Brazilian Journal of Biology®

ISSN 1519-6984 (Print)
ISSN 1678-4375 (Online)

Original Article

Structural conservation and SSR-driven divergence in *Lepidium* chloroplast genomes: balancing genomic stability with adaptive plasticity

Conservação estrutural e divergência impulsionada por SSR em genomas de cloroplastos de *Lepidium*: equilibrando estabilidade genômica com plasticidade adaptativa

Raed Al-Saharin[a]*

[a]Tafila Technical University, Department of Applied Biology, Tafila, Jordan

Abstract

The chloroplast genomes of 13 *Lepidium* species were analyzed to explore how structural conservation and SSR (Simple Sequence Repeat) dynamics balance genomic stability with adaptive divergence. While genome size and GC content remained highly conserved, reflecting strong functional constraints on photosynthetic machinery, repetitive elements exhibited striking lineage-specific diversity. Mid-length, AT-rich SSRs dominated across species, with conserved motifs (e.g., pentanucleotide TCCAT) likely underpinning structural integrity. In contrast, clade-specific innovations, such as absent octanucleotide repeats in *Lepidium cordatum* and *Lepidium draba* or unique pentanucleotide expansions, highlight adaptive plasticity in non-essential regions. These findings reveal a dual evolutionary strategy: chloroplast genomes maintain core functions through conserved architecture while permitting localized SSR-driven innovations, potentially enhancing ecological resilience. This interplay underscores how genomic stability and plasticity coexist to enable adaptation in dynamic environments, offering new insights into the evolutionary mechanisms shaping plant organelle genomes.

Keywords: chloroplast genomes, *Lepidium*, Simple Sequence Repeats (SSRs), genome evolution, adaptive divergence.

Resumo

Os genomas de cloroplastos de 13 espécies de *Lepidium* foram analisados para explorar como a conservação estrutural e a dinâmica de SSR (repetição de sequência simples) equilibram a estabilidade genômica com a divergência adaptativa. Enquanto o tamanho do genoma e o conteúdo de GC permaneceram altamente conservados, refletindo fortes restrições funcionais na maquinaria fotossintética, elementos repetitivos exibiram uma impressionante diversidade específica da linhagem. SSRs de comprimento médio e ricos em AT dominaram todas as espécies, com motivos conservados (por exemplo, pentanucleotídeo TCCAT) provavelmente sustentando a integridade estrutural. Em contraste, inovações específicas de clado, como repetições de octanucleotídeos ausentes em *Lepidium cordatum* e *Lepidium draba* ou expansões únicas de pentanucleotídeos, destacam a plasticidade adaptativa em regiões não essenciais. Essas descobertas revelam uma estratégia evolutiva dupla: os genomas de cloroplastos mantêm funções essenciais por meio de arquitetura conservada ao mesmo tempo que permitem inovações localizadas impulsionadas por SSR, potencialmente aumentando a resiliência ecológica. Essa interação ressalta como a estabilidade genômica e a plasticidade coexistem para permitir a adaptação em ambientes dinâmicos, oferecendo novos insights sobre os mecanismos evolutivos que moldam os genomas das organelas vegetais.

Palavras-chave: genomas de cloroplastos, *Lepidium*, Repetições de Sequência Simples (SSRs), evolução do genoma, divergência adaptativa.

1. Introduction

Lepidium is a genus of flowering plants in the *Brassicaceae* family (Mustard), involving over 200 globally distributed species. Commonly referred to as pepperworts or peppergrasses, they produce small white or yellow flowers and have pungent leaves and seeds that have a sharp peppery taste (Bona, 2020; German, 2014; Said and Kassahun, 2015). *Lepidium* species are highly adaptable and can be found in diverse habitats, including temperate regions, arid deserts, and coastal areas (Orth et al., 2006). Some species, such as *Lepidium sativum* (garden cress), are cultivated for their culinary and medicinal uses (Ghante et al., 2011; Shabbir et al., 2018; Shah et al., 2021).

*e-mail: raed.alsaharin@ttu.edu.jo
Received: February 12, 2025 – Accepted: April 30, 2025
Editor: Takako Matsumura Tundisi

Interest in the genus has been fuelled by its economic value, ecological role, and medicine potential. For example, *Lepidium meyenii* (Maca) is a well-known species native to the Andes, valued for its nutritional and aphrodisiac properties (Gonzales, 2012; Lee et al., 2011; Ruiz-Luna et al., 2005). Additionally, *Lepidium* species are studied for their phytochemical composition, including glucosinolates, flavonoids, and alkaloids, which contribute to their antioxidant, anti-inflammatory, and anticancer properties (Chatoui et al., 2016; Girmay et al., 2014; Hashmi et al., 2021).

Microsatellites, also known as simple sequence repeats (SSRs) are short repeat sequences of DNA that are tandemly repeated and 1-6 base pairs in length. These repeated elements are found across a broad range of organisms, from eukaryotes to prokaryotes, and even to some viruses (Alhawatema, 2023; Li et al., 2002). Underlined by the importance of simple sequence repeats in genetic studies is the fact that they are highly polymorphic. SSRs have emerged as one of the most common molecular markers because they exist in the genome in large numbers, are co-dominantly inherited, and are highly reproducible (Ellegren 2004; Karaoglu et al., 2005).

