scientific reports

Check for updates

OPEN

Lack of ITS sequence homogenization in *Erysimum* species (Brassicaceae) with different ploidy levels

Carolina Osuna-Mascaró[1,2,5]✉, Rafael Rubio de Casas[2,3], Modesto Berbel[1], José M. Gómez[2,4] & Francisco Perfectti[1,2]✉

The internal transcribed spacers (ITS) exhibit concerted evolution by the fast homogenization of these sequences at the intragenomic level. However, the rate and extension of this process are unclear and might be conditioned by the number and divergence of the different ITS copies. In some cases, such as hybrid species and polyploids, ITS sequence homogenization appears incomplete, resulting in multiple haplotypes within the same organism. Here, we studied the dynamics of concerted evolution in 85 individuals of seven plant species of the genus *Erysimum* (Brassicaceae) with multiple ploidy levels. We estimated the rate of concerted evolution and the degree of sequence homogenization separately for ITS1 and ITS2 and whether these varied with ploidy. Our results showed incomplete sequence homogenization, especially for polyploid samples, indicating a lack of concerted evolution in these taxa. Homogenization was usually higher in ITS2 than in ITS1, suggesting that concerted evolution operates more efficiently on the former. Furthermore, the hybrid origin of several species appears to contribute to the maintenance of high haplotype diversity, regardless of the level of ploidy. These findings indicate that sequence homogenization of ITS is a dynamic and complex process that might result in varying intra- and inter-genomic diversity levels.

Concerted evolution is an evolutionary process by which sequences from the same gene family show higher sequence similarity to each other than to orthologous genes in related species[1,2]. Hence, genes evolved in a concerted manner present low polymorphism in their sequences, i.e., the sequences are homogenized. Concerted evolution is particularly notable in multicopy nuclear genes, where homogenization is mainly achieved by unequal crossing over and gene conversion[3,4]. One of the best characterized multicopy gene families is the 45S nuclear ribosomal DNA (nrDNA)[5]. It appears arranged as tandem repeated units with hundreds to thousands of copies in one or several loci per genome. These units are composed of the 18S rDNA, internal transcribed spacer 1 (ITS1), 5.8 S rDNA, internal transcribed spacer 2 (ITS2), and 26S rDNA, separated by longer non-transcribed intergenic spacers[6]. Among all these units, the internal transcribed spacers (ITS1 and ITS2) are the best-characterized nrDNA sequences[7] partly because ITS sequences show characteristics advantageous for phylogenetic studies, such as biparental inheritance, short length, and high evolution rate[4,8,9].

ITS sequences usually present fast concerted evolution with low levels of intra-genomic sequence variation and very few polymorphic positions[10,11]. However, in some animals (e.g.,[12]) and especially in plants[13–18], sequence homogenization remains incomplete across ITS sequences, resulting in relatively high intra-genomic polymorphism. This ITS diversity is often linked to hybridization events[9,19–22]. Different ITS sequences may meet after hybridization and become homogenized after a time, but this homogenization may not be consistent among descendant lineages[23]. As concerted evolution tends to homogenize sequences rapidly[8], evidence of non-concerted evolution is mostly expected in recently-formed hybrid species, where both parental ITS sequences may still be present. This phenomenon should be particularly conspicuous in recent allopolyploid species, where the occurrence of different ITS sequences located in distinct chromosomes tends to delay this homogenization[24].

[1]Departamento de Genética, Universidad de Granada, Granada, Spain. [2]Research Unit Modeling Nature, Universidad de Granada, Granada, Spain. [3]Departamento de Ecología, Universidad de Granada, Granada, Spain. [4]Departamento de Ecología Funcional y Evolutiva, Estación Experimental de Zonas Áridas (EEZA-CSIC), Almería, Spain. [5]Present address: Department of Biology, University of Nevada-Reno, Reno, NV, USA. ✉email: cosuna@unr.edu; fperfect@ugr.es

Taxon	Population code	Location	Elevation (m.a.s.l.)	Geographical coordinates	Ploidy level
E. baeticum	Ebb07	Sierra Nevada, Almería, Spain	2128	37°05′46″ N, 3°01′01″ W	8x
	Ebb10	Sierra Nevada, Almería, Spain	2140	37°05′32″ N, 3°00′40″ W	8x
	Ebb12	Sierra Nevada, Almería, Spain	2264	37°05′51″ N, 2°58′06″ W	8x
E. bastetanum	Ebt01	Sierra de Baza, Granada, Spain	1990	37°22′52″ N, 2°51′49″ W	4x
	Ebt12	Sierra de María, Almería, Spain	1528	37°41′03″ N, 2°10′51″ W	4x
	Ebt13	Sierra Jureña, Granada, Spain	1352	37°57′10″ N, 2°29′24″ W	8x
E. fitzii	Ef01	Sierra de la Pandera, Jaén, Spain	1804	37°37′56″ N, 3°46′46″ W	2x
E. lagascae	Ela07	Sierra de San Vicente, Toledo, Spain	516	44°05′49″ N, 4°40′40″ W	2x
E. mediohispanicum	Em21	Sierra Nevada, Granada, Spain	1723	37°08′04″ N, 3°25′43″ W	2x
	Em39	Sierra de Huétor, Granada, Spain	1272	37°19′08″ N, 3°33′11″ W	2x
	Em71	Sierra Jureña, Granada, Spain	1352	37°57′10″ N, 2°29′24″ W	4x
E. nevadense	En05	Sierra Nevada, Granada, Spain	2074	37°06′35″ N, 3°01′32″ W	2x
	En10	Sierra Nevada, Granada, Spain	2321	37°06′37″ N, 3°24′18″ W	2x
	En12	Sierra Nevada, Granada, Spain	2255	37°05′37″ N, 2°56′19″ W	2x
E. popovii	Ep16	Jabalcuz, Jaén, Spain	796	37°45′26″ N, 3°51′02″ W	4x
	Ep20	Sierra de Huétor, Granada, Spain	1272	37°19′08″ N, 3°33′11″ W	10x
	Ep27	Llanos del Purche, Granada, Spain	1470	37°07′46″ N, 3°28′48″ W	4x

Table 1. Taxonomic assignment, population code, location, elevation, and ploidy level for the *Erysimum* spp. populations sampled.

