region of 92494, 92940bp and 92156 bp, covering 56.7%, 56.8% and 56.6% in the genomes, correspondingly; (b) Short single copy (SSC) region of 18880, 18865 and 18882 bp, covering 11.6%, 11.5% and 11.6% of the genome, and (c) two Inverted repeats (IR) regions divide the LSC and SSC regions, that were 25931, 25931 base pairs and 25932 base pairs in size, that was of 15.8 % of the totality of the genome. The *P. j, P. c* and *P. f* CPG encode 126, 126 and 127 genes including the same number of coding for proteins 82, 36 tRNA (*P. c; P. j*) and 37 tRNA genes (*P. f*), and 8 rRNA genes, respectively (Figure 1 and Table 1). The CPG of the three studied *prosopis* species contain about 35.9% of GC that comprised 35.6% in *Leucaena trichandra*. The majority of the genes were detected in the three *Prosopis* species being 22 genes

distributed in the different genomic regions as shown in Table 2. In this the detected gene's structure including the exon and intronic regions and other related regions are illustrated. Three genes (*trn*G-GCC, *trn*I-GAU and *trn*V-UAC) out of the 22 were just only found in the *P. farcta* and one gene *acc*D was only detected in *P. juliflora*. The *ycf*1 gene was found only in *P. cineraria* and characterized by found only in both regions SSC-IR (Table 2). The number of exons and introns of the detected genes were similar in all studied *Prosopis* species being of two exons and a single intron, furthermore three exons and two introns were detected in Protase and *ycf*3 genes. Few differences or variation were observed in the size of exons and introns of the three *Prosopis* species.

Genes found in the prosopis genomes: Among the annotated genes, the structure of nineteen genes (table 2), that detected in the *Prosopis* CPG were only of one intron, the full number of exons and introns are shown in Table 2. The *ycf1* gene was trans-spliced, with the 5′ end exon found in the LSC region and the 3′ exon found in the IR region. Fourteen ribosomal protein encoding genes (*rps*2 to rps19), ten genes encoding large ribosomal proteins (*rpl*2 to *rpl*36), fourteen genes were found to be related to photosystem tow (*psb*A, to *psb*Z, and *ycf*3), five genes encoding photosystem I components (*psa*A, to *psa*J), and six genes encoding the synthesis of the ATP and the component of the transport electron chain (*atp*A to, *atp*I) (Table 3). Table (3) shows the genes name that divided into groups, each group of this genes are responsible for specific function or category, group of genes included large / small subunit of ribosomal proteins, DNA dependent RNA polymerase, rRNA genes and tRNA genes that stand for self-replication function. Also, Photosynthesis category that resulted from group of genes (Photosystem I, Photosystem II, Cytochrome b6/f complex, ATP synthase, Rubisco, and Subunits of NADH-dehydrogenase). In addition to, some group of genes that controlled other functions such as Maturase, Protease, Envelop membrane protein, Subunit of Acetyl-CoA-carboxylase and c-type cytochrome synthesis gene. Other genes called conserved open reading frames until now considered unknown function (Table 3).

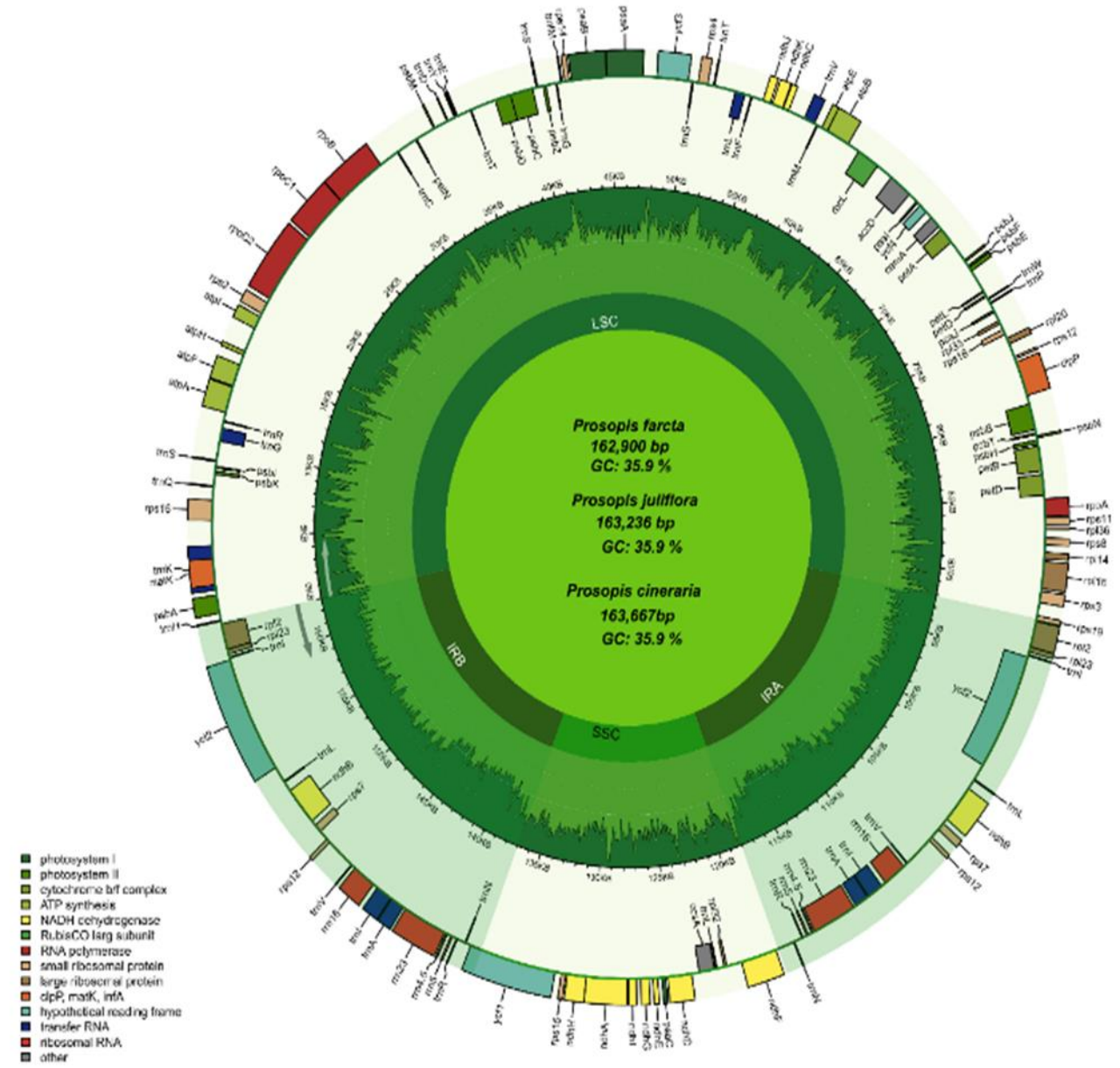


Figure 1. *P. juliflora, P. farcta, and P. cineraria* genome maps (CPG). Thick lines represent the extent of the inverted repeat regions (IRs), which divide the CPG into large (LSC) and small (SSC) single copy regions. Genes drawn within the circle are transcribed clockwise, whereas those drawn outside the circle are transcribed counterclockwise. Different functional groups of genes are color coded. The inner circle's dark green represents GC content, while the outer circle's light green represents AT content

Table 1. Summary of Prosopis complete chloroplast genomes features compared to other published related species.