SSRs can be classified as mono-, di-, tri-, tetra-, penta- or hexanucleotide repeats based on their repeat motifs. They can occur in coding regions, non-coding regions, or regulatory regions of the genome, influencing gene expression, protein function, and genome evolution. Due to their high mutation rate, SSRs are often associated with genetic diversity and adaptation. However, they can also contribute to genetic disorders and diseases when mutations occur in functional regions (Alhawatema, 2023; Kalia et al., 2011).

Advancements in sequencing technologies and bioinformatics tools have revolutionized the development of SSR markers, enabling large-scale discovery and analysis. SSRs are now extensively used in genetic mapping, marker-assisted selection, and conservation genetics (Li et al., 2002).

In this work, we investigated structural variations of chloroplast genomes across 13 species of *Lepidium* in relation to the repetitive elements. Our goal is to locate simple sequence repeats (SSRs) in these species' chloroplast genomes and then characterize their distribution, motif configurations, availability, and putative evolutionary conservation. We hypothesize that distinct traits of chloroplast genomes such as genome size, GC content, and SSR frequency differ in species, suggesting ecological and evolutionary adaptations.

2. Material and Methods

2.1. Chloroplast genomic sequence of Lepidium species

Chloroplast genomes from 13 *Lepidium* species were selected and downloaded from NCBI GenBank (NCBI, 2025). The chloroplast genome sequences of *Lepidium* species were downloaded in files as FASTA format (Table 1).

2.2. Identification of SSRs

The identification of simple sequence repeats (SSRs) in the chloroplast genomes of *Lepidium* species was conducted using RepeatMasker software (version 4.1.5; Smit et al., 2015). The chloroplast genome sequences of each *Lepidium* species were uploaded to the software in FASTA format. Upon completion of the analysis, four output files were generated for each chloroplast genome, containing details on repeat locations, types, sequences, and units. The output files were imported into Microsoft Excel for further processing to facilitate the study of repeats across all chloroplast genomes of 13 *Lepidium* species.

2.3. Repeats analysis

The repetitive sequences in the chloroplast genomes of *Lepidium* species were analyzed for mono-, di-, tri-, tetra-, penta-, hexa-, hepta-, and octanucleotide repeats. All repeats were evaluated based on their number, frequency, and abundance. Additionally, the repeat density, and longest repeat sequences were determined. Furthermore, conserved repeats across all chloroplast genomes of *Lepidium* species, interspersed repeats, and the proportion of bases masked by these interspersed repeats were also examined.

3. Results

The chloroplast genomes of 13 *Lepidium* species were comprehensively characterized to evaluate genomic features, including the distribution of genomic elements, total SSR (Simple Sequence Repeat) counts, SSR density, longest SSR motifs, and SSR conservation patterns across all species, as detailed in Tables 1-6.

GenBank ID, genome size (bp), and GC content percentage are shown for the chloroplast genome of 13 *Lepidium* species (Table 1). The chloroplast genome sizes of the analyzed *Lepidium* species range from 153,177 bp (*L. cordatum*) to 154,997 bp (*L. sativum*), with a difference of approximately 1,820 bp, which is not considered a substantial variation. In terms of GC content, all species fall within a narrow range of 36.35% (*L. meyenii*) to 36.51% (*L. cartilagineum*, *L. cordatum*, *L. draba*, and *L. latifolium*). Interestingly, *L. sativum*, which has the largest chloroplast genome, exhibits a mid-range GC content of 36.42%, while *L. cordatum*, with the smallest chloroplast genome, has one of the higher GC values at 36.51%. This indicates no clear relationship between these species' chloroplast genome size and GC content.

The analysis of genomic elements in *Lepidium* chloroplast genomes (Table 2, Figure 1) reveals conserved and variable features. All species contain two Short Interspersed Nuclear Elements (SINE) with a length range between (126-127 bp), except *L. sativum*, which has a slightly longer SINE (134 bp) (Figure 2). Simple repeats vary in frequency (35-45), with *L. cordatum* and *L. draba* showing the highest counts (45), while low-complexity regions range from 8 (*L. draba*, *L. virginicum*) to 14 (*L. appelianum*). Small RNA regions, critical for gene regulation, mask the largest proportion of bases, with frequencies spanning 7 (*L. echinatum*) to 10 (multiple species) and masked base lengths between 3,333 bp (*L. echinatum*) and 3,651 bp (*L. appelianum*, *L. chalepense*). The percentage of masked bases ranges narrowly from 3.73% (*L. echinatum*) to 4.18% (*L. draba*), reflecting subtle

Table 1. List of the analyzed chloroplast genomes of *Lepidium* species in the study.

Organism	GenBank ID	Genome Size (bp)	GC Content (%)
Lepidium apetalum	NC_051540.1	154680	36.44
Lepidium appelianum	NC_077510.1	153803	36.48
Lepidium cartilagineum	NC_077509.1	153991	36.51
Lepidium chalepense	NC_077508.1	153810	36.47
Lepidium cordatum	NC_077507.1	153177	36.51
Lepidium draba	NC_077506.1	153675	36.51
Lepidium echinatum	NC_049660.1	154504	36.46
Lepidium ferganense	NC_077505.1	154754	36.43
Lepidium latifolium	NC_053830.1	153989	36.51
Lepidium meyenii	NC_034363.1	154403	36.35
Lepidium ruderale	NC_077504.1	154846	36.40
Lepidium sativum	NC_047178.1	154997	36.42
Lepidium virginicum	NC_009273.1	154743	36.47

Table 2. Shows the frequency of different genomic elements in the chloroplast genomes of 13 *Lepidium* species.