Erysimum l. (Brassicaceae) comprises more than 200 species[25], mainly from Eurasia, with some species inhabiting North America and North Africa[26,27]. The Baetic Mountains (SE Iberia) constitute one of the most important glacial refugia in Europe and a hotspot for this group, with ~ 10 *Erysimum* species occurring in this small area[28,29]. In particular, these *Erysimum* species show characteristics that may facilitate hybridization and inter-specific gene-flow, such as occasional sympatry and a generalist pollination system[32]. Thus, previous studies have suggested that several of these taxa could have a hybrid origin[30–32]. Ploidy levels vary among and, in some cases, within species[28,33], suggesting that a detailed understating of hybridization and allopolyploidization is necessary to shed light on the evolution of this group. However, the effects of hybridization and polyploidization on the genomes of these species are far from being fully understood.

In this study, we explore the homogenization dynamics of ITS, taking into account the interacting effects of concerted evolution, hybridization, and polyploidization. For this purpose, we analyzed polymorphisms at the species, population, and individual levels in ITS1 and ITS2 for seven *Erysimum* species. These species are closely related and belong to an Iberian clade within this genus[34]. We sequenced ITS1 and ITS2 by NGS to recover all the ITS copies present in the different genomes[11,35]. With these sequences, we then proceeded to quantify the degree of sequence homogenization in both ITS1 and ITS2 within individuals, populations, and species; and the concerted evolution levels in *Erysimum* spp. These species have been previously studied, showing a mainly outcrossing mating system with weak prezygotic barriers among them[32,33]. A cpDNA phylogeny has shown a recent origin (< 2 Mya) for these species[36]. Moreover, other phylogenies for these same species have shown reticulated patterns with a lack of species clustering in some cases and evidence of cytonuclear discordance, suggesting a recent hybridization scenario with allopolyploidization[33]. Due to their recent evolution, we hypothesize that polyploid species of this genus will have less ITS homogenization than diploid species. Any insight into ITS evolution in plants needs to consider the concomitant effects of hybridization and polyploidization on the rates of concerted evolution.

Materials and methods

Taxon sampling. We collected fresh leaves from polyploid and diploid *Erysimum* species. To determine DNA ploidy levels and assess genome size for each population, we used flow cytometry (see Table 1 for details on species ploidy levels). Specific details on the flow cytometry analyses could be found in Osuna-Mascaró et al.[32,33].

In particular, we sampled leaves from five individuals belonging to three different populations of *Erysimum baeticum*, *E. bastetanum*, *E. mediohispanicum*, *E. nevadense*, and *E. popovii*, and five individuals from one population of *E. lagascae* and five from the microendemic *E. fitzii*. A total of 85 samples (= leaves of each individual) were dried and preserved in silica gel until DNA extraction. Table 1 shows the code, location, and ploidy levels of all samples.

DNA extraction. We used at least 60 mg of dry plant material from each sample. We disrupted the tissues in liquid N2 using a mortar and pestle. Then, total genomic DNA was isolated using the GenElute Plant Genomic DNA Miniprep kit (Sigma-Aldrich, St. Louis, MO) following the manufacturer's protocol [https://www.sigmaaldrich.com/ES/es/technical-documents/protocol/genomics/dna-and-rna-purification/genelute-plant-genomic-dna-purification-kit]. The quantity and quality of the obtained DNA were checked using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, United States), and the integrity of the extracted DNA was checked on agarose gel electrophoresis.

ITS1 and ITS2 amplification. We independently amplified ITS1 and ITS2 in each sample. The ITS PCR reactions were performed in 25 µl with the following composition: 5 µL 5 × buffer containing MgCl2 at 1.5 mM (New England Biolabs), 0.1 mM each dNTP, 0.2 µM each primer, and 0.02 U Taq high fidelity DNA-polymerase (Q5 High-Fidelity DNA Polymerase, New England Biolabs). We used a set of long primers developed to have a 5' flanking sequence complementary to the Nextera XT DNA index to facilitate adapter ligation during library construction:

>ITS1-Flabel (for ITS1 amplification)
TCG TCG GCA GCG TCA GAT GT GTA TAA GAG ACA GTC CGT AGG TGA ACC TGC GG
>ITS1-Rlabel (for ITS1 amplification)
GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGC TGC GTT CTT CAT CGA TGC
>ITS3-Flabel (for ITS2 amplification)
TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GCA TCG ATG AAG AAC GCA GC
>ITS4-Rlabel (for ITS2 amplification)
GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTC CTC CGC TTA TTG ATA TGC

Reactions included 30 cycles with the following conditions: 94 °C 15 s, 60 °C 30 s, and 72 °C 30 s. Amplified fragments were purified using spin columns (GenElute TM PCR Clean-Up Kit, Sigma-Aldrich) and were checked on agarose gel electrophoresis. Finally, we quantified the starting DNA concentration using the Infinite M200 PRO NanoQuant spectrophotometer (TECAN, Männedorf, Switzerland).

Library construction. We constructed two libraries, one for ITS1 amplicons and one for ITS2 amplicons. The libraries were prepared using the Nextera XT DNA Sample Preparation Kit. In brief, the DNA was tagged by adding a unique adapter label combination to the 3' and 5' ends of the DNA sequence. Then, the DNA was amplified via a nine-cycle PCR. The total volume reaction was 25 µl with the following composition: 5 µL 10 × buffer at 1.0 mM (New England BioLabs), 0.1 mM each dNTP, 0.2 µM each Nextera primer, 0.02 U Taq high fidelity DNA-polymerase (Q5, NEB), and 5 × Q5 High GC Enhancer (NEB). PCR thermocycling conditions were 98 °C during 5 s, 55 °C for 10 s, and 72 °C for 10 s. After that, we purified both libraries using the GenElute PCR Clean-Up Kit (Sigma) to remove short library fragments. Finally, we generated equal volumes of the libraries to prepare equimolar libraries for sequencing, and the final concentration of each library was quantified using the Infinite M200 PRO NanoQuant spectrophotometer (TECAN, Männedorf, Switzerland).

Library sequencing. ITS1 and ITS2 library sequencing were carried out by Novogene Bioinformatics Technology Co., Ltd, with an Illumina MiSeq platform (Illumina, USA) using a paired-end 150 bp sequence read run. The ITS libraries of *E. mediohispanicum* were sequenced twice due to an unexpected low sequencing output (we constructed new libraries as explained above). This sequencing was done using the Illumina Miseq platform and paired-end chemistry in the Center for Scientific Instrumentation (CIC) of the University of Granada, Spain.

Data analysis. FASTQ files were demultiplexed, and read quality was checked in FastQC v0.11.5[37]. Then, we did a trimming procedure using first cutadapt v1.15[38] to trim the adapters, followed by a quality trimming using Sickle v1.33[39]. Forward and reverse reads were paired in Geneious R.11[40]. Using the function "Set pair read" with default parameters for Illumina paired-end read technology. Then the paired reads were merged using BBMerge v37.64[41] with a "Low Merge rate" to decrease false positives. Then, to reduce redundancy and noise caused by sequencing errors and tag switching events, we did a cluster analysis using CD-HIT v4.6.8[42]. We clustered the sequences from each sample using an identity threshold of 0.99 (i.e., we merged the sequences with similarity ≥ 99%) and discarded the clusters that included < 5% of the total reads[43]. This step reduced the contribution of sequencing errors to the reported sequence diversity.