	P. j (New)	*P. c (New)*	*P. f*	*D. c*	*L.t*	*P. j (old)*	*P. g (old)*	*P. c (old)*
Size (bp)	**163236**	**163667**	**162900**	**161240**	**164692**	**163237**	**163040**	**163677**
Overall GC contents (%)	35.9	35.9	35.9	**9.35**	**6.35**	**9.35**	**9.35**	**9.35**
LSC size in bp	**92494**	**92940**	**92156**	**90426**	**93690**	**92495**	**92310**	**92937**
SSC size in bp	**18880**	**18865**	**18882**	**18526**	**18890**	**18880**	**19132**	**18878**
IR size in bp	**25931**	**2531**	**25932**	**26144**	**26056**	**25931**	**25931**	**25931**
Protein coding regions (bp)	**75774**	**75633**	**77880**	**77958**	**78759**	**78421**	**78039**	**78883**
tRNA size (bp)	**2725**	**2725**	**2810**	**2793**	**2815**	**2927**	**2810**	**2868**
rRNA size (bp)	**9052**	**9052**	**9052**	**9068**	**90949**	**9052**	**9052**	**9052**
Number of genes	**126**	**126**	**127**	**128**	**129**	**132**	**128**	**131**
Number of protein coding genes	**82**	**82**	**82**	**83**	**84**	**85**	**83**	**85**
Number of rRNA	**8**	**8**	**8**	**8**	**8**	**8**	**8**	**8**
Number of tRNA	**36**	**36**	**37**	**37**	**37**	**39**	**37**	**38**
Genes with introns	**21**	**23**	**21**	**23**	**22**	**21**	**23**	**21**

P. j= *Prosopis juliflora*, P. c= *Prosopis cineraria*, P. f= *Prosopis farcta* P. g= *Prosopis glandulosa*, D. c=*Dichrostachys cinerea*, L. t= *Leucaena trichandra*. GC Contents: Guanine-Cytosine content, LSC: Large single copy, SSC: Short single copy, IR: Inverted repeat, tRNA: Transfer RNA, rRNA: Ribosomal RNA *(33)(34)(35).

Table 2. Genes introns and exons length found in the three Prosopis species CPG

Gene	Location	Gene size(bp)			Exon I (bp)			Intron I (bp)			Exon II (bp)			Intron II (bp)			Exon III (bp)		
		P.j	*P.c*	*P.f*	*P.j*	*P.c*	*P.f*	*P.j*	*P.c*	*P.f*	*P.j*	*P.c*	*P.f*	*P.j*	*P.c*	*P.f*	*P.j*	*P.c*	*P.f*
*acc*D	LSC	1542			668			33			841								
*ycf*1*	SSC-IR		3315			2009			123			1183							
*trn*G-GCC	LSC			771			23			699			49						
*trn*I-GAU [a]	IR			1028			42			951			35						
*trn*V-UAC	LSC			690			39			614			37						
*atp*F	LSC	1279	1279	1280	145	145	145	727	727	728	407	407	407						
*clp*P	LSC	2019	2019	2036	71	71	71	786	786	797	294	294	291	642	642	651	226	226	226
*Ndh*A	SSC	2541	2536	2549	553	553	553	1449	1444	1457	539	539	539						
*ndh*B [a]	IR	2218	2218	2218	775	775	775	685	685	685	758	758	758						
*Pet*B	LSC	1463	1463	1454	6	6	6	815	815	806	642	642	642						
*Pet*D	LSC	1203	1203	1203	8	9	8	720	720	720	475	474	475						
*rpl*16	LSC	1581	1581	1578	9	9	9	1173	1173	1170	399	399	399						
*rpl*2 [a]	IR	1490	1490	1495	391	391	397	665	665	667	434	434	431						
*rpo*C1	LSC		2851	2854		430	430		802	805		1619	1619						
*rps*16	LSC	1169	1168	1161	40	40	40	884	883	885	245	245	236						
*trn*A-UGC [a]	IR	875	875	875	37	37	38	802	802	802	36	36	35						
*trn*E-UUC [a]	IR	1025	1025		32	32		953	953		40	40							
*trn*K-UUU	LSC	2558	2558	2566	37	37	37	2486	2486	2500	35	35	29						
*trn*L-UAA	LSC	623	623	623	35	35	37	538	538	536	50	50	50						
*trn*T-CGU	LSC	778	776		35	35		699	697		44	44							
*ycf*3	LSC	1974	1974	1966	124	124	124	729	728	724	230	230	230	753	754	735	138	138	153

P. j= *Prosopis juliflora*, P. f = *Prosopis farcta*, P. c= *Prosopis cineraria*, IR: Inverted repeat, LSC: Large single copy, SSC: Short single copy) a (Duplicated gene.) *) The *ycf1* coding sequence is divided into 5, -ycf1and 3, -ycf1, which are located in the small single-copy region and inverted repeat region, respectively

Table 3. Genes function in the chloroplast genomes Prosopis species

Category	Group of Genes	Name of Genes
Self-replication	Large subunit of ribosomal proteins	rpl14, rpl16, rpl2, rpl2, rpl20, rpl23, rpl23, rpl32, rpl33, rpl36
	Small subunit of ribosomal proteins	rps11, rps12, rps12, rps14, rps15, rps16, rps18, rps19, rps2, rps3, rps4, rps7, rps7, rps8
	DNA dependent RNA polymerase	rpoA, rpoB, rpoC1, rpoC2
	rRNA genes	rrn4.5, rrn5, rrn16, rrn23
	tRNA genes	trnA-UGC, trnE-UUC trnF-GAA, trnfM-CAU, trnG-UCC, trnD-GUC, trnC-GCA, trnY-GUA trnH-GUG, trnI-CAU, trnL-UAA, trnM-CAU, trnI-GAU, trnN-GUU, trnL-CAA, trnK-UUU, trnL-UAG, trnP-GGG, trnP-UGG, trnQ-UUG, trnR-ACG,trnR-UCU, trnS-GCU, trnS-GGA,trnS-UGA, trnT-GGU, trnT-UGU, trnV-GAC, trnV-UAC, trnW-CCA
Photosynthesis	Photosystem I	psaA, psaB, psaC, psaI, psaJ
	Photosystem II	psbA, psbB, psbC, psbD, psbE, psbF, psbI, psbJ, psbK, psbM, psbN, psbT, psbZ, ycf3
	Cytochrome b6/f complex	petA, petB, petD, petG, petL, petN
	ATP synthase	atpA, atpB, atpE, atpF, atpH, atpI
	Rubisco	rbcL
	Subunits of NADH-dehydrogenase	ndhA, ndhB, ndhB, ndhC, ndhD, ndhE, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK
Other Genes	Maturase	matK
	Protease	clpP
	Envelop membrane protein	cemA
	Subunit of Acetyl-CoA-carboxylase	accD
	c-type cytochrome synthesis gene	ccsA
Unknown	Conserved open reading frames	ycf1, ycf2, ycf2, ycf4

The size and the borders of each gene found in the three prosopis species relatively to the whole CPG: It was revealed the start bp and the end bp for each studied gene that found inside the three *Prosopis* species genome, in addition to, present and absent of some genes in each species, also it included documented the charge of each gene according to the site encoded for its specific codon. Furthermore, the amino acids and their sites that contain negative charges or positive charge for each gene is illustrated in which has been published in National Center for Biotechnology Information (NCBI) website for 3 species and documented under accession numbers: MZ073640 for *Prosopis juliflora*, MZ073639 for *Prosopis farcta* and MZ073638 for *Prosopis cineraria*. Within the genome of the three *Prosopis* species we indicated different codons and its encoded amino acid in addition to its frequency throughout the genomes. In these CPG the most abundant codons were ATT (n= 2066, n=1924, n=1981), AAA (n = 1995, n=1705, n=2015) and TTT (n= 1887, n=1753, n=1859) *in P. farcta*, *P. cineraria* and *P. juliflora*, respectively, which encodes Isoleucine, lysine and Phenylalanine, respectively. The least frequently used codons were TAG (n = 127, n= 135, n=107) and TGA (n = 121, n= 138, n= 125) encode stop codon in *P. farcta*, *P. cineraria* and *P. juliflora*, respectively (CPG).