Organism	Analyzed	SINE		Simple Repeats		Low Complexity		Small RNA		Bases Masked (%)
	Genome	Number	Length (bp)	Number	Length (bp)	Number	Length (bp)	Number	Length (bp)	
Lepidium apetalum	154680	2	126	41	2053	12	616	10	3476	4.05
Lepidium appelianum	153803	2	126	39	2001	14	644	10	3651	4.17
Lepidium cartilagineum	153991	2	127	36	1835	10	531	9	3432	3.85
Lepidium chalepense	153810	2	126	41	2060	10	531	10	3651	4.14
Lepidium cordatum	153177	2	126	45	2178	13	582	9	3434	4.12
Lepidium draba	153675	2	126	45	2301	8	392	9	3607	4.18
Lepidium echinatum	154504	2	126	35	1797	11	515	7	3333	3.73
Lepidium ferganense	154754	2	126	41	2014	10	390	9	3432	3.85
Lepidium latifolium	153989	2	127	36	1834	10	531	9	3432	3.84
Lepidium meyenii	154403	2	126	44	2091	12	508	9	3432	3.98
Lepidium ruderale	154846	2	126	42	2070	8	376	10	3476	3.90
Lepidium sativum	154997	2	134	41	1852	12	484	10	3490	3.84
Lepidium virginicum	154743	2	126	42	2144	8	364	9	3432	3.92

interspecific differences in repetitive content (Figure 3). These patterns highlight conserved structural elements like SINEs alongside species-specific variability in repetitive regions, suggesting evolutionary constraints on core genomic features balanced with adaptive flexibility in non-coding regions.

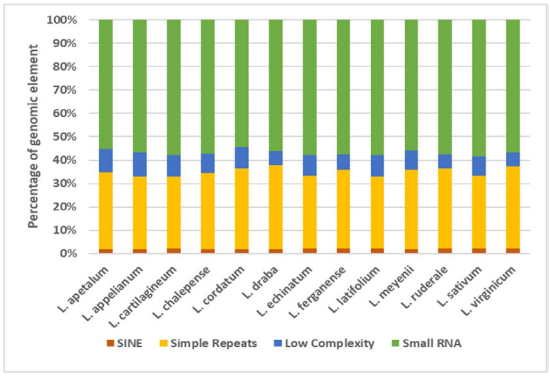

Figure 1. The percentage distribution of different types of genomic elements in the chloroplast genomes of several *Lepidium* species.

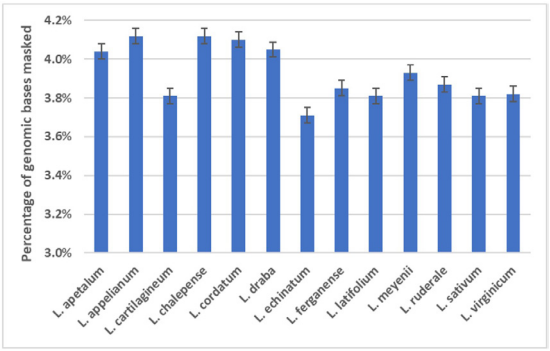

Figure 3. Comparison of the percentage (%) of total genomic bases masked by all genomic elements in the chloroplast genomes of various *Lepidium* species. The error bars are shown in each column.

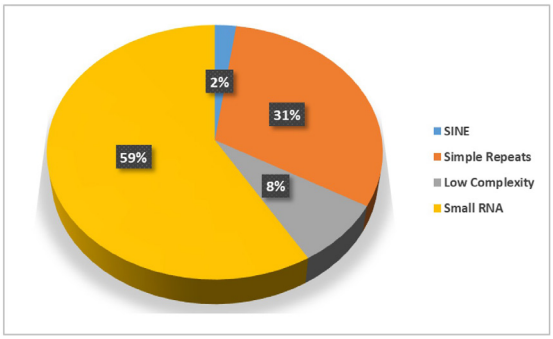

Figure 2. The percentage of different types of genomic elements in the chloroplast genome of *Lepidium sativum*.

The chloroplast genomes of *Lepidium* species show distinct patterns in SSR (simple sequence repeat) distribution (Table 3). Hexanucleotide (Hexa) repeats are the most frequent overall (90 total), followed by tetranucleotide (Tetra: 89) and trinucleotide (Tri: 88) repeats, indicating a preference for mid-length SSRs. Dinucleotide (Di: 84) and pentanucleotide (Penta: 83) repeats are moderately common, while mononucleotide (Mono: 45), heptanucleotide (Hepta: 38), and octanucleotide (Octa: 11) repeats are rare. Species-specific variations stand out: *L. cordatum* and *L. draba* have the highest Hexa counts (10 each), with *L. cordatum* also showing an unusually high Penta count (10). *L. sativum* has the most Mono repeats (6), while *L. virginicum* has the fewest Hepta repeats (1). Notably, *L. cordatum* and *L. draba* lack Octa repeats entirely, unlike other species that retain at least one. These trends suggest evolutionary or functional pressures shaping SSR diversity, balancing shared genomic features with lineage-specific adaptations.