We aligned the sequences from each sample using MAFFT v7.450[44] with default parameters, generating one alignment per species and marker. We trimmed the alignments using trimAl v1.2[45], removing poorly aligned regions with the "gappyout" method. We estimated population genetic parameters at intra-species, intra-population, and intra-individual levels using the R package PEGAS v0.1[46]. We used the "nuc.div" function to calculate nucleotide diversity (π), estimated as the average number of nucleotide differences per site between two sequences[47,48]. We constructed boxplots in R to depict the nucleotide diversity (π) of each sample for ITS1 and ITS2 using the package ggplot2[49]. Moreover, we estimated the haplotype diversity (Hd), with the "hap.div" function, as the probability of differentiation between two randomly chosen haplotypes. We then used the "haplotype" function to calculate the total number of haplotypes and the haplotype frequency distribution for each species,

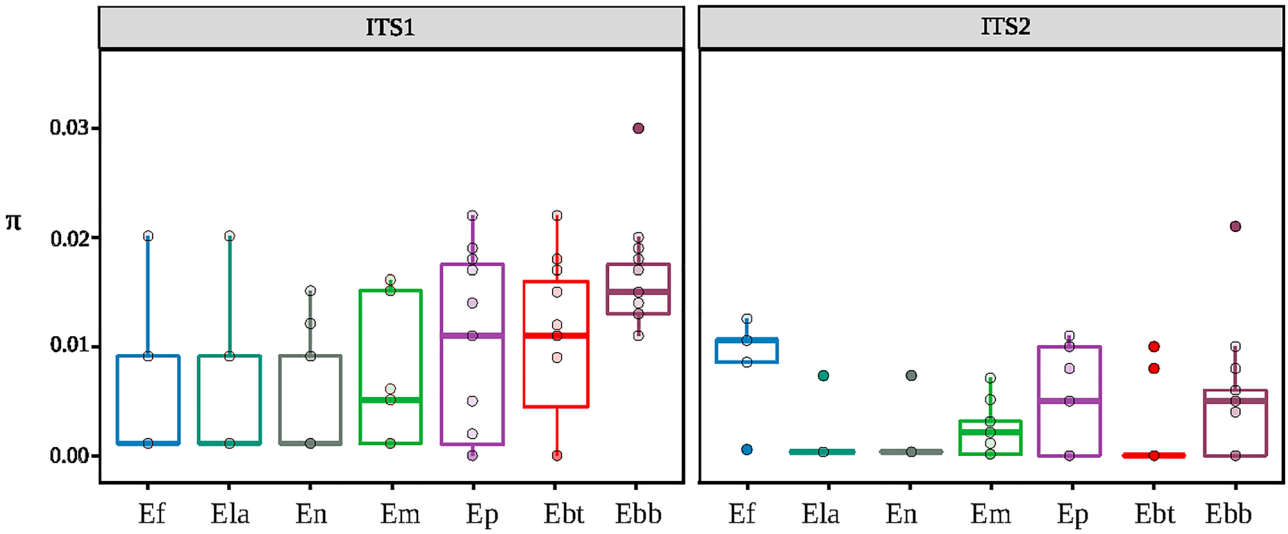

Figure 1. Boxplot depicting the nucleotide diversity (π) for ITS1 and ITS2 samples. Nucleotide polymorphism was estimated for each *Erysimum* individual as the average number of nucleotide differences per site between two sequences (Nei and Li[47]). *E. baeticum* (Ebb, ploidy 8x), *E. bastetanum* (Ebt, ploidy 4x and 8x), *E. popovii* (Ep, ploidy 4x and 10x), and one population of *E. mediohispanicum* (Em, ploidy 4x) are polyploids. *E. nevadense* (En), *E. fitzii* (Ef), two populations of *E. mediohispanicum* (Em), and *E. lagascae* (Ela) are diploids.

population, and individual. We represented the number of ITS1 and ITS2 haplotypes per sample for diploid and polyploid species with a boxplot generated in R using ggplot2[49]. We checked for normality using Shapiro–Wilk's method and then compared the nucleotide and haplotype diversity and the number of haplotypes among polyploid and diploid species and among ITS1 and ITS2 using the Mann–Whitney–Wilcoxon test. All statistical analyses were done in R v 4.1.0 using the package stats v3.6.1[50].

We investigated potential correlations among ploidy levels and haplotype and nucleotide diversity for ITS1 and ITS2 samples. Also, as these species were described as frequently hybridizing, we studied if there were shared haplotypes among different populations of the same species and among different species. To explore that, we estimated the total number of ITS1 and ITS2 haplotypes and their frequencies.

We analyzed the genetic structure of ITS1 and ITS2 by performing a hierarchical analysis of molecular variance (AMOVA;[51]). We used the "amova" function from the R package PEGAS v0.1[46] to explore the genetic variation explained by populations (i.e., at the population level), among individuals within populations (i.e., at the individual level), and within individuals (i.e., at the intra-genome level). We run an AMOVA for each species, including all the population samples, regardless of population ploidy levels. Moreover, we analyzed the amount of genetic variation in ITS1 and ITS2 explained by interspecific differences by partitioning the variance into three levels: among species, among populations within species, and within populations (i.e., among individuals). For that, we run two different AMOVA analyses first, all the sequences of ITS1 and then all the sequences of ITS2, regardless of the species and ploidy level.

Research involving plants. We obtained permission for collecting plant material from: Junta de Andalucía, Consejería de Medioambiente y Ordenación del Territorio. The sampling complied with all institutional, national, and international guidelines and legislations.

Results

From the initial 85 individuals, we obtained good-quality sequences for a total of 84 ITS1 and 81 ITS2 samples, with 10,156 ± 1233 sequences per individual for ITS1 and 49,428 ± 7678 sequences for ITS2 (Table S1).