Simple Sequence Repeats (SSRs) provides a glimpse into the genome: The simple sequence repeats (SSRs) found in CPG were determined, in population genetics and evolutionary studies, SSRs were frequently used as genetic markers. SSRs, also known as microsatellite markers, are made up of one to sex bp repeat units in the sequence. The three Prosopis species CPG and two other CPG from the subfamily Caesalpinioideae were examined for simple sequence repetitions. SSRs per species ranged in total number from 184 in *L. trichandra* to 212 in *P. juliflora* the detected repeats were of higher percentage of mononucleotide repeats, 2-6 dinucleotide repeats, 2-21 trinucleotide repeats, 1-3 tetranucleotide repeats, pentanucleotide repeats, 4-6 and the hexanucleotide repeats 1-4 was only observed in *Prosopis Farcta* and *Leucaena trichandra* CPG (Figure 2). The most SSRs were discovered in *P. juliflora* (212 SSRs), while the fewest SSRs were discovered in *L. trichandra* (184 SSRs) (Figure 2). The majority of SSRs in these CPG were

discovered to be mononucleotide repeats (*P. juliflora*, *P. farcta*, and *P. cineraria*), comprising 92%, 89.2% 91.6%, of total SSRs, respectively (Figure 3).

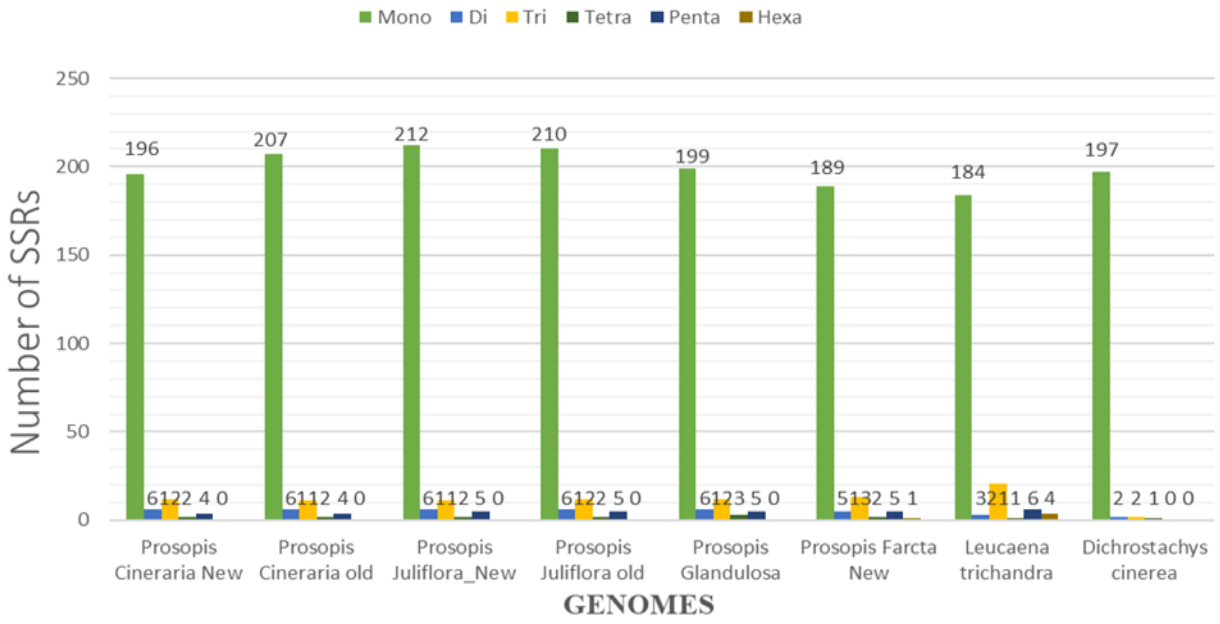


Figure 2. simple sequence repeats (SSRs) analysis in the, *P. farcta* and *P. cineraria* and related CPG species. SSR counts for both the entire genome and the coding area

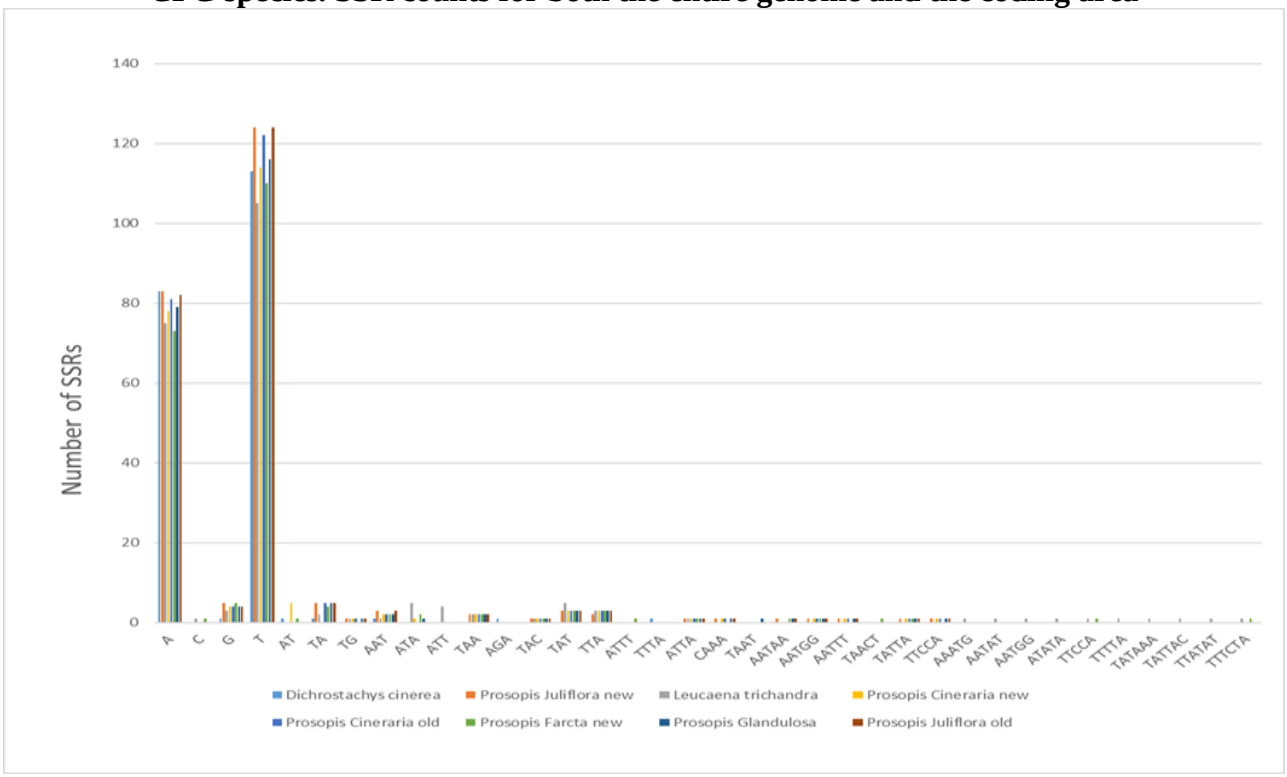
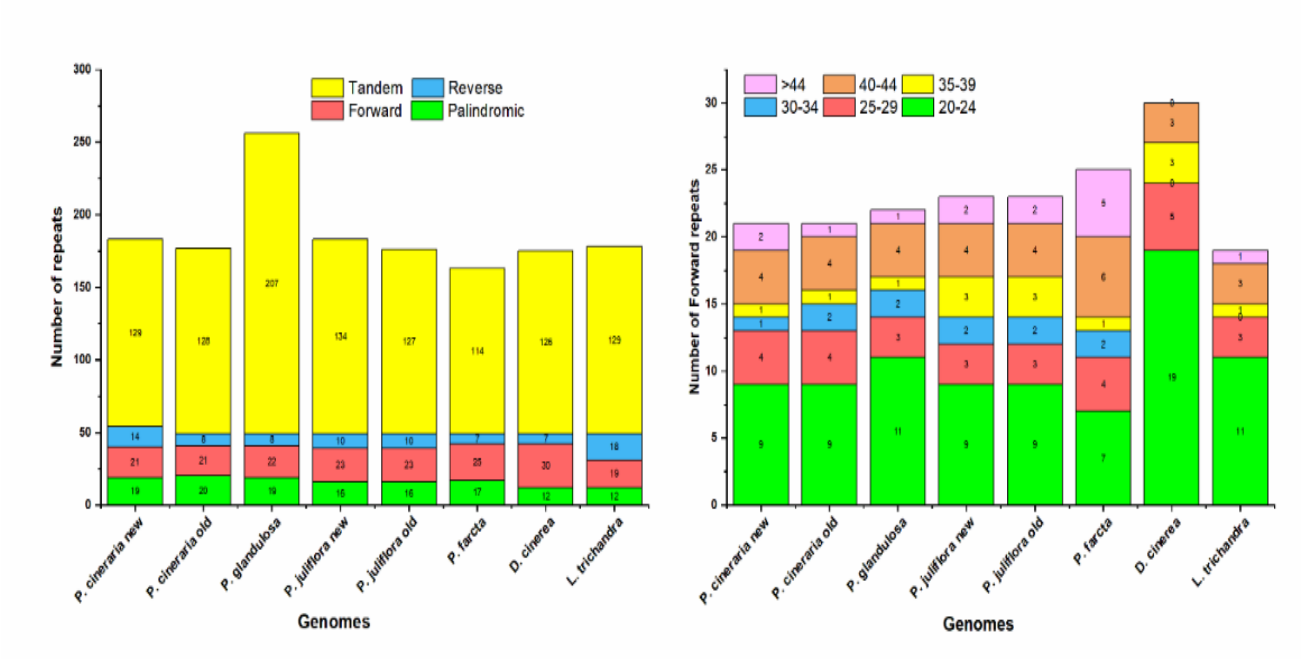