The relative density of SSR repeats in *Lepidium* chloroplast genomes (Table 4) reveals a clear dominance of dinucleotide (Di) repeats (total: 5,744 bp), followed by hexanucleotide (Hexa: 4,872 bp) and tetranucleotide (Tetra: 4,362 bp) repeats. Mid-length SSRs (Di, Hexa, Tetra) collectively form the majority, while shorter (mononucleotide: 1,166 bp) and longer repeats (heptanucleotide: 1,926 bp; octanucleotide: 616 bp) are less frequent (Table 4). Species-specific differences are pronounced: *L. ruderale* shows the highest Di density (600 bp), whereas *L. draba* and *L. cordatum* exhibit elevated Hexa repeats (550 and 473 bp) but lack Octa repeats entirely.

L. virginicum stands out with the highest Tri (407 bp) and Penta (432 bp) densities but the lowest Hepta (40 bp). These variations suggest evolutionary or functional pressures influencing SSR accumulation, with AT-rich motifs (e.g., Di and Hexa) dominating, likely due to chloroplast genome biases. Overall, patterns reflect a balance between conserved genomic architecture and lineage-specific adaptations shaping SSR distribution.

The longest SSR motifs in *Lepidium* chloroplast genomes show a strong prevalence of dinucleotide (Di) repeats, particularly TA (e.g., TA (148 bp) in *L. appelianum*), with AT and TT also recurring frequently (Table 5). Trinucleotide (Tri) motifs like TAT, TTT, and TAA appear across species, while tetranucleotide (Tetra) repeats such as AATA and TTTA dominate in most genomes. Notably, *L. ferganense* displays an exceptionally long pentanucleotide (Penta) motif ATTAA (121 bp), and *L. virginicum* features a unique Penta repeat TATTA (112 bp). Hexanucleotide (Hexa) sequences like TGATAT and ATTAGA are widespread, with *L. cartilagineum* and *L. appelianum* showing extended variants (e.g., TGATAT 83 bp). Mononucleotide (Mono) repeats are primarily T or A homopolymers, reaching up to 35 bp in *L. ruderale*. While certain motifs (e.g., AATA in Tetra) are conserved, others vary significantly between species, such as *L. sativum*'s distinct Hexa motif ATACAA (39 bp). These patterns emphasize AT-rich biases in chloroplast DNA and suggest evolutionary fine-tuning, where conserved motifs maintain structural roles, while species-specific repeats may reflect adaptive or mutational dynamics unique to individual lineages.

Conserved SSR motifs across *Lepidium* chloroplast genomes reveal striking uniformity in core repeat types, with subtle length variations (Table 6). Dinucleotide (Di) repeats TA (124-148 bp) and TT (60-70 bp) are universally present, alongside the trinucleotide (Tri) motif TAA (43 bp). Tetranucleotide (Tetra) repeats predominantly feature AATA (66 bp), except in *L. cordatum* (74 bp). Pentanucleotide (Penta) motifs are strictly conserved as TCCAT (41 bp), while hexanucleotide (Hexa) repeats include ATACAA (39 bp),

Table 3. Total repeats of SSRs in chloroplast Genome of 13 *Lepidium* species.

Organisms	Mono	Di	Tri	Tetra	Penta	Hexa	Hepta	Octa
Lepidium apetalum	3	6	8	6	6	7	4	1
Lepidium appelianum	3	6	8	7	6	6	2	1
Lepidium cartilagineum	4	4	5	7	4	8	3	1
Lepidium chalepense	2	7	7	8	7	6	3	1
Lepidium cordatum	3	7	7	4	10	10	4	0
Lepidium draba	4	6	7	9	6	10	3	0
Lepidium echinatum	2	7	3	7	7	4	4	1
Lepidium ferganense	3	9	6	4	7	8	3	1
Lepidium latifolium	4	4	5	7	4	8	3	1
Lepidium meyenii	5	5	7	10	6	7	3	1
Lepidium ruderale	3	9	7	7	6	6	3	1
Lepidium sativum	6	7	9	5	6	5	2	1
Lepidium virginicum	3	7	9	8	8	5	1	1
TOTAL	45	84	88	89	83	90	38	11

Table 4. Relative Density of SSR Repeats in the chloroplast genome of 13 *Lepidium* species.

Organisms	Mono	Di	Tri	Tetra	Penta	Hexa	Hepta	Octa
Lepidium apetalum	72	434	375	309	256	388	165	53
Lepidium appelianum	101	414	349	340	266	379	100	52
Lepidium cartilagineum	99	317	182	377	164	466	173	57
Lepidium chalepense	47	478	273	400	325	349	136	52
Lepidium cordatum	70	433	348	169	426	473	259	0
Lepidium draba	112	427	331	471	228	550	182	0
Lepidium echinatum	47	503	164	294	317	208	200	64
Lepidium ferganense	68	517	209	246	387	406	124	57
Lepidium latifolium	98	317	182	377	164	466	173	57
Lepidium meyenii	142	372	320	424	241	362	173	57
Lepidium ruderale	79	600	309	341	256	304	124	57
Lepidium sativum	138	462	374	251	258	235	77	57
Lepidium virginicum	93	470	407	363	432	286	40	53
Total	1166	5744	3823	4362	3720	4872	1926	616

TGATAT (65 bp), and TATCAA (65-83 bp), with elongated TATCAA in species like *L. appelianum* and *L. cartilagineum*. Mononucleotide (Mono) motifs are dominated by T repeats (22-29 bp), with *L. ruderale* having the shortest (22 bp) and *L. sativum* the longest (29 bp). These conserved motifs are overwhelmingly AT-rich (e.g., TA, TAA, AATA), aligning with chloroplast genome composition. Minor variations in motif lengths, such as TA in *L. appelianum* (148 bp) versus *L. ferganense* (124 bp), suggest localized evolutionary tweaks. The widespread conservation of these SSRs implies critical roles in genome structure and regulation, while subtle differences reflect lineage-specific genomic adjustments under functional constraints.

