Turczaninowia 27, 3: 97–109 (2024)
DOI: 10.14258/turczaninowia.27.3.10
http://turczaninowia.asu.ru

ISSN 1560–7259 (print edition)
TURCZANINOWIA
ISSN 1560–7267 (online edition)

УДК 582.998.4+575.174.015.3(470)

Evaluation of intragenomic polymorphism of ribosomal DNA ITS1 region in *Taraxacum* (Asteraceae: *Crepidinae*) by targeted NGS sequencing

P. G. Efimov[1, 2*], V. V. Domashkina[1, 3], E. M. Machs (†), P. M. Zhurbenko[1, 4]

[1] *Komarov Botanical Institute RAS, Prof. Popova St., 2, St. Petersburg, 197022, Russian Federation*

[2] *E-mail: efimov@binran.ru, efimov81@mail.ru; ORCID iD: https://orcid.org/0000-0003-2926-255X*

[3] *E-mail: domvalya@gmail.com; ORCID iD: https://orcid.org/0000-0001-9222-9504*

[4] *E-mail: pj_28@mail.ru; ORCID iD: https://orcid.org/0000-0002-2102-4568*

* *Corresponding author*

Keywords: apomixis, DNA homogenization, DNA phylogenetics, reticulate evolution, ribosomal DNA, *Taraxacum*.

Summary. The picture of intragenomic polymorphism of ITS1 region of rDNA repeats revealed by targeted high-throughput sequencing is a modern and simple method useful for phylogenetic purposes, and it can shed light on recent hybridizations. However, up to the present time it was not applied in *Taraxacum*, a taxonomically intricate genus with largely reticulate evolutionary pattern. We tested this approach on an exploratory set composed of 20 samples (8 species from 4 sections, including 5 species from the largest section *Taraxacum*), which were situated in different evolutionary and geographic distances from each other. Our results have shown that intragenomic polymorphism of ITS1 region of rDNA in *Taraxacum* may be successfully used for phylogenetic reconstructions. Species belonging to different sections were strikingly different by intragenomic polymorphism of rDNA, and substantial variability was also observed between some (but not all) specimens of section *Taraxacum*, giving preliminary evidence that this group may be not monomorphic. In the same time, 8 of 13 studied specimens from the section *Taraxacum* had ITS1 region substantially homogenized (dominant variant was at least 7 times more frequent than any other), which is considered connected with the age of the last hybridization event in corresponding lineages. Possibly, microspecies with substantially homogenized rDNA represent more strongly stabilized taxonomic units.

Оценка внутригеномного полиморфизма района ITS1 рибосомальной ДНК видов рода *Taraxacum* (Asteraceae: *Crepidinae*) методами высокопроизводительного секвенирования

П. Г. Ефимов, В. В. Домашкина, Э. М. Мачс (†), П. М. Журбенко

Ботанический институт им. В. Л. Комарова РАН, ул. Проф. Попова, д. 2, г. Санкт-Петербург, 197022, Россия

Ключевые слова: апомиксис, гомогенизация ДНК, молекулярная филогенетика, рибосомальная ДНК, сетчатая эволюция, *Taraxacum*.

Аннотация. Выявление внутригеномного полиморфизма района ITS1 рибосомальной ДНК методами высокопроизводительного секвенирования перспективно при исследованиях филогении близкородственных таксонов, а также для идентификации гибридов. Однако до настоящего времени эта методика не использовалась в роде *Taraxacum*, таксономически очень сложной группе с широко распространённой сетчатой эволюцией. Мы применили данный метод на тестовой выборке из 20 образцов (8 видов из 4 секций, включая 5 видов

Поступило в редакцию 03.05.2024
Принято к публикации 17.10.2024

Submitted 03.05.2024
Accepted 17.10.2024

из наиболее крупной секции *Taraxacum*), находящихся на различных эволюционных и географических дистанциях друг от друга. Использованный метод оказался эффективным для изучения межвидового генетического полиморфизма. Так, виды, относящиеся к различным секциям рода, сильно отличались друг от друга по набору и частотам различных вариантов исследованного фрагмента рДНК. Существенное разнообразие было выявлено также в пределах секции *Taraxacum*, что говорит об её неоднородности. В то же время, у 8 из 13 образцов этой секции исследованный фрагмент рДНК был сильно гомогенизирован (преобладающий вариант встречался не менее чем в 7 раз чаще любого другого), что можно связывать с более отдалённым во времени случаем последней гибридизации в их эволюционной истории. Вероятно, микровиды с гомогенизированной рДНК представляют собой более стабилизированные таксономические единицы.