Figure 3. shows the frequency of SSR motifs in various repeat class types

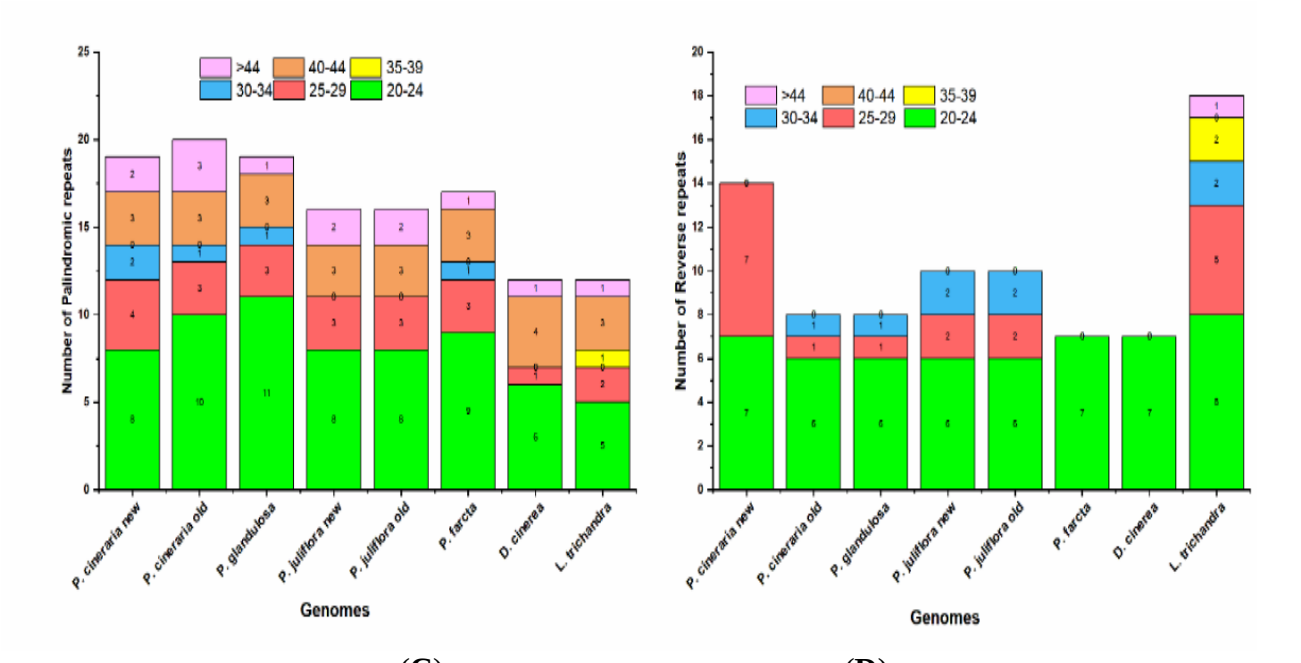
In the *P. juliflora*, *P. cineraria*, and *P. farcta* CPGs, respectively, 183, 183, and 163 repetitions were found. The *P. juliflora* genome includes 23 forward, 16 palindromic,10 reverse, and 134 tandem repeats, while *P. cineraria* CPG comprises 21 forward,14 reverse, 129 and 19 tandem and palindromic repeats, respectively while 25, 7, 17, and 114 forward, reverse, palindromic and tandem repeats were detected, respectively in *P. farcta* CPG (Figure 4). Likewise, about 177,256,176,175 and 178 Additionally, complete repeats in connected CPGs were found (4) in old *P. cineraria*, *P. glandulosa*, old *P. juliflora*, *D. cinerea*, and *L. trichandra*, respectively (Figure 4). With 20 palindromic

repeats, old *P. cineraria* includes the highest count of palindromic, while *D. cinerea* and *P. farcta* comprises the most repetitions moving forward (30,25), and the uppermost repeats of tandem was observed in *P. glandulosa* (207). We additionally noted that *D. cinerea, and L. trichandra* contained the lowest count of palindromic repeats (12) while *L. trichandra* has the minimum number of forward repeats (19) and high reverse repeats (18) (Figure 4).