4. Discussion

The chloroplast genomes of *Lepidium* species exhibit a conserved structural framework, with genome sizes (153,177-154,997 bp) and GC content (36.35-36.51%) showing minimal variation, consistent with patterns observed in other angiosperms (Dobrogojski et al., 2020; Wicke et al., 2011; Zhou et al., 2022) (Table 1). This stability arises from the strong functional constraint on the genes crucial for photosynthesis and genome maintenance. (Mower and Vickrey, 2018; Ruhlman and Jansen, 2014). Nonetheless, the complexity of a variety of repeating elements that comprise simple repeats, low-complexity

Table 5. Longest SSR Motifs in the chloroplast genomes of 13 *Lepidium* species.

Organism	Mono	Di	Tri	Tetra	Penta	Hexa
Lepidium apetalum	T (24)	TA (138)	TAT (100)	ATTT (82)	CTATT (59)	TTTATA (75)
	A (24)	AT (78)	TTT (66)	TTTA (66)	ATTTA (52)	TGATAT(65)
	A (24)	TT (66)	TAA (43)	AATA (66)	TCCAT (41)	TTTTTA (65)
Lepidium appelianum	T (50)	TA (148)	TAT (76)	TTTA (66)	TATAT (67)	ATTAGA(74)
	T (27)	TT (69)	TTT (64)	AATA (66)	AATAA (56)	TTAATA(72)
	T (24)	TA (66)	TAA (43)	TTTA (60)	TCCAT (41)	TGATAT(65)
Lepidium cartilagineum	A (27)	TA (144)	TAA (43)	AATA (78)	TCTTT (49)	TGATAT(83)
	T (25)	TT (65)	ATA (38)	TTCT (69)	TCCAT (41)	TATCAA(83)
	T (24)	TT (64)	ATT (34)	TATA (67)	TATCA (37)	TTTTTA (64)
Lepidium chalepense	T (28)	TA (142)	TTT (64)	TTTA (66)	TATAT (67)	ATTAGA(74)
	T (19)	AT (72)	TAA (43)	AATA (66)	AATAA (56)	TTAATA(72)
	-	TT (68)	ATT (34)	TTTA (60)	TTATT (49)	TGATAT(65)
Lepidium cordatum	T (26)	TA (138)	TAT (103)	AATA (74)	CTATT (59)	TGATAT(65)
	A (22)	AT (76)	TTT (64)	ATAA (38)	AATAA (52)	TTTTTA (65)
	T (22)	TT (69)	TAA (43)	TTCA (30)	TTCTA (48)	TATCAA(65)
Lepidium draba	T (34)	TA (145)	TAT (79)	AATA (77)	ATTCT (59)	ATTAGA(74)
	T (29)	AT (80)	TTT (64)	TTTA (66)	TCCAT (41)	TTAATA(72)
	A (25)	TT (70)	ATT (46)	AATA (66)	ATTCT (37)	TGATAT(65)
Lepidium echinatum	T (25)	TA (128)	ATT (99)	TTTA (66)	CTATT (59)	TGATAT(65)
	A (22)	TA (94)	ATT (34)	AATA (66)	TATAT (57)	TATCAA(65)
	-	AT (72)	TAA (31)	TTTC (52)	ATTTT (54)	TTTTAA (39)
Lepidium ferganense	T (27)	TA (124)	TAA (43)	AATA (91)	ATTAA (121)	TTTTTA (71)
	A (24)	AT (66)	ATA (38)	TTTA (66)	CTATT (59)	TGATAT(65)
	A (17)	TT (66)	ATT (34)	AATA (66)	ATTTA (52)	TATCAA(65)
Lepidium latifolium	A (26)	TA (144)	TAA (43)	AATA (78)	TCTTT (49)	TGATAT(83)
	T (25)	TT (65)	ATA (38)	TTCT (69)	TCCAT (41)	TATCAA(83)
	T (24)	TT (64)	ATT (34)	TATA (67)	TATCA (37)	TTTTTA (64)
Lepidium meyenii	A (35)	TA (136)	TAT (106)	AATA (70)	CTATT (59)	TGATAT(65)
	T (31)	TA (68)	TAA (43)	TTTA (66)	TCCAT (41)	TATCAA(65)
	T (28)	TT (67)	ATA (38)	AATA (66)	AAGTA (37)	TTTTTA (64)
Lepidium ruderale	A (35)	TA (124)	TAT (100)	AATA (70)	CTATT (59)	TGATAT(65)
	T (22)	TA (107)	TAA (43)	TTTA (66)	ATTTA (52)	TTTTTA (65)
	A (22)	AT (67)	ATA (38)	AATA (66)	TCCAT (41)	TATCAA(65)
Lepidium sativum	T (29)	TA (145)	TAT (67)	TTTA (66)	TTATT (61)	TGATAT(65)
	A (25)	TA (92)	TTT (64)	AATA (66)	ATTTA (52)	TATCAA(65)
	A (24)	TT (68)	TAA (43)	TTTA (62)	TCCAT (41)	ATACAA(39)
Lepidium virginicum	T (32)	TA (144)	TAT (100)	TTTA (66)	TATTA (112)	TATATT (74)
	A (32)	AT (84)	TTT (64)	AATA (66)	AAGTA (67)	TGATAT(65)
	T (29)	TT (66)	TAA (43)	TATA (57)	AATAA (56)	TATCAA(65)

regions, and SINEs indicate that these regions have a history of conservation and divergence (Table 2, Figure 1). The near-uniform presence of two SINEs (126-127 bp) across the chloroplast genome of all species, except for a longer variant in *L. sativum* (134 bp) (Figure 2), suggests a strong selection on these non-coding elements, possibly for roles in genome organization or regulation. Conversely, the more variable simple repeat (35-45) and low-complexity region

Table 6. Conserved SSR Motifs in all chloroplast genomes of 13 *Lepidium* species.