Polyploid species (*E. baeticum*, *E. bastetanum*, *E. popovii*) tended to have higher nucleotide diversity than diploid species (Fig. 1) for both ITS1 (Wilcoxon test = 655, *p* value: 0.04; mean π ± SE; polyploid: 0.012 ± 0.007, diploid: 0.004 ± 0.006) and ITS2 (Wilcoxon test = 663, *p* value: 0.03; polyploids: 0.003 ± 0.004, diploids: 0.002 ± 0.003). In addition, the polyploid population of *E. mediohispanicum* (Em71, 4x) showed higher nucleotide diversity than the two diploid populations of this species, marginally significant for ITS1 (Wilcoxon test = 10, *p* value: 0.05; Em71: mean π = 0.011 ± 0.006; Em39: mean π = 0.006 ± 0.007 for ITS1; Em21: mean π = 0.0004 ± 0.001; Fig. S1) and more pronounced for ITS2 (Wilcoxon test = 8.5, *p* value: 0.04; Em71: mean π = 0.003 ± 0.002; Em39: mean π = 0.0003 ± 0.001 for ITS2; Em21: mean π = 0.001 ± 0.002 for ITS2; Fig. S1). Furthermore, the correlation between ploidy level and nucleotide diversity was highly significant for ITS1 (Spearman's rho: 0.48, *p* value: 2.10×10^{-6}; Fig. S2) and marginally significant for ITS2 (Spearman's rho: 0.20, *p* value: 0.06). The difference in the degree of association of ITS1 and ITS2 polymorphisms with ploidy levels might be a consequence of overall diversity. ITS2 samples presented significantly lower nucleotide diversities than ITS1 ones (Wilcoxon test = 5165.5, *p* value: 3.33×10^{-7}). Nucleotide diversity values for ITS1 and ITS2 at the three levels of analysis (species, population, individual) are shown in Tables S2–S8.

Species	ITS 1	ITS2
E. baeticum	0.983 (1–0.963)	0.897 (1–0.933)
E. bastetanum	0.983 (1–0.969)	0.893 (1–0.666)
E. fitzii	0.944 (1–0)	0.872 (1–0)
E. lagascae	0.936 (1–0)	0.733 (1–0)
E. mediohispanicum	0.969 (1–0.400)	0.805 (1–0.866)
E. nevadense	0.938 (1–0.785)	0.941 (1–0.833)
E. popovii	0.984 (1–0.888)	0.943 (1–0.866)

Table 2. Average haplotype diversity (Hp) per species, estimated for ITS1 and ITS2 samples. Maximum and minimum values (in parentheses) refer to individual samples.

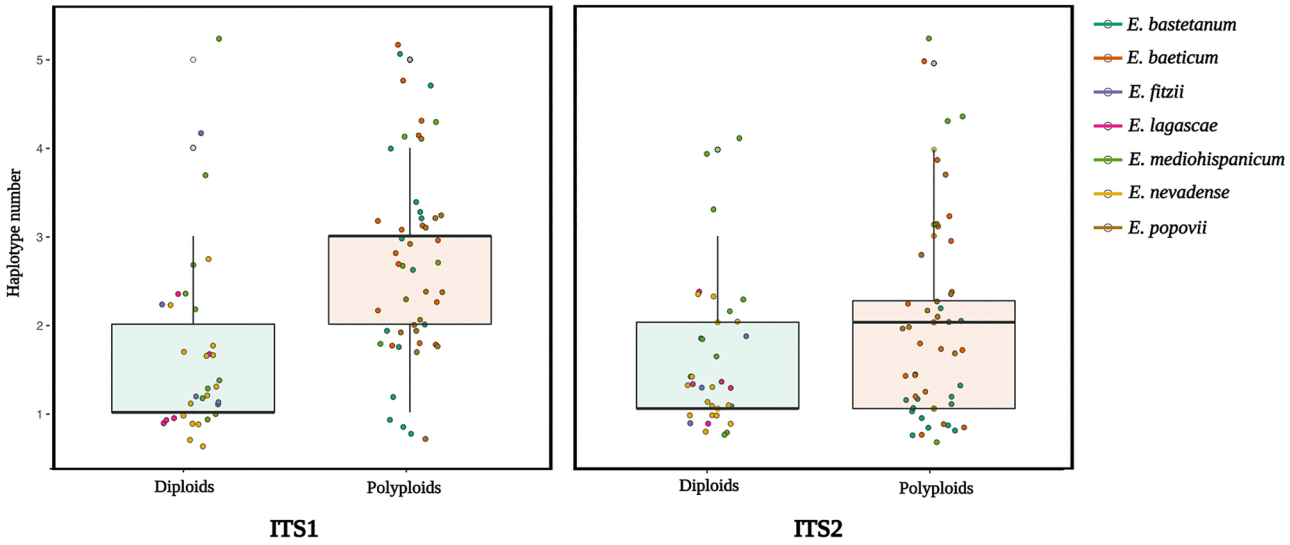

Figure 2. Boxplot depicting the number of ITS1 (left) and ITS2 (right) haplotypes per sample for diploid and polyploid species.

Haplotype diversity showed a similar pattern to that of the nucleotide diversity, with higher haplotype diversity for polyploid species than diploid species, for ITS1 (Wilcoxon test = 343, p value: 2.16×10^{-6}; mean Hd = 0.89 ± 0.38 for polyploid; mean Hd = 0.50 ± 0.49 for diploid) and marginally significant for ITS2 (Wilcoxon test = 632.5, p value = 0.059; mean Hd = 0.39 ± 0.49 for polyploid; mean Hd = 0.28 ± 0.45 for diploid). Moreover, the degree of association between haplotype diversity and ploidy level seemed to differ between ITSs, being highly significant for ITS1 (Spearman's rho: 0.43, p value: 2.96×10^{-5}) but only marginally significant for ITS2 (Spearman's rho: 0.18, p value: 0.09). The values of haplotype diversity for both ITS and three levels are shown in Tables S2–S8. ITS2 presented lower haplotype diversity than ITS1 in terms of haplotype numbers (Wilcoxon test = 4458, p value 0.002; Table 2). ITS2 diversity was reduced to a single haplotype (i.e., no polymorphism was detected) in 49 individuals (Tables S2–S8). Conversely, only 30 individuals showed no nucleotide diversity in ITS1 (Tables S2–S8).

Polyploid species showed higher number of haplotypes than diploid species (Fig. 2). Moreover, several ITS1 haplotypes were shared across species, particularly among some populations of *E. bastetanum*, *E. fitzii*, *E. mediohispanicum*, and *E. nevadense*. Specifically, we found that the three populations of *E. bastetanum* studied in this article shared haplotypes with two *E. mediohispanicum* populations (Em39, Em71) and with the three populations of *E. nevadense*. In addition, *E. bastetanum* populations and one population of *E. nevadense* (En05) shared haplotypes with the *E. fitzii* population included in the analyses. Conversely, no ITS2 haplotypes were found to be shared across different species (Tables S10, S11, S13, and S14).