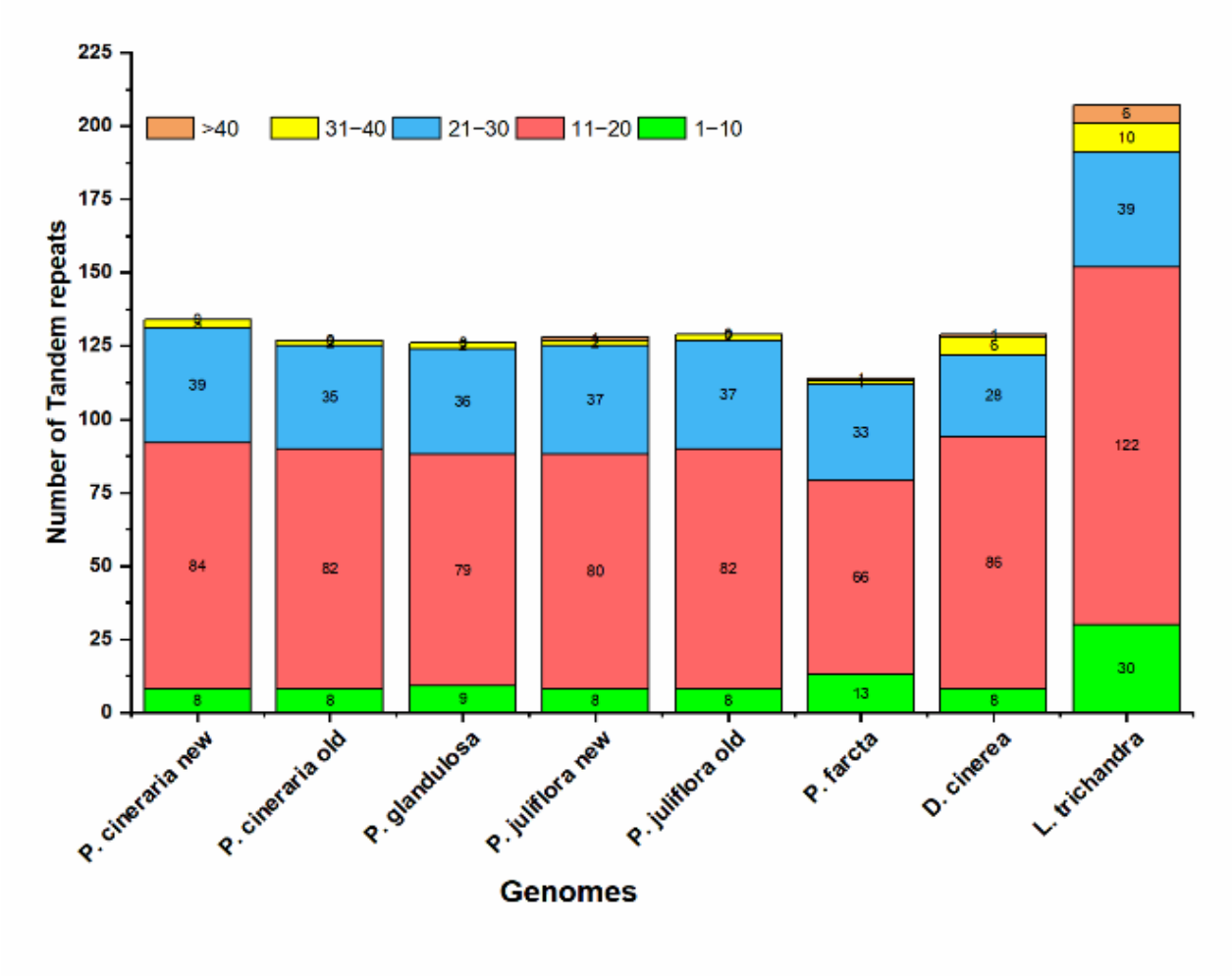
(**A**)

(**B**)



(**C**)

(**D**)



(F)

Figure 4. The analysis of repetitive sequences in *P. juliflora, P. farcta, P. cineraria*, and associated CPG. Numbers of the four repeat types, frequencies of palindromic repeats by length, forward repeats by length, reverse repeats by length, and frequencies of tandem repeats by length are shown in the graphs in (A), (B), (C), (D), and (E), respectively(F).

Phylogenetic Relationships: Using 27 whole chloroplast genomes, the phylogenetic link between *P. cineraria, P. juliflora*, and *P. farcta* was established within the Caesalpinioideae (Leguminosae) subfamily (Figure 5). The maximum parsimony (MP) approach was used for the phylogenetic analysis. Furthermore, the results obtained indicated that the species of the genus Prosopis are monophyletic and closely linked to *Adenanthera microsperma, Leucaena trichandra*, and *Dichrostachys cinerea* within the subfamily Caesalpinioideae (Figure 5), the resulting tree from Maximum parsimony analyses were congruent equal 100 in most boot strap value all relationships (Figure 5). The three studied species were clustered in one clade and are separated into two sub clades. Sub clade 1 which is includes *Prosopis farcta* and *Prosopis glandulosa*, in sub clade 2 containing *Prosopis cineraria* and *Prosopis juliflora* in the same node. Within the second clade is found sister species such as *D. cinerea* and *L. trichandra.*