Organism	Mono	Di	Tri	Tetra	Penta	Hexa
Lepidium apetalum	T (24)	TA (138) TT (66)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium appelianum	T (24)	TA (148) TT (69)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (83)
Lepidium cartilagineum	T (24)	TA (144) TT (65)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (83) TATCAA (83)
Lepidium chalepense	T (28)	TA (142) TT (68)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (83)
Lepidium cordatum	T (26)	TA (138) TT (69)	TAA (43)	AATA (74)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium draba	T (29)	TA (145) TT (70)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (83)
Lepidium echinatum	T (25)	TA (128) TT (60)	TAA (31)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium ferganense	T (27)	TA (124) TT (66)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium latifolium	T (24)	TA (144) TT (65)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (83) TATCAA (83)
Lepidium meyenii	T (28)	TA (136) TT (67)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium ruderale	T (22)	TA (124) TT (66)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium sativum	T (29)	TA (145) TT (68)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)
Lepidium virginicum	T (29)	TA (144) TT (66)	TAA (43)	AATA (66)	TCCAT (41)	ATACAA (39) TGATAT (65) TATCAA (65)

(8-14) frequencies indicate plasticity in non-essential regions, which might be driven by mutation accumulation or adaptive fine-tuning (Ellegren, 2004; Provan et al., 2001). A region where small RNA is located masking 3.73-4.18% of the genomes again highlights specific regulatory requirements for individual species, as repetitive elements in chloroplasts maintain stability and flexibility (Smith and Keeling, 2015).

SSR distribution in *Lepidium* chloroplasts demonstrates a clear bias toward mid-length repeats (Hexa, Tetra, Tri), likely due to replication slippage favoring these motifs, while shorter (Mono) and longer (Hepta, Octa) repeats are rarer, as observed in other plant genomes (Wheeler et al., 2014) (Table 3). AT-rich SSRs such as TA, TAA, and AATA prevail, and this mirrors the general AT, as opposed to GC, bias of chloroplast DNA and may help with DNA bending or exclusion of nucleosomes from regulatory regions (Provan et al., 2001; Mower and Vickrey, 2018). Species-specific deviations, like elevated Hexa/Penta counts in *L. cordatum* and *L. draba* or the complete absence of Octa repeats in these species, suggest lineage-specific evolutionary pressures. For example, the absence of Octa repeats may be reflective of purifying selection against large indels in compact chloroplast genomes, in which even slight perturbations can impact critical functions (Wicke et al., 2011; Smith and Keeling, 2015). It indicates that mutation biases and selection jointly shape SSR landscapes, permitting adaptation while maintaining essential cellular functions (Provan et al., 2001).

Conserved SSR motifs, such as TCCAT (Penta) and TATCAA (Hexa), across chloroplast genomes of all 13 *Lepidium* species point to critical roles in genome stability or regulation (Ebert and Peakall, 2009) (Table 6). The tight conservation of TCCAT (41 bp) and the extension of TATCAA (65-83 bp) in select taxa indicate that these elements may serve as originating points for replications or bind to nuclear-encoded proteases (Provan et al., 2001). Minor variations, like TA length differences (*L. appelianum*: 148 bp vs. *L. ferganense*: 124 bp), hint at localized evolutionary adjustments, possibly driven by environmental stressors or drift. The sheer abundance of AT-rich motifs coincides with chloroplast genome architecture, as AT-rich regions correlate with replication efficiency and transcriptional activity (Daniell et al., 2016; Zhou et al., 2017). Unique features, such as the exceptionally long ATTAA (121 bp) in *L. ferganense*, may represent adaptive expansions influencing gene expression under specific ecological conditions. The results suggest that SSRs represent a conserved and variable strategy for chloroplast genomes to maintain structural integrity through conserved motifs whilst hosting lineage-specific innovations as part of a dynamic process. This process is critical for balancing genomic stability against the adaptive potential that is required for diverse environments (Daniell et al., 2016; Smith and Keeling, 2015; Wicke et al., 2011).

The observed lineage-specific variations in SSR motifs—such as the absence of octanucleotide repeats in *L. cordatum* and *L. draba* or the expansion of pentanucleotide ATTAA in *L. ferganense*—underscore the interplay between phylogenetic divergence and potential ecological adaptation. While conserved SSRs (e.g., TCCAT, TATCAA) likely maintain structural and regulatory stability across *Lepidium* (Provan et al., 2001; Ebert and Peakall, 2009), lineage-specific repeats may reflect divergent evolutionary pressures tied to habitat-specific stressors, such as aridity or soil salinity, which are known to influence chloroplast genome plasticity in arid-adapted Brassicaceae (Zhou et al., 2022; Mower and Vickrey, 2018). However, explicit functional or environmental data linking these SSRs to adaptive traits (e.g., stress tolerance) remain limited in this study. For instance, while AT-rich motifs align with chloroplast genome architecture and replication efficiency (Daniell et al., 2016), their role in ecological resilience—such as buffering against environmental fluctuations—requires experimental validation through comparative fitness assays or gene expression analyses under controlled stress conditions. Phylogenetic comparative methods (e.g., Phylogenetic Generalized Least Squares PGLS) could further clarify whether SSR divergence correlates with historical climate shifts or biogeographic patterns in *Lepidium*.