Introduction

The genus *Taraxacum* F. H. Wigg. (Asteraceae: *Crepidinae* Cass. ex Dumort.) is notable for high taxonomic diversity and a variety of evolutionary mechanisms involved in diversification, such as different modes of reproduction, polyploidy, complex hybridity.

Currently we know ca. 2500 species in 60 sections in the genus (Kirschner et al., 2022), however, it is considered that many species are yet undescribed, and there is still no phylogenetically approved classification for the genus as a whole. Main obstacle for phylogenetic reconstructions and classifications in *Taraxacum* is complicated reticulate pattern of evolution (Kirschner et al., 2015).

In 2016 we started taxonomical inventory research of *Taraxacum* focused on North and North-West European Russia. Taxa which have been already reported from here fall at least in 14 presently approved sections (based on Tzvelev, 1989; Lundevall, Øllgaard, 1999; Sennikov, 2007 and unpublished data), but many sections may represent only preliminarily delimited morphogroups instead of monophyletic phylogenetic clades, and many are unclearly defined.

It is essential that taxonomic studies of such complicated genera like *Taraxacum* are accompanied by independent phylogenetic evidence, like that provided by molecular phylogenetic criteria. Molecular phylogenetic data are crucially important for providing correct classifications on the above-species level, and in theory, may be helpful to define individual apomictic lineages, which otherwise may remain very subjectively defined.

Many kind of molecular phylogenetic studies have been tested in *Taraxacum* (their short reviews were earlier provided by Kirschner et al. (2015) and Salih et al. (2017)). The first genus-level phylogenies revealed relatively weak and inconsistent differentiation within the genus. Thus, phylogenetic reconstruction after plastome uncoding sequences and RFLP sites by Kirschner et al. (2003) using taxa from 44 sections (or species aggregates) rendered generally low congruence between chloroplast and morphological data, which was interpreted as a consequence of reticulation.

Nuclear rDNA variation in *Taraxacum* was studied by several authors (Mes et al., 2002; Uhlemann et al., 2009; Záveská Drábková et al., 2009; Kirschner et al., 2015; Machackova et al., 2020; etc.) and also provided limited sectional-level resolution, but evidenced close relationship in several cases, e. g. between sections *Erythrosperma* (H. Lindb.) Dahlst. and *Erythrocarpa* Hand.-Mazz., *Arctica* Dahlst. and *Antarctica* Hand.-Mazz. Kirschner et al. (2015) also pointed out an association of presumably ancestral Asian sections, and Uhlemann et al. (2009) argued for the clade consisting of 'derived European sections'. All taxa studied in the current contribution fall in the latter group. In the context of our study it is important that almost all studies of nuclear rDNA variation cited above have shown smaller intragenomic variation in sexual species comparing to apomicts, which was explained by suppressed concerted evolution of nrDNA (this thesis is again discussed in the present contribution, in relation to section *Taraxacum* only).

The studies focused on clonal structure of dandelion populations and differentiation of separate apomictic clones by molecular methods were successful by variety of approaches. Chloroplast DNA markers, microsatellite loci, AFLPs, ScoT markers, and other methods were applied for that purpose (Mes et al., 2002; Majeský et al., 2015; Kirschner et al., 2016; Wolanin et al., 2023). Of them, study by Kirschner et al. (2016) was a milestone for it confirmed the possibility for correct identification of individual apomictic lineages by morphology in the most variable and taxonomically intricate section *Taraxacum*. In that study, all of the nine studied apomictic species were shown to be determined consistently by different specialists across Europe: grouping after 9 polymorphic microsatellite loci was almost identical with that based on preliminary determinations (with only three exceptions among

125 samples). Complete chloroplast sequences were also illustrative for clonality (from maternal lineages), as in two of three species from the section *Taraxacum* plastids were completely identical (Salih et al., 2017). Additional complete chloroplast sequences of 20 *Taraxacum* specimens from North Europe may be assembled in future from the genome skim database of the PhyloNorway project (Alsos et al., 2020).