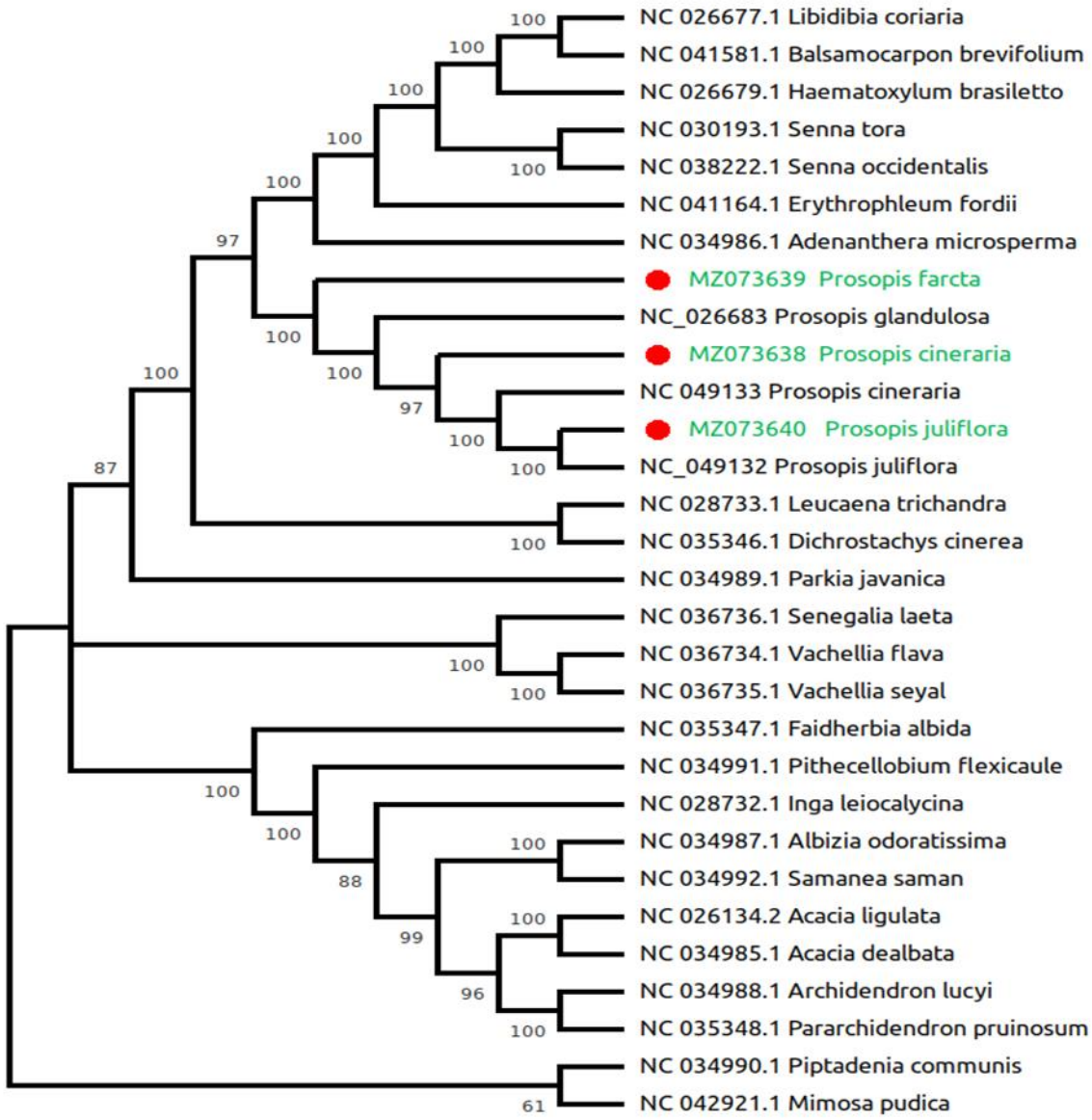


Figure 5. shows the phylogenetic trees of the Caesalpinioideae subfamily species P. juliflora, P. farcta, and P. cineraria (Leguminosae). MP techniques were used to evaluate the complete genomic data collection. In the MP trees, numbers upper and lower the branches reflect bootstrap values. The places of P. cineraria, P. farcta, and P. juliflora are indicated by the green color and red dots

This study revealed the first complete CPG sequence for three Prosopis species in Jordan; *P. juliflora*, *P. farcta* and *P. cineraria* using Ion Torrent S5 sequencing methods. The detected genomes were also compared with other species (within the subfamily Caesalpinioideae) genomes that are available in the NCBI. Similar findings were claimed by Wicke et al., (52) when investigating the CPGs in land plants such as *Cycas*, its size was ranged in length between 92 to 162 kb in angiosperms and had a circular and quadrilateral structure, which consisted of two copies of the inverted repeat regions (IR), a small single copy region (SSC) and a large single copy region (LSC). The same structure including the size and the shape was reported previously for the Leguminosae family lima bean *Phaseolus lunatus L.* CPGs (50). Wang et al., (51) studied the CPGs of some Prosopis species and found to be range in size from 161,240 bp in *D. cinerea* to 164,692 bp in *L. trichandra*, which encoded 128 *and* 129 in genes *D. cinerea and L. trichandra*, respectively. Asaf (5) reported *P. juliflora* to be163237 bp in size that account for 132 genes while *D. cinerea*, was 163677 bp of 131 genes. The *P. glandulosa* was of 163040 bp in size that has 128 genes (14). The length of the inverted repeat region in tobacco (*Nicotiana*

tabacum) was about 20 - 28 kb (46)v that was almost similar to the findings of Chumley et al., (10) who indicated it to be within the reported range. Our results indicated the inverted repeat region is very close, 26 kb in length, including 25,931 bp in all species. Asaf et al., (4) reported a length of 26,100 in *D. cinerea* and *L. trichandra*, which. The observed variation in length may resulted from between the different studied species may explained by either changes in the extent (expansions or contractions) due to idles (insertion or deletion) of certain segments in the genome as a result of mutation(s) that appeared through evolutionary process that also affected by selection, genetic drift, mutation and migration (16).In addition to the changes in genome size may also be explained by to the existed difference in the LSC and SSC areas, instead of the IR region contracting and expanding (47).As described previously by Qian et al., (42) on *Salvia miltiorrhiza* the SSC and LSC in two inverted repeat regions discovered in our study were more divergent than regions in all Prosopis species. The CPG of *P. juliflora* was found to encode 126 functional genes, including 82, 36 and 8 protein-coding, tRNA and rRNA genes, respectively as appeared in our study, all of which are comparable to the numbers found in other related species (*Stryphnoden dronadstrin*) (12). Moreover, twenty-one genes (thirteen coding proteins genes and eight tRNA genes) having introns were noticed in though our genetic analysis, interestingly, those detected two genes *(clp*P and *ycf*3) that contain two introns each, *ycf*3 gene has a functionally regulate the buildup of the photosystem I and plays a key role in photosynthesis (26,13). While *clp*P play an important role in regulating protease enzyme secretion (12). Our research revealed that the *ycf1* gene is unevenly split, with one copy of the 3' intron and exon being revealed in the reagon of IR and the 5' exon being detected in the SSC region. At the contrary Asaf et al., (5) found *rps*12 the gene was divided into two regions, and this gene is considered as a trans-spliced gene, and quite prevalent in plant CPG (35,55). Our results indicated the *acc*D gene to be existed in *P. juliflora* CPG while absent in *P. cineraria* and *P. farcta* CPGs, whole the *P. farcta* CPG contains *trn*G-GCC,*trn*I-GAU and *trn*V-UAC genes and was not detected in both *P. juliflora* and *P. cineraria* CPGs. The ycf1 was only observed in *P. cineraria*. This is the first to be published and no single record previously mention this finding. single sequence repeats (SSRs) in chloroplasts are an essential molecular marker, which are widely utilized in biogeographic studies, plant population genetics, polymorphism studies and evolutionary research (33,34). The benefits of the SSRs detected in this study can be exploited in evolutionary studies of the Prosopis genus and can be extended as a useful material in conservation strategies of the genus. A total number of 212 in *P. juliflora* ,196 in *P. cineraria*, and 189 in *P. farcta* were detected in our study. Similarly, Asaf et al., (3) studied the number of SSRs of some Prosopis species and found to be between 210 in *P. juliflora*, 207 in *P. cineraria*, and 184 in *L. trichandra.* Our findings in the SSRs number were more complex than that had been anticipated in the earlier study (4), as we register repetitions of repeat 2-6 dinucleotide repeats, 2-21 trinucleotide repeats, 1-3 tetranucleotide repeats, pentanucleotide repeats, 4-6 and the hexanucleotide repeats 1-4 was only observed in *P. farcta* and *L. trichandra* CPG while he didn't observe the hexanucleotide repeats in his studied species. *P. farcta* genome was investigated for the first time in our study, therefore, there are no previous studies to compare it with other, and the fact that this plant is considered one of the native plants that grow in Jordan. Ebert and Peakall (15) reported that, intra-species difference in CPG are expected to be mononucleotide. The obtained results were consistent with earlier observations that CPG SSRs often contain tandem adenine or thymine repeats and infrequently tandem guanine or cytosine repeats. (56), the detection of AT-rich SSRs in all species CPGs were similar to other plant species (15). In our study, a number of 183, 183 and 163 repeats were detected in the *P. j, P. c*, and *P. f* CPGs, respectively. The total repeats were also indicated by Asaf et al., (5) who claimed them to be related to CPG showing 177, 256, 176, 175 and 178 repeats in *P. cineraria, P. glandulosa, old P. juliflora, D. cinerea,* and *L. trichandra*, respectively. The