Thus, while our findings suggest that SSRs provide a genomic substrate for balancing conservation and innovation, claims of adaptive plasticity should be tempered until mechanistic links to ecological performance are established (Wicke et al., 2011; Smith and Keeling, 2015). Future studies integrating environmental gradients, phenotypic data, and population-level SSR dynamics will be critical to disentangling neutral drift from adaptive evolution in chloroplast genome evolution.

5. Conclusion

Our comparative analysis of the chloroplast genomes of 13 *Lepidium* species reveals a remarkable balance between evolutionary conservation and adaptive diversification. Despite a highly conserved genomic architecture—evidenced by minimal variation in genome size and GC content—the distribution and configuration of SSRs display both ubiquitous motifs and distinct, species-specific patterns. AT-rich, intermediate-length repeats are prevalent and seem to play a role in the maintenance of total chloroplast genome stability and regulation. while the absence of octanucleotide repeats in some species and unique repeat expansions in some other species, indicate to lineage-specific rather than universal selection pressures. These insights enhance our understanding of how conserved genomic elements can accommodate local innovations, ultimately contributing to the ecological resilience and evolutionary success of *Lepidium* species.

Data Availability Statement

The entire data set that supports the results of this study was published in the article itself.

References

ALHAWATEMA, M., 2023. DendroSSR: SSRs and sequence alignment as tools for building phylogeny trees. *Brazilian Journal of Biology = Revista Brasileira de Biologia*, vol. 83, e275386. http://doi.org/10.1590/1519-6984.275386.

BONA, M., 2020. Numerical taxonomic analysis on some *Lepidium* L. taxa (Brassicaceae) from Turkey. *European Journal of Biology*, vol. 79, no. 2, pp. 132-143. http://doi.org/10.26650/EurJBiol.2020.0020.

CHATOUI, K., TALBAOUI, A., ANEB, M., BAKRI, Y., HARHAR, H. and TABYAOUI, M., 2016. Phytochemical screening, antioxidant and antibacterial activity of *Lepidium sativum* seeds from

Morocco. *Journal of Materials Environmental Science*, vol. 7, no. 8, pp. 2938-2946.

DANIELL, H., LIN, C.S., YU, M. and CHANG, W.J., 2016. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biology*, vol. 17, no. 1, pp. 134. http://doi.org/10.1186/s13059-016-1004-2. PMid:27339192.

DOBROGOJSKI, J., ADAMIEC, M. and LUCIŃSKI, R., 2020. The chloroplast genome: a review. *Acta Physiologiae Plantarum*, vol. 42, no. 6, pp. 98. http://doi.org/10.1007/s11738-020-03089-x.

EBERT, D. and PEAKALL, R., 2009. Chloroplast simple sequence repeats (cpSSRs): technical resources and recommendations for expanding cpSSR discovery and applications to a wide array of plant species. *Molecular Ecology Resources*, vol. 9, no. 3, pp. 673-690. http://doi.org/10.1111/j.1755-0998.2008.02319.x. PMid:21564725.

ELLEGREN, H., 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews. Genetics*, vol. 5, no. 6, pp. 435-445. http://doi.org/10.1038/nrg1348. PMid:15153996.

GERMAN, D.A., 2014. Taxonomic remarks on some Asian *Lepidium* s. l. (Brassicaceae). *Phytotaxa*, vol. 186, no. 2, pp. 97-105. http://doi.org/10.11646/phytotaxa.186.2.4.

GHANTE, M.H., BADOLE, S.L. and BODHANKAR, S.L., 2011. Health benefits of garden cress (*Lepidium sativum* Linn.) seed extracts. In: V.R. PREEDY, R.R. WATSON and V. B. PATEL, eds. Nuts and seeds in health and disease prevention. London: Academic Press, pp. 521-525. http://doi.org/10.1016/B978-0-12-375688-6.10062-3.

GIRMAY, S., DEKOBO, A., BEREHE, S.G. and DEKOBO BORU, A., 2014. Phytochemical screening and antimicrobial activities of crude extract of *Lepidium sativum* seeds grown in Ethiopia. *International Journal of Pharmaceutical Sciences and Research*, vol. 5, no. 10, pp. 4182-4187.

GONZALES, G.F., 2012. Ethnobiology and ethnopharmacology of *Lepidium meyenii* (Maca), a plant from the Peruvian highlands. *Evidence-Based Complementary and Alternative Medicine*, vol. 2012, pp. 193496. http://doi.org/10.1155/2012/193496. PMid:21977053.

HASHMI, H.F., BIBI, S., ANWAR, M. and RASHID KHAN, M., 2021. Qualitative and quantitative analysis of phytochemicals in *Lepidium* pinnatifidum Ledeb. *Scholars International Journal of Traditional and Complementary Medicine*, vol. 4, no. 5, pp. 67-75.