Molecular characterization of lineages, as well as intrasectional structure in *Taraxacum* remains very poorly investigated. At the same time, such studies are very important for gradual accumulation of phylogenetic evidence in complicated groups like *Taraxacum*, when phylogeny of the genus cannot be immediately outlined. Therefore, it is very valuable to develop a convenient method for identification of the groups of taxa which are closely related to each other, possibly shared by similar origin, at the 'above-clonal' taxonomic level in *Taraxacum*. In the current study, we tested the phylogenetic utility of intragenomic polymorphism of the short fragment of rDNA repeats for that purpose.

Ribosomal DNA consists of tandem repeats which may be different from each other first of all by SNP sites or indels. The relative abundance of different variants within the genome is dynamic due to their loss or multiplication, provided by homologous recombination and other mechanisms (e. g. Nelson et al., 2019). With time, some variants may be completely lost in genome, and new variants may appear in the course of continuous mutations, recombination and other processes. Relative frequencies of all rDNA variants and the presence of specific minor variants can be revealed by targeted NGS (next-generation sequencing) and provide important phylogenetic information, especially for closely related, weakly diverged taxa.

Existing studies of intragenomic rDNA polymorphism in *Taraxacum* (Záveská Drábková et al., 2009; Machackova et al., 2020), although based on relatively broad species and sectional sampling, didn't focuse on the evaluation of taxonomic utility of intragenomic spectrum of rDNA variants, and revealed only several major variants in each genome. Current study differs from them by applying targeted NGS sequencing approach, recovering detailed picture of intragenomic rDNA polymorphism and providing deeper insight of its phylogenetic significance.

Materials and methods

Selection of specimens for analysis

For the present study we selected a testing set of 20 samples. Main idea was to include specimens positioned at various phylogenetic and geographical distances from each other (different sections, well-delimited microspecies, poorly delimited microspecies, distant localities, closely situated localities). The source of all samples were herbarium specimens, the complete list of vouchers is presented in Table 1. The images of all voucher herbarium specimens may be downloaded from https://herbariumle.ru/?t=occ (to be searched after barcode numbers).

Table 1

Samples *Taraxacum* used in the present study

	Sampled as	Section according to: Lundevall, Øllgaard, 1999	Voucher herbarium label
1	*T. ingens* Palmgr.	*Taraxacum*	Pskov Region, Plyussa District, near Sheregi, along the road to Zapesen'ye, forest edge, N58°25', E29°39', 21 V 2022. P. G. Efimov. 2022/446 (plant 2) (LE 01202491)
2	*T. ingens* Palmgr.	*Taraxacum*	Pskov Region, Plyussa District, Zapol'ye, near the church, roadside near the old park, N58°22', E29°41', 21 V 2022. P. G. Efimov. 2022/457 (LE 01202503)
3	*T. ingens* Palmgr.	*Taraxacum*	Pskov Region, Porkhov District, Pochepovo, meadows and roadsides in the village, N57°57', E29°41', 21 V 2022. P. G. Efimov. 2022/481 (LE 01202444)
4	*T. alatum* H. Lindb.	*Taraxacum*	Republic of Karelia, Segezha District, Nadvoitsy, near Voitskiy Padum waterfall, along trail, N63°53', E34°19', 18 VI 2022. P. G. Efimov, V. V. Kuropatkin, M. V. Legchenko and M. B. Sheludyakova. 2022/1005 (LE 01190346)

Table 1 (continued)