The hierarchical AMOVA showed that interspecific differences were a significant source of variation for both ITS (Table 3). The species-level explained 52.63 and 73.50% of the variance for ITS1 (p value < 0.001, $\Phi = 0.48$) and ITS2 (p value < 0.001, $\Phi = 0.70$) respectively, implying ample genetic divergence among species. Conversely, differences among populations were not significant and absorbed a relatively low amount of molecular variance

Sequence	Source of variation	df	Variance (sigma^2)	% Variance	Φ statistics	*p* value
ITS1	Species	6	1.96×10^{-5}	52.63	0.48	<0.01
	Populations within species	10	2.23×10^{-6}	6.00	0.55	0.58
	Within populations	197	1.54×10^{-5}	41.36	–	–
ITS2	Species	6	1.10×10^{-4}	73.50	0.7	<0.01
	Populations within species	10	1.32×10^{-5}	8.84	0.8	0.99
	Within populations	128	2.64	17.64	–	–

Table 3. Hierarchical AMOVA results for ITS1 and ITS2 regions.

Species	Source of variation	ITS1					ITS2				
		df	Variance (sigma^2)	Variance (%)	Φ	p	df	Variance (sigma^2)	Variance (%)	Φ	*p*
E. baeticum	Populations	2	1.48×10^{-5}	8.93	0.11	0.09	2	1.38×10^{-7}	0.3	0.01	0.46
	Individuals within populations	12	0	0	0	0.99	12	0	0	0	0.99
	Within individuals	24	1.51×10^{-4}	91.06	–	–	2	4.46×10^{-5}	99.69	–	–
E. bastetanum	Populations	2	0	0	0	0.96	2	8.05×10^{-5}	70.87	0.75	<0.01
	Individuals within populations	12	0	0	0	0.61	10	0	0	0.69	0.99
	Within individuals	31	2.33×10^{-4}	100	–	–	29	3.30×10^{-5}	29.12	–	–
E. fitzii	Individuals within populations	4	8.37×10^{-5}	55.03	0	0.14	3	0	0	0	0.9
	Within individuals	4	6.48×10^{-5}	44.96	–	–	5	6.08×10^{-5}	100	–	–
E. lagascae	Individuals within populations	4	6.50×10^{-8}	0.07	0	0.55	4	0	0	0	0.94
	Within individuals	2	8.48×10^{-5}	99.92	–	–	4	0.06	100	–	–
E. mediohispanicum	Populations	2	4.58×10^{-5}	48.07	0.5	<0.01	2	0	0	0	0.97
	Individuals within populations	12	0	0	0.45	0.69	12	0	0	0	0.28
	Within individuals	42	4.95×10^{-5}	51.92	–	–	10	5.08×10^{-5}	100	–	–
E. nevadense	Populations	2	1.67×10^{-5}	18.29	0.25	<0.01	2	0	0	0	0.99
	Individuals within populations	11	0	0	0	0.8	11	0	0	0	0.88
	Within individuals	7	7.47×10^{-5}	81.7	–	–	3	5.12×10^{-5}	100	–	–
E. popovii	Populations	2	3.13×10^{-5}	19.01	0.2	0.02	2	1.15×10^{-3}	7.73	0.09	0.56
	Individuals within populations	12	0	0	0.13	0.72	12	0	0	0	0.41
	Within individuals	20	1.33×10^{-4}	80.98	–	–	12	0.01	92.26	–	–

Table 4. ITS1 and ITS2 hierarchical AMOVA results for *E. baeticum*, *E. bastetanum, E. fitzii, E.lagascae, E. mediohispanicum*, *E. nevadense,* and *E. popovii.*

(< 9% for both ITS1 & ITS2; Table 3). When the genetic structure was separately analyzed for each species, we found more complex results. Most of the variance (44.96–100% for ITS1; 29.12–100% for ITS2) resided within-individuals (see Table 4). Differences among populations varied from 0 to 48.07% for ITS1 and from 0 to 70.87% for ITS2. Moreover, the differences were only significant in *E. mediohispanicum*, *E. nevadense*, and *E. popovii* for ITS1 and *E. bastetanum* for ITS2 (Table 4).

Discussion

We observed incomplete sequence homogenization for the 45S rDNA regions in the *Erysimum* species studied here. Our analyses were based on stringent trimming to avoid false polymorphisms due to sequencing errors. However, despite being so restrictive, we found high nucleotide and haplotype diversities overall, especially for ITS1, and a significant genetic structure that may inform the evolutionary history of these species.

Polyploid *Erysimum* species presented lower ITS homogenization levels than diploid species. Specifically, polyploid species presented higher nucleotide and haplotype diversity and a higher number of haplotypes, congruent with the hypothesis that polyploids harbor greater genetic diversity even within gene families[52]. The lack of concerted evolution in polyploid species has been previously described in several plant species in which an absence of sequence homogenization could be related to a recent allopolyploid origin[34,53–56]. Moreover, some studies have suggested that the number of rDNA loci, usually located in different chromosomes, is expected to be higher in polyploids, hindering sequence homogenization[57,58]. The number of rDNA loci and their chromosomic locations in these *Erysimum* species is unknown. In the genome of the diploid *E. cheiranthoides*[34], the rDNA appears in eight locations in chromosomes 3, 6, 7, and 8, which may be related to the number of rDNA loci for the diploid *Erysimum* species studied here. In any case, a relatively higher number of rDNA loci is expected for polyploid *Erysimum* species. Although the number of rDNA loci may coincide with the sum of those of its parents in young allopolyploids, it could be more variable in older polyploids, where some loci are usually lost[59–62].

We also detected limited sequence homogenization in diploid species, particularly in ITS1. The high molecular variance within diploid genomes (Table 4) could be the result of past hybridization events, which might result in the coexistence of multiple ITS families within individual genomes, particularly if hybrids are young[34,63]. This result is congruent with previous studies, in which the genomes of the diploid *Erysimum* species studied here were found to exhibit signatures of recent hybridization and introgression[33]. Moreover, *Erysimum* phylogenies based on ITS sequences[64–66] showed a variable degree of phylogenetic incongruence compatible with hybridization. Here, the influence of hybridization on ITS diversity is further supported by the significant molecular variance among populations detected in some species (i.e., *E. mediohispanicum*, *E. nevadense*, for ITS1), showing a non-consistent homogenization pattern in the population level. Thus, these results suggest a different history of hybridization for each population, in concordance with previous studies[32,33].