presence of these high differences between chloroplast genomes is very important and has excellent value for further studies on plant breeding and molecular markers (11). Molecular, evolutionary, and phylogenetic research have all benefited from CPG. There are numerous techniques that compare entire genome sequences although the many difficulties were associated in the phylogenetic analysis of complicated node level of but recently explained and simplified by Hohmann et al., (21) who have enriched our knowledge in understanding of angiosperm evolutionary resemblance. The phylogenetic relationships of *P. juliflora*, *P. cineraria*, and *P. farcta* Using entire chloroplast genomes from 27 plant CPGs, were identified within the subfamily Caesalpinioideae (Leguminosae). Asaf et al., (3) studied the phylogenetic trees of *P. juliflora*, *P. glandulosa*, and *P. cineraria* the resulted showed that High bootstrap support indicates that *P. juliflora* is more resemble to *P. cineraria than P. glandulosa* in the subfamily Caesalpinioideae. In our study we found that the three species bunched in one clade and branched into two sub clades. Where the tree showed that *P. cineraria* and *P. juliflora* are closer to each other, then the *P. farcta*. The *Prosopis* species are closely related to *Adenanthera microsperma*, *Leucaena trichandra* and *Dichrostachys cinerea*, the resulting tree from Maximum parsimony analyses were congruent equal 100 in most boot strap value all relationships. All the three Prosopis species (*P. j, P. f,* and *P. c*) split into two major subclades and are strongly supported in one clade. The similarity in same node appeared equal 100 in Sub clade 1 which is includes *P. farcta* and prosopis glandulosa, in sub clade 2 containing *Prosopis cineraria* and *Prosopis juliflora* in the same node appeared equal 97. within the second clade is found sister species such as *D. cinerea* and *L. trichandra*. The observed results presented by (53) are in agreement with this research results in inverted repeat region in three species (*P. j, P. c,* and *P. f*) 25931, 25931, and 25932, respectively. While there were differences in the length of the total genome and the number of genes. In the other hand, the GC content reach in *A. crassicarpa* of 35.3% While it was 35.9% in the three species of Prosopis (*P. j, P. c*, and *P. f*) as shown by (57). The observed results are in agreement with (2) who presented that in small single copy region in three species (*P. j, P. c*, and *P. f*) 18880, 18865, and 18882, respectively. While there were differences in the length of the total genome and the number of genes. In our study, we found that the number of protein coding genes was 82 in the three studied species, while in the study Asaf et al., it was 89 in *V. nilotica* and 87 in *S. Senegal*, while there was a similarity in the number of rRNA, which reached 8, a similarity also appeared in the number of genes in *P. farcta* and *S. Senegal*, which reached 127. The obtained results revealed for the first time the complete CPG of *Prosopis* species found in Jordan and its relationship with other published data. The detected sequence, genes, and mutations could be used in genome association studies. Furthermore, the deep molecular findings open the floor for the researchers to compare their local species with ours for any evolutionary forces that may alter the genome of local trees or shrubs as a reflect of many environmental or genetic forces/factors.

Acknowledgments: The authors would like to thank the Deanship of Research at Jordan University of Science and Technology for supporting and fund the research project (project number 217/2021).

REFERENCES

1. Al-Soqeer, A.A., Q.D. Alsubaie, M. I. Motawei, H.M. Mousa and A.M. Abdel-Salam. 2017. Isolation and identification of allergens and biogenic amines of Prosopis juliflora genotypes. Electron J Biotechnol, 30: 24–32

2. Asaf, S, A. Khan, AL . Khan, A. Al-Harrasi, and A. Al-Rawahi. 2019. Complete chloroplast genomes of Vachellia nilotica and Senegalia senegal: comparative genomics and phylogenomic placement in a new generic system. PLoS One,14(11):e0225469

3. Asaf, S, AL. Khan, A. Khan, and A. Al-Harrasi. 2020. Unraveling the chloroplast genomes of two prosopis species to identify its genomic information, comparative analyses and phylogenetic relationship. Int J Mol Sci, 21(9):3280

4. Asaf, S, AL. Khan, A. Khan, and A. Al-harrasi. 2020. Unraveling the Chloroplast

Genomes of Two Prosopis Species to Identify Its Genomic Information, Comparative Analyses and Phylogenetic Relationship. International Journal of Molecular Sciences, 21(9), 3280

5. Asaf, S, AL. Khan, A. Khan, and A. Al-harrasi. 2020. Unraveling the Chloroplast Genomes of Two Prosopis Species to Identify Its Genomic Information , Comparative Analyses and Phylogenetic Relationship. International Journal of Molecular Sciences, 21(9), 3280

6. Asaf, S, AL. Khan, AR. Khan, M. Waqas, S-M. Kang, M.A. Khan, S-M. L. Lee and I. Lee. 2016. Complete Chloroplast Genome of Nicotiana otophora and its Comparison with Related Species. Frontiers in plant science, (7):843

7. Asaf, S, M. Waqas, A.L. Khan, M.A. Khan, S-M.Kang, Q.M. Imran, R. Shahzad, S. Bilal, B-W. Yun and I-J. Lee. 2017. The complete chloroplast genome of wild rice (Oryza minuta) and its comparison to related species. Front Plant Sci, 8: 304

8. Beier, S., T. Thiel, T. Münch, U. Scholz, and M. Mascher. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics, 33(16):2583–5

9. Burkart, A. 1976. A monograph of the genus Prosopis (Leguminosae subfam. Mimosoideae). Catalogue of the recognized species of Prosopis. Journal of the Arnold arboretum, 57: 450-525

10. Chumley, T.W, J.D. Palmer, J.P. Mower, H.M. Fourcade, P.J Calie, J.L. Boore, and R.K. Jasen. 2006. The complete chloroplast genome sequence of Pelargonium× hortorum: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. Mol Biol Evol,23 (11):2175–90

11. Daniell, H., C.S. Lin, M .Yu, and W.J Chang. 2016. Chloroplast genomes: Diversity, evolution, and applications in genetic engineering. Genome Biol [Internet], 17(1):1–30. Available from: http://dx.doi.org/10.1186/s13059-016-1004-2

12. de Souza, U.J.B, R . Nunes, C.P. Targueta, J.A.F. Diniz-Filho and M.P. de Campos Telles. 2019. The complete chloroplast genome of Stryphnodendron adstringens (Leguminosae-caesalpinioideae): Comparative analysis with related Mimosoid species. Sci Rep, 9(1):1–12

13. Dong, W., C. Xu, C.Li, J. Sun, Y. Zuo and S. Shi.2015. ycf1, the most promising plastid DNA barcode of land plants. Sci Rep, 5 (1):1–5.

14. Dugas, D.V, D. Hernandez, E.J.M. Koenen, E. Schwarz, S. Straub and C.E. Hughes. 2015. Mimosoid legume plastome evolution: IR expansion, tandem repeat expansions and accelerated rate of evolution in clpP. Sci Rep,5 (1):1–13

15. Ebert, D. and R.O.D. Peakall. 2009. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. Mol Ecol Resour, 9 (3):673–90

16. Falconer, D.S and T.F.C. Mackay. 1996. Introduction to quantitative genetics. Longmans Green, Harlow, Essex, UK. Introd to Quant Genet 4th ed Longmans Green, Harlow, Essex, UK, Essex, England, 254-256

17. Felker, P., G.H. Cannell, P.R. Clark, J.F. Osborn and P. Nash. 1983. Biomass production of Prosopis species (mesquite), leucaena, and other leguminous trees grown under heat/ drought stress. For Sci, 29 (3):592–606

18. Frailey, D.C., S.R. Chaluvadi, J.N. Vaughn, C.G. Coatney and J.L. Bennetzen. 2018. Gene loss and genome rearrangement in the plastids of five Hemiparasites in the family Orobanchaceae. BMC Plant Biol, 18 (1):30

19. Frazer, K.A., L. Pachter, A. Poliakov, E.M. Rubin and I. Dubchak. 2004. VISTA: computational tools for comparative genomics. Nucleic Acids Res, 32 (suppl_2):W273–9

20. Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol, 14 (7):685–95

21. Hohmann, N., R. Schmickl, T-Y.Chiang, M. Lučanová, F. Kolář and K. Marhold. 2014. Taming the wild: resolving the gene pools of non-model Arabidopsis lineages. BMC Evol Biol, 14 (1):1–21.