	Sampled as	Section according to: Lundevall, Øllgaard, 1999	Voucher herbarium label
5	*T. alatum* H. Lindb.	*Taraxacum*	Leningrad Region, Luga District, between Vitovo and Mokrovo villages, roadside, dark shade, N58°32', E29°43', 21 V 2022. P. G. Efimov. 2022/440 (LE 01202483)
6	*T. alatum* H. Lindb.	*Taraxacum*	Leningrad Region, Slantsy District, between villages Kamenka and Pechurki, meadow near the rural road, N59°09', E28°01', 28 V 2021. P. G. Efimov. 2021/469, plant 1 (LE 01149630)
7	*T. trilobatum* Palmgr. (determination may be incorrect according to the present study)	*Taraxacum*	Republic of Komi, Vuktyl District, Vuktyl town, a lawn in the courtyard, N63°51', E57°18', 3 VII 2022. P. G. Efimov, V. V. Kuropatkin and M. A. Makarova. 2022/1298 (LE 01149981)
8	*T. trilobatum* Palmgr.	*Taraxacum*	Arkhangelsk Region, Onega District, Onega, central part of the town, shady place under the trees at the coast of Onega river, N63°54', E38°06', 24 VI 2022. P. G. Efimov, V. V. Kuropatkin, M. V. Legchenko and M. B. Sheludyakova. 2022/1129 (LE 01190292)
9	*T. trilobatum* Palmgr.	*Taraxacum*	Arkhangelsk Region, Maloshuyka District, Abramovskaya, meadow near the old churches, N63°45', E37°25', 20 VI 2022. P. G. Efimov, V. V. Kuropatkin, M. V. Legchenko and M. B. Sheludyakova. 2022/1062 (LE 01190215)
10	*T. planum* Raunk. (determination may be incorrect according to the present study)	*Taraxacum*	Republic of Karelia, Suoyarvi District, south part of Porosozero village, ruderal place along the road, N62°43', E32°43', 11 VI 2022. P. G. Efimov, G. Yu. Konechnaya and M. V. Legchenko. 2022/845 (LE 01202198)
11	*T. planum* Raunk.	*Taraxacum*	Saint-Petersburg, Frunzenskiy District, along Bukharestskaya street, near the house no. 114, more or less abandoned lawn, N59°51', E30°24', 13 V 2021. P. G. Efimov. 2021/133 (LE 01149272)
12	*T. planum* Raunk.	*Taraxacum*	Saint-Petersburg, Frunzenskiy District, along Bukharestskaya street, near the house no. 114, more or less abandoned lawn, N59°51', E30°24', 13 V 2021. P. G. Efimov. 2021/134 (LE 01149273)
13	*T. perattenuatum* H. Lindb.	*Boreigena*	Murmansk Region, Kandalaksha District, 2 km SE from Kandalaksha, road to Umba, a viewpoint, meadow-like vegetation near the road, N67°06', E32°11', 15 VI 2022. P. G. Efimov, G. Yu. Konechnaya and M. V. Legchenko. 2022/965(plant 3) (LE 01190203)
14	*T. perattenuatum* H. Lindb.	*Boreigena*	Arkhangelsk Region, Mezen' District, ca. 4 km N from Mezen' village, meadow, N65°55', E44°11', 27 VI 2022. P. G. Efimov, V. V. Kuropatkin and M. A. Makarova. 2022/1193 (LE 01202339)
15	*T. perattenuatum* H. Lindb.	*Boreigena*	Arkhangelsk Region, Mezen' District, near Bereznik village, meadow, N65°28', E45°01', 28 VI 2022. P. G. Efimov, V. V. Kuropatkin and M. A. Makarova. 2022/1203 (LE 01200571)
16	*T. lojoense* H. Lindb.	*? Borea*	Saint-Petersburg, Frunzenskiy District, Geroyev-Pozharnykh Park close to Sofiyskaya street, near the fence, at the limit of the lawn and waste-ground, N59°51', E30°25', 17 V 2021. P. G. Efimov. 2021/269 (LE 01149419)
17	*T. lojoense* H. Lindb.	*? Borea*	Saint-Petersburg, Kurortnyi District, area adjacent to the Gulf of Finland ca. 1.5 km SW from the Sestroretsk railway station, sparse woodland or ruderal place along the trail, N60°05', E29°56', 21 V 2021. P. G. Efimov. 2021/378 (LE 01149533)
18	*T. lojoense* H. Lindb.	*? Borea*	Leningrad Region, Priozersk District, Konevets Island in the Ladoga Lake, S part of the island, roadside, N60°51', E30°36', 6 VI 2021. P.G. Efimov. 2021/629 (LE 01149804)

Table 1 (continued)