Our results indicated that sequence homogenization was heterogeneous across the 45S rDNA regions within a general scenario of high diversity (Tables S2–S8). The degree of polymorphism exhibited by ITS1 was much higher than that of ITS2, suggesting that concerted evolution is operating more efficiently on the latter. This result agrees with previous studies that have shown that ITS1 is, on average, more variable than ITS2, which has been described as a very conserved marker[67–72]. This variation between the two spacers might help to analyze evolutionary patterns at different scales. While ITS1 variation might throw light on divergence at the population- or individual-level, our AMOVA results (Table 3) suggest that ITS2 could be useful for species-level characterization, at least in *Erysimum* spp.

Because of their sensitivity to hybridization, ITS markers have been previously used to identify the parental contributors of hybrid taxa[17,53,64,73,74]. Our study found shared haplotypes among diploid and polyploid species (specifically among *E. bastetanum*—a polyploidy—and the diploid species *E. fitzii*, *E. mediohispanicum*, and *E. nevadense*), which could be the result of incomplete lineage sorting or the effect of recent hybridization events. However, we have not found decisive evidence of whether these diploid species could be considered parental species of the polyploid taxon. Moreover, our results indicate hybridization across taxonomic levels (i.e., from individuals to species) since they are more congruent with multiple backcrossings across populations and taxa than with a single, "original" allopolyploidization event. Reticulated evolution seems to be the norm in this genus[33,36,59,75–77]. Thus, for the species analyzed in this study, Osuna-Mascaro et al.[33] have found genomic evidence of rampant introgression between species, including both lilac- and yellow-flowered species. Future studies identifying the alleles co-located on the same chromosome through phased haplotypes[78] or using PacBio single-molecule sequencing and the PURC method (Pipeline for Untangling Reticulate Complexes;[79]) could be used to identify parental species of the different hybrid taxa and trace back the evolutionary patterns of these *Erysimum* species.

Despite their evident versatility as molecular evolution markers, the analysis of ITS sequences needs to be undertaken to realize that concerted evolution might often be insufficient to ensure sequence homogenization[80]. Both ITS and, especially, ITS2 have for a long time been used as phylogenetic and barcoding markers in plants[8,69,81–83]. However, many studies have pointed out that evolutionary inferences based on these markers might lead to misleading or erroneous conclusions in species where sequence homogenization is lacking due to hybridization or other genome rearrangement events[9,11]. In this study, our results indicate that allopolyploidization and hybridization have severely impaired ITS sequence homogenization in *Erysimum*, implying that ITS-based phylogenies of this genus should be considered with prudence. Given that these causes of genomic rearrangement are widespread and prevalent among flowering plants[11], caution is advised when using ITS for phylogenetic studies without prior knowledge of haplotype distribution, even for diploid species. Hence, intragenomic variation for ITS sequences could be used as an indication of possible recent hybridization.

Data availability

The raw data from this project were submitted to NCBI Sequence Read Archive (SRA) and can be found by the BioSample ID: SUB11440702, BioProject PRJNA835881, and the following accession numbers: *E. baeticum* (Ebb07: SAMN28120146, Ebb10: SAMN28120147, Ebb12: SAMN28120148); *E. bastetanum* (Ebt01: SAMN28120149, Ebt12: SAMN28120150, Ebt13: SAMN28120151); *E. fitzii:* Ef01 (SAMN28120152); *E. lagascae* (Ela07: SAMN28120153); *E. mediohispanicum* (Em21: SAMN28120154, Em39: SAMN28120155, Em71: SAMN28120156); *E. nevadense* (En05: SAMN28120157, En10: SAMN28120158, En12: SAMN28120159); *E. popovii* (Ep16: SAMN28120160, Ep20: SAMN28120161, Ep27: SAMN28120162).

Received: 29 May 2022; Accepted: 9 September 2022
Published online: 07 October 2022

References

1. Elder, J. F. Jr. & Turner, B. J. Concerted evolution of repetitive DNA sequences in eukaryotes. *Q. Rev. Biol.* **70**(3), 297–320 (1995).
2. Eickbush, T. H. & Eickbush, D. G. Finely orchestrated movements: Evolution of the ribosomal RNA genes. *Genetics* **175**, 477–485 (2007).
3. Dover, G. Concerted evolution, molecular drive, and natural selection. *Curr. Biol.* **4**(12), 1165–1166 (1994).
4. Ganley, A. R. & Kobayashi, T. Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res.* **17**(2), 184–191 (2007).
5. Feliner, G. N. & Rosselló, J. A. Concerted evolution of multigene families and homoeologous recombination. In *Plant Genome Diversity* (eds Wendel, J. F. *et al.*) 171–194 (Springer-Verlag, 2012).
6. Sone, T. *et al.* Bryophyte 5S rDNA was inserted into 45S rDNA repeat units after the divergence from higher land plants. *Plant Mol. Biol.* **41**(5), 679–685 (1999).
7. Long, E. O. & Dawid, I. B. Repeated genes in eukaryotes. *Annu. Rev. Biochem.* **49**(1), 727–764 (1980).
8. Baldwin, B. G. *et al.* The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* https://doi.org/10.2307/2399880 (1995).