22. Karaman, M.W. and J.G. Hacia. 2003. Genomes. Oxford University Press, 432-433

23. Katoh, K. and H. Toh. 2010. Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics, 26 (15):1899–900

24. Kearse, M., R. Moir, A.Wilson, S. Stones-Havas, M. Cheung and S. Sturrock. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28 (12):1647–9

25. Khan, A., S. Asaf, AL Khan, A. Al-Harrasi, O. Al-Sudairy and N.M. AbdulKareem. 2019. First complete chloroplast genomics and comparative phylogenetic analysis of Commiphora gileadensis and C. Foliacea: Myrrh producing trees, PLoS One, 14(1), e0208511

26. Kikuchi, S., J. Bédard, M. Hirano, Y. Hirabayashi, M. Oishi and M. Imai. 2013. Uncovering the protein translocon at the chloroplast inner envelope membrane. Science, 339(6119):571–4

27. Kumar, S., M. Nei, J. Dudley and K. Tamura. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinformatics, 9 (4), 299-306

28. Kurtz, S., J.V. Choudhuri, E. Ohlebusch, C. Schleiermacher, J. Stoye and R. Giegerich. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Research, 29 (22):4633–42

29. Küster, E. and S.T. Williams. 1964. Selection of media for isolation of streptomycetes. Nature, 202 (4935):928–9

30. Langmead, B. and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods, 9 (4):357

31. Liu, D., Y. Cui, S. Li, G. Bai, Q. Li and Z. Zhao.2019. A New Chloroplast DNA Extraction Protocol Significantly Improves the Chloroplast Genome Sequence Quality of Foxtail Millet (Setaria italica (L.) P. Beauv.). Sci Rep, 9 (1):1–9

32. Lohse, M., O. Drechsel and R. Bock. 2007. Organellar Genome DRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet, 52 (5–6):267–74

33. Martin, G., F-C. Baurens,C. Cardi, J-M. Aury and A. D'Hont. 2013. The complete chloroplast genome of banana (Musa acuminata, Zingiberales): insight into plastid monocotyledon evolution. PLoS One, 8 (6):e67350

34. Melotto-Passarin, D.M., E.V. Tambarussi, K. Dressano, V.F. De Martin and H .Carrer. 2011. Characterization of chloroplast DNA microsatellites from Saccharum spp and related species. Genet Mol Res, 10 (3):2024–33

35. Menezes, A.P.A., L. C. Resende-Moreira, R.S.O. Buzatti, A.G. Nazareno, M. Carlsen and F.P. Lobo. 2018. Chloroplast genomes of Byrsonima species (Malpighiaceae): comparative analysis and screening of high divergence sequences. Sci Rep, 8 (1):1–12

36. Merilä. J. and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology, 14 (6), 892-903

37. Michel-López, C. Y., D. González-Mendoza and O. Grimaldo-Juarez. 2013. Fast protocol for extraction of DNA from Prosopis spp leaves (plant adapted to arid environment) without liquid nitrogen. Genetics and Molecular Research, 12 (3):4090–4

38. Nazir, A., Z. Dong, J. Liu, X. Zhang, R.A. Tahir and N. Ashraf. 2020. Sequence Analysis of a Jumbo Bacteriophage, Xoo-sp14, That Infects Xanthomonas oryzae pv. oryzae. Microbiology Resource Announcements, 9 (48), e01072-20

39. Neuhaus, H.E and M.J. Emes. 2000. Nonphotosynthetic metabolism in plastids. Annual review of plant biology, 51, 111.=40. Pasiecznik, N.M., P. Felker, P.J.C. Harris, L.n. Harsh, G. Cruz and J.C. Tewari. 2001. The Prosopis juliflora-Prosopis pallida complex: a monograph, (172). HDRA Coventry

41. Posada, D. 2008. jModelTest: phylogenetic model averaging. Molecular biology and evolution, 25 (7), 1253-1256

42. Qian, J., J. Song, H. Gao, Y. Zhu, J. Xu and X. Pang. 2013. The Complete Chloroplast Genome Sequence of the Medicinal Plant Salvia miltiorrhiza. PLoS One, 8 (2), e57607

43. Rodríguez-Ezpeleta, N., H. Brinkmann, S.C. Burey, B. Roure, G. Burger and W. Löffelhardt. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Current biology, 15 (14):1325–30.

44. Ronquist, F. and J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics, 19 (12), 1572-1574

45. Shi, C., N. Hu, H. Huang, J. Gao, Y-J. Zhao and L-Z. Gao.2012. An improved chloroplast DNA extraction procedure for whole plastid genome sequencing. PLoS One, 7 (2):e31468

46. Shinozaki, K., M. Ohme, M. Tanaka, T. Wakasugi, N. Hayashida and T. Matsubayashi. 1986. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. The EMBO journal, 5 (9), 2043-2049

47. Su, Q., L. Liu, M. Zhao, C. Zhang, D. Zhang and Y. Li. 2020. The complete chloroplast genomes of seventeen Aegilops tauschii: genome comparative analysis and phylogenetic inference. PeerJ, 8:e8678

48. Swofford, D. L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4.,(Sinauer Associates Inc.: Sunderland, MA, USA.).

49. Tadros, M.J., A. Al-Assaf, Y.A. Othman, Z. Makhamreh and H. Taifoure.2020. Evaluating the Effect of Prosopis juliflora, an Alien Invasive Species, on Land Cover Change Using Remote Sensing Approach. Sustainability, 12 (15):5887

50. Tian, S., P. Lu, Z. Zhang, J.Q. Wu, H. Zhang and H. Shen. 2021. Chloroplast genome sequence of Chongming lima bean (Phaseolus lunatus L.) and comparative analyses with other legume chloroplast genomes. BMC Genomics, 22 (1):1–15.

51. Wang, Y-H., X-J. Qu, S-Y. Chen, D-Z. Li, and S. Yi. 2017. Plastomes of Mimosoideae: structural and size variation, sequence divergence, and phylogenetic implication. Tree Genet genomes, 13 (2):41

52. Wicke, S., G. M. Schneeweiss, C. W. Depamphilis, K. F. Müller, and D. Quandt. 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. Plant molecular biology, 76 (3), 273-297

53. Williams, A.V., L.M. Boykin, K.A. Howell, P.G. Nevill and I. Small. 2015. The complete sequence of the Acacia ligulata chloroplast genome reveals a highly divergent clpP1 gene. PLoS One, 10 (5):e0125768

54. Wyman, S, K., R. K. Jansen, and J. L. Boore. 2004. Automatic annotation of organellar genomes with DOGMA. Bioinformatics, 20 (17):3252–5

55. Yang, Z., T. Zhao, Q. Ma, L. Liang and G. Wang. 2018. Comparative genomics and phylogenetic analysis revealed the chloroplast genome variation and interspecific relationships of Corylus (Betulaceae) species. Frontiers in plant science, 9, 927

56. Yi, X., L. Gao, B. Wang, Y-J. Su and T. Wang. 2013. The complete chloroplast genome sequence of Cephalotaxus oliveri (Cephalotaxaceae): evolutionary comparison of Cephalotaxus chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. Genome biology and evolution, 5 (4), 688-698

57. Yue, X., Y. Yu, W. Gao, S. Chen, Z. Weng, and G. Ye. 2021. Complete chloroplast genome sequence of Acacia crassicarpa (Fabaceae). Mitochondrial DNA Part B, 6(8), 2249-2250.