KALIA, R.K., RAI, M.K., KALIA, S., SINGH, R. and DHAWAN, A.K., 2011. Microsatellite markers: an overview of the recent progress in plants. *Euphytica*, vol. 177, no. 3, pp. 309-334. http://doi.org/10.1007/s10681-010-0286-9.

KARAOGLU, H., LEE, C.M.Y. and MEYER, W., 2005. Survey of simple sequence repeats in completed fungal genomes. *Molecular Biology and Evolution*, vol. 22, no. 3, pp. 639-649. http://doi.org/10.1093/molbev/msi057. PMid:15563717.

LEE, M.S., SHIN, B.C., YANG, E.J., LIM, H.J. and ERNST, E., 2011. Maca (*Lepidium meyenii*) for treatment of menopausal symptoms: a systematic review. *Maturitas*, vol. 70, no. 3, pp. 227-233. http://doi.org/10.1016/j.maturitas.2011.07.017. PMid:21840656.

LI, Y.C., KOROL, A.B., FAHIMA, T., BEILES, A. and NEVO, E., 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, vol. 11, no. 12, pp. 2453-2465. http://doi.org/10.1046/j.1365-294X.2002.01643.x. PMid:12453231.

MOWER, J.P. and VICKREY, T.L., 2018. Structural diversity among plastid genomes of land plants. *Advances in Botanical Research*, vol. 85, pp. 263-292. http://doi.org/10.1016/bs.abr.2017.11.013.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION – NCBI, 2025 [viewed 12 February 2025]. GenBank [online]. Available from: http://www.ncbi.nlm.nih.gov

ORTH, J.F., GAMMON, M., ABDUL-BASIR, F., STEVENSON, R.D., TSIRELSON, D., EBERSOLE, J., SPEAK, S. and KESSELI, R., 2006. Natural history, distribution, and management of *Lepidium* latifolium (Brassicaceae) in New England. *Rhodora*, vol. 108, no. 934, pp. 103-118. http://doi.org/10.3119/05-12.1.

PROVAN, J., POWELL, W. and HOLLINGSWORTH, P.M., 2001. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution*, vol. 16, no. 3, pp. 142-147. http://doi.org/10.1016/S0169-5347(00)02097-8. PMid:11179578.

RUHLMAN, T.A. and JANSEN, R.K., 2014. The plastid genomes of flowering plants. *Methods in Molecular Biology (Clifton, N.J.)*, vol. 1132, pp. 3-38. http://doi.org/10.1007/978-1-62703-995-6_1. PMid:24599844.

RUIZ-LUNA, A.C., SALAZAR, S., ASPAJO, N.J., RUBIO, J., GASCO, M. and GONZALES, G.F., 2005. *Lepidium meyenii* (Maca) increases litter size in normal adult female mice. *Reproductive Biology and Endocrinology*, vol. 3, no. 1, pp. 16. http://doi.org/10.1186/1477-7827-3-16. PMid:15869705.

SAID, M. and KASSAHUN, T., 2015. Molecular genetic diversity study of *Lepidium sativum* population from Ethiopia as revealed by inter simple sequence repeat (ISSR) markers. *African Journal of Biotechnology*, vol. 14, no. 17, pp. 1461-1470. http://doi.org/10.5897/AJB2015.14533.

SHABBIR, F., EDDOUKS, M., NADEEM, F. and AZEEM, M.W., 2018. A brief review on bioactivities and therapeutic potentials of garden cress (*Lepidium sativum* L.). *International Journal of Chemical and Biochemical Science*, vol. 13, pp. 36-45.

SHAH, M.B., DUDHAT, V.A. and GADHVI, K.V., 2021. *Lepidium sativum*: a potential functional food. *Journal of Ayurvedic and Herbal Medicine*, vol. 7, no. 2, pp. 140-149. http://doi.org/10.31254/jahm.2021.7213.

SMIT, S.A., HUBLEY, R. and GREEN, P., 2015 [viewed 12 February 2025]. RepeatMasker Open-4.0. Version 4.1.1 [online]. Available from: http://www.repeatmasker.org

SMITH, D.R. and KEELING, P.J., 2015. Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 33, pp. 10177-10184. http://doi.org/10.1073/pnas.1422049112. PMid:25814499.

WHEELER, G.L., DORMAN, H.E., BUCHANAN, A., CHALLAGUNDLA, L. and WALLACE, L.E., 2014. A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. *Applications in Plant Sciences*, vol. 2, no. 12, pp. 1400059. http://doi.org/10.3732/apps.1400059. PMid:25506520.

WICKE, S., SCHNEEWEISS, G.M., DEPAMPHILIS, C.W., MÜLLER, K.F. and QUANDT, D., 2011. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology*, vol. 76, no. 3-5, pp. 273-297. http://doi.org/10.1007/s11103-011-9762-4. PMid:21424877.

ZHOU, Q., CHEN, Y., WANG, J., LI, M., ZENG, W., WANG, Y., LI, Y. and ZHAO, H., 2022. A comparative study of the chloroplast genomes of five *Lepidium* species with high medicinal value. *Functional Plant Biology*, vol. 50, no. 1, pp. 29-45. http://doi.org/10.1071/FP22052. PMid:36043226.

ZHOU, T., YANG, Y., HU, Y., ZHANG, X., BAI, G. and ZHAO, G., 2017. Characterization of the complete chloroplast genome sequence of *Lepidium meyenii* (Brassicaceae). *Conservation Genetics Resources*, vol. 9, no. 3, pp. 405-408. http://doi.org/10.1007/s12686-017-0695-3.