9. Álvarez, I. & Wendel, J. F. Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* **29**(3), 417–434 (2003).
10. Xu, B. *et al.* ITS non-concerted evolution and rampant hybridization in the legume genus *Lespedeza* (Fabaceae). *Sci. Rep.* **7**, 40057 (2017).
11. Nieto-Feliner, G. & Rosselló, J. A. Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Mol. Phylogenet. Evol.* **44**(2), 911–919 (2007).
12. Teruel, M. *et al.* Disparate molecular evolution of two types of repetitive DNAs in the genome of the grasshopper *Eyprepocnemis plorans*. *Heredity* **112**(5), 531–542 (2014).
13. Buckler, E. S. & Holtsford, T. P. Zea systematics: Ribosomal ITS evidence. *Mol. Biol. Evol.* **13**(4), 612–622 (1996).
14. Mayol, M. & Rosselló, J. A. Why nuclear ribosomal DNA spacers (ITS) tell different stories in Quercus. *Mol. Phylogenet. Evol.* **19**(2), 167–176 (2001).
15. Popp, M. & Oxelman, B. Evolution of an RNA polymerase gene family in *Silene* (Caryophyllaceae) incomplete concerted evolution and topological congruence among paralogues. *Syst. Biol.* **53**(6), 914–932 (2004).
16. Harpke, D. & Peterson, A. Non-concerted ITS evolution in Mammillaria (Cactaceae). *Mol. Phylogenet. Evol.* **41**(3), 579–593 (2006).
17. Denk, T. & Grimm, G. W. The oaks of western Eurasia: Traditional classifications and evidence from two nuclear markers. *Taxon* **59**(2), 351–366 (2010).
18. Xiao, L. Q., Möller, M. & Zhu, H. High nrDNA ITS polymorphism in the ancient extant seed plant Cycas: Incomplete concerted evolution and the origin of pseudogenes. *Mol. Phylogenet. Evol.* **55**(1), 168–177 (2010).
19. Bailey, J. A., Liu, G. & Eichler, E. E. An Alu transposition model for the origin and expansion of human segmental duplications. *Am. J. Hum. Genet.* **73**(4), 823–834 (2003).
20. Won, H. & Renner, S. S. The internal transcribed spacer of nuclear ribosomal DNA in the gymnosperm Gnetum. *Mol. Phylogenet. Evol.* **36**(3), 581–597 (2005).
21. Zheng, X., Cai, D., Yao, L. & Teng, Y. Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in Pyrus. *Mol. Phylogenet. Evol.* **48**(3), 892–903 (2008).
22. Drábková, L. Z., Kirschner, J., Štěpánek, J., Záveský, L. & Vlček, Č. Analysis of nrDNA polymorphism in closely related diploid sexual, tetraploid sexual and polyploid species. *Plant Syst. Evol.* **278**(1–2), 67–85 (2009).
23. Okuyama, Y. *et al.* Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian Mitella (Saxifragaceae). *Mol. Biol. Evol.* **22**(2), 285–296 (2004).
24. Soltis, P. S. & Soltis, D. E. The role of hybridization in plant speciation. *Annu. Rev. Plant. Biol.* **60**, 561–588 (2009).
25. Polatschek A. *Erysimum*. *Mountain flora of Greece* (ed. Strid, A.) 239–247 (Cambridge University Press, 1986).
26. Warwick, S. I., Francis, A., Al-Shehba, I. & A,. Brassicaceae: species checklist and database on CD-Rom. *Plant Syst. Evol.* **259**, 249–258 (2006).
27. Al-Shehbaz, I. A. A generic and tribal synopsis of the Brassicaceae (Cruciferae). *Taxon* **61**, 931–954 (2012).
28. Nieto-Feliner, G. in *Erysimum L. Flora Ibérica* (Vol. IV. Cruciferae-Monotropaceae) 48–76 (Real Jardín Botánico, CSIC, Madrid, 1993).
29. Médail, F. & Diadema, K. Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *J. Biogeogr.* **36**(7), 1333–1345 (2009).
30. Abdelaziz, M. How Species are Evolutionarily Maintained? Pollinator-Mediated Divergence and Hybridization in *Erysimum mediohispanicum* and *Erysimum nevadense*. Doctoral Dissertation, Universidad de Granada (2013).
31. Muñoz-Pajares, A. J. *Erysimum mediohispanicum* at the Evolutionary Crossroad: Phylogrography, Phenotype, and Pollinators. Doctoral Dissertation, Universidad de Granada (2013).
32. Osuna Mascaró, C. Hybridization as an Evolutionary Driver for Speciation: A Case in the Southern European Erysimum Species. Doctoral Dissertation, Universidad de Granada (2020).
33. Osuna-Mascaró, C. *et al.* Hybridization and introgression are prevalent in Southern European *Erysimum* (Brassicaceae) species. *Ann. Bot.* https://doi.org/10.1093/aob/mcac048 (2022).
34. Züst, T. *et al.* Independent evolution of ancestral and novel defenses in a genus of toxic plants (*Erysimum*, Brassicaceae). *Elife* **9**, e51712 (2020).
35. Rauscher, J. T., Doyle, J. J. & Brown, A. H. D. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the Glycine (Leguminosae) polyploid complex. *Mol. Ecol.* **11**(12), 2691–2702 (2002).
36. Osuna-Mascaró, C., Rubio de Casas, R., Landis, J. B. & Perfectti, F. Genomic resources for *Erysimum* spp (Brassicaceae): Transcriptome and chloroplast genomes. *Front. Ecol. Evol.* **9**, 620601 (2021).
37. Andrews, S. FastQC: A quality control tool for high throughput sequence data. Available at: www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).
38. Martin, M. Cutadapt removes adapter sequences from highthroughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
39. Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) (2011).
40. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**(12), 1647–1649 (2012).
41. Bushnell, B., Rood, J. & Singer, E. BBMerge–accurate paired shotgun read merging via overlap. *PLoS ONE* https://doi.org/10.1371/journal.pone.0185056 (2017).
42. Li, W. & Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**(13), 1658–1659 (2006).
43. Esling, P., Lejzerowicz, F. & Pawlowski, J. Accurate multiplexing and filtering for high-throughput amplicon- sequencing. *Nucleic Acids Res.* **43**(5), 2513–2524 (2015).
44. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**(14), 3059–3066 (2002).
45. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**(15), 1972–1973 (2009).
46. Paradis, E. Pegas: A R package for population genetics with an integrated–modular approach. *Bioinformatics* **26**(3), 419–420 (2010).
47. Nei, M. & Li, W. H. Mathematical model for studying genetic variation in terms of restriction endonucleases. *PNAS* **76**(10), 5269–5273 (1979).
48. Nei, M. & Jin, L. Variances of the average numbers of nucleotide substitutions within and between populations. *Mol. Biol. Evol.* **6**(3), 290–300 (1989).
49. Villanueva, R. A. M. & Chen, Z. J. ggplot2: Elegant graphics for data analysis 160–167 (2019).
50. Team, R. C. R Core Team R: A language and environment for statistical computing. *Foundation for Statistical Computing* (2020).
51. Excoffier, L., Smouse, P. E. & Quattro, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* **131**(2), 479–491 (1992).
52. Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
53. Koch, M. A., Dobeš, C. & Mitchell-Olds, T. Multiple hybrid formation in natural populations: Concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Mol. Biol. Evol.* **20**(3), 338–350 (2003).
54. Kovarik, A. *et al.* Concerted evolution of 18–5.8–26S rDNA repeats in Nicotiana allotetraploids. *Biol. J. Linn. Soc.* **82**(4), 615–625 (2004).

55. Lunerová, J., Renny-Byfield, S., Matyášek, R., Leitch, A. & Kovařík, A. Concerted evolution rapidly eliminates sequence variation in rDNA coding regions but not in intergenic spacers in *Nicotiana tabacum* allotetraploid. *Biol. J. Linn. Soc.* **303**(8), 1043–1060 (2017).
56. Morales-Briones, D. F. & Tank, D. C. Extensive allopolyploidy in the neotropical genus *Lachemilla* (Rosaceae) revealed by PCR-based target enrichment of the nuclear ribosomal DNA cistron and plastid phylogenomics. *Am. J. Bot.* **106**(3), 415–437 (2019).
57. Wendel, J. F. Genome evolution in polyploids. *Plant Mol. Evol*. 225–249 (2000).
58. Kovarik, A. *et al.* Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* **169**(2), 931–944 (2005).
59. Clarkson, J. J. *et al.* Long-term genome diploidization in allopolyploid *Nicotiana* section Repandae (Solanaceae). *New Phytol.* **168**, 241–252 (2005).
60. Chester, M. *et al.* Extensive chromosomal variation in a recently formed natural allopolyploid species, Tragopogon miscellus (Asteraceae). *Proc. Natl. Acad. Sci.* **109**, 1176–1181 (2012).
61. Rebernig, C. A. *et al.* The evolutionary history of the white-rayed species of Melampodium (Asteraceae) involved multiple cycles of hybridization and polyploidization. *Am. J. Bot.* **99**, 1043–1057 (2012).
62. Weiss-Schneeweiss, H., Emadzade, K., Jang, T. & Schneeweiss, G. Evolutionary consequences, constraints and potential of polyploidy in plants. *Cytogenet. Genome Res.* **40**, 137–150 (2013).
63. Nieto-Feliner, G., Gutiérrez Larena, B. & Fuertes Aguilar, J. Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in Armeria (Plumbaginaceae). *Ann. Bot.* **93**(2), 189–200 (2004).
64. Abdelaziz, M. *et al.* Phylogenetic relationships of *Erysimum* (Brassicaceae) from the Baetic Mountains (se Iberian peninsula). *An. Jard. Bot. Madr.* **71**, 005 (2014).
65. Gómez, J. M., Perfectti, F. & Klingenberg, C. P. The role of pollinator diversity in the evolution of corolla-shape integration in a pollination-generalist plant clade. *Philos. Trans. R. Soc. B.* **369**, 20130257 (2014).
66. Moazzeni, H. *et al.* Phylogenetic perspectives on diversification and character evolution in the species-rich genus *Erysimum* (Erysimeae; Brassicaceae) based on a densely sampled ITS approach. *Bot. J. Linn. Soc.* **175**(4), 497–522 (2014).
67. Hershkovitz, M. A. & Zimmer, E. A. Conservation patterns in angiosperm rDNA ITS2 sequences. *Nucleic Acids Res.* **24**(15), 2857–2867 (1996).
68. Coleman, A. W. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet.* **19**(7), 370–375 (2003).
69. Chen, S. *et al.* Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* https://doi.org/10.1371/journal.pone.0008613 (2010).
70. Buchheim, M. A. *et al.* Internal transcribed spacer 2 (nu ITS2 rRNA) sequence-structure phylogenetics: Towards an automated reconstruction of the green algal tree of life. *PLoS ONE* https://doi.org/10.1371/journal.pone.0016931 (2011).
71. Wang, X. C. *et al.* ITS1: A DNA barcode better than ITS 2 in eukaryotes?. *Mol. Ecol. Resour.* **15**(3), 573–586 (2015).
72. Yang, R. H. *et al.* Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS ONE* https://doi.org/10.1371/journal.pone.0206428 (2018).
73. Sun, K., Ma, R., Chen, X., Li, C. & Ge, S. Hybrid origin of the diploid species *Hippophae goniocarpa* evidenced by the internal transcribed spacers (ITS) of nuclear rDNA. *Belg. J. Bot*. 91–96 (2013).
74. Hodač, L., Scheben, A. P., Hojsgaard, D., Paun, O. & Hörandl, E. ITS polymorphisms shed light onhybrid evolution in apomictic plants: a case study on the *Ranunculus auricomus* complex. *PLoS ONE* https://doi.org/10.1371/journal.pone.0103003 (2014).
75. Clot, B. Caryosystématique de quelques *Erysimum* L. dans le nord de la Péninsule Ibérique. *Anal. Jardín Botán. Madrid* **49**, 215–229 (1992).
76. Marhold, K. & Lihová, J. Polyploidy, hybridization and reticulate evolution: Lessons from the Brassicaceae. *Plant Syst. Evol.* **259**(2), 143–174 (2006).
77. Turner, B. L. Taxonomy and nomenclature of the *Erysimum asperum-E.capitatum* complex (Brassicaceae). *Phytologia* **88**, 279–287 (2006).
78. Browning, S. R. & Browning, B. L. Haplotype phasing: Existing methods and new developments. *Nat. Rev. Genet.* **12**(10), 703–714 (2011).
79. Rothfels, C. J., Pryer, K. M. & Li, F. W. Next-generation polyploid phylogenetics: Rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytol.* **213**(1), 413–429 (2017).
80. Song, H. X. *et al.* The evolution and utility of ribosomal ITS sequences in Bambusinae and related species: Divergence, pseudogenes, and implications for phylogeny. *J. Genet.* **91**(2), 129–139 (2012).
81. Hughes, C. E., Eastwood, R. J. & Donovan Bailey, C. From famine to feast? Selecting nuclear DNA sequence loci for plant species-level phylogeny reconstruction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **361**(1465), 211–225 (2006).
82. Mishra, P., Kumar, A., Rodrigues, V., Shukla, A. K. & Sundaresan, V. Feasibility of nuclear ribosomal region ITS1 over ITS2 in barcoding taxonomically challenging genera of subtribe Cassiinae (Fabaceae). *PeerJ* **4**, e2638 (2016).
83. Cheng, T. *et al.* Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Mol. Ecol. Resour.* **16**(1), 138–149 (2016).

Acknowledgements

The authors thank the Tatiana López Pérez and the Evoflor group for helping us during several phases of the study. We also thank the Sierra Nevada National Park headquarters for providing access to sampling in the National Park.

Author contributions

C.O.M., R.R., and F.P. conceived and designed the study. C.O.M. and M.B. carried out the laboratory procedures and field sampling, with the help of F.P. and J.M.G. C.O.M. analyzed the data with the help of F.P. C.O.M. wrote the first draft. The final version of the MS was redacted with the contribution of all the authors.

Funding

This research is supported by grants from FEDER/Junta de Andalucía-Consejería de Economía y Conocimiento A-RNM-505-UGR18 and P18-FR-3641. This research was also funded by the Spanish Ministry of Science and Innovation (CGL2013-47558-P and PID2021-126456NB-C22), including EU FEDER funds. COM was supported by the Ministry of Economy and Competitiveness (BES-2014-069022). This is a contribution to the Research Unit Modeling Nature, funded by the Consejería de Economía, Conocimiento, Empresas y Universidad, and European Regional Development Fund (ERDF), reference QUALIFICA 00011.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-20194-8.

Correspondence and requests for materials should be addressed to C.O.-M. or F.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.