	Sampled as	Section according to: Lundevall, Øllgaard, 1999	Voucher herbarium label
19	*T. hollandicum* Soest	*Palustria*	Leningrad Region, [Gatchina District], near Gatchina, N[S] part of the Gatchina Park [Zverinets Park], meadow on the gentle slope of the hill, [N59°57', E30°11'], 4 VI 1974, N. Tzvelev. 19 (LE 01229552)
20	*T. pectinatiforme* H. Lindb.	*Taraxacum*	Leningrad Region, Slantsy District, between villages Kamenka and Pechurki, meadow near the rural road, N59°09', E28°01', 28 V 2021. P. G. Efimov. 2021/472 (LE 01149634)

All specimens were determined by the first author based on the existing literature (most important sources: Uhlemann, 2003; Trávníček et al., 2010; Uhlemann et al., 2016) and after comparisons with herbarium collections previously determined by renown experts in the field, kept both in the Herbarium of the Komarov Botanical Institute (LE) and in the other botanical institutions (accessed by GBIF. URL: www.gbif.org and other sources), as well as after personal expertise. The *Taraxacum hollandicum* specimen was earlier identified by Jan Kirschner (no. det. 3562) and Nikolai Tzvelev.

The testing set includes taxa from 4 sections (*Palustria* (H. Lindb.) Dahlst., *Boreigena* G. E. Haglund, *Borea* A. J. Richards, *Taraxacum*), of which section *Palustria* was taken as the most distantly related to the others. On the opposite, sections *Borea* and *Taraxacum* were taken as presumably the most closely related, known for several taxa to oscillate in between (Lundevall, Øllgaard, 1999). Only one species was taken from each of the sections *Palustria, Boreigena* and *Borea*, viz. *T. hollandicum*, *T. perattenuatum* and *T. lojoense*, whereas 5 taxa were taken from section *Taraxacum*. The latter 5 taxa fall into 3 or 4 informal species groups ('morphogroups') accepted (but variously named) by *Taraxacum* experts. Of them, *T. alatum* and *T. ingens* represented a pair of morphologically similar species, incorporated in one informal morphogroup (Uhlemann, 2003; Uhlemann et al., 2016) or their similarity was directly stated (Trávníček et al., 2010). *Taraxacum trilobatum* and *T. planum* were taken as another pair of similar taxa, classified in one informal group by Uhlemann (2003) and Richards (2021), but they were recently distinguished in different morphogroups by Uhlemann et al. (2016). *Taraxacum pectinatiforme* was accepted and classified in a separate morphogroup in all of the above-mentioned classifications. Pairs of closely related species were taken bearing in mind that misdeterminations are possible; here, the main idea was to include one more level, intermediate between 'well-distinguishable microspecies' and 'specimens of one microspecies'.

All species, except *T. hollandicum* and *T. pectinatiforme*, were taken in three accessions: two from closely situated localities and one from geographically distant one. The distance between samples of one species ranged from 600 km to several meters.

All samples were collected recently (in 2021–2022), the herbarium was prepared personally using electric or diesel heater aimed with high-quality drying with minimal DNA degradation. The sole exception was *T. hollandicum* specimen collected as far as in 1974, this herbarium specimen was prepared by traditional methods (slow drying, green color of the leaves faded). *Taraxacum hollandicum* was collected in the locality where it was alien, probably as a result of unintentional introduction from Central Europe with cultivated flowering plants (Tzvelev, 1989) or with hay during World War II. Nowadays the plants are not extant in the locality (pers. obs.).

Existing data about species sampled for the current study (Trávníček et al., 2010) and the patterns of distribution of sexual taxa in the genus (Kirschner et al., 2022) argue that all (or almost all) plants sampled in the current study are apomictic triploids.

DNA isolation and NGS sequencing

DNA was extracted from herbarium specimens using the CTAB-method (Doyle J. J., Doyle J. L., 1987). The target DNA represented part of 35S rDNA, consisting of the 3'-end of 18S rRNA gene, the complete sequence of ITS1, and the 5'-end of 5.8S rRNA gene (further referred to as 'ITS1 region').

ITS1 region was enriched by PCR using primers *ITS2* (White et al., 1990) and *ITS1P* (Ridgway et al., 2003). Library preparation and pair-end sequencing with 300-bp reads were performed on the Illumina MiSeq at the Core Centre of "Genomic Technologies, Proteomics and Cell Biology" of the All-Russian Research Institute for Agricultural Microbiology (Saint-Petersburg, Russia).

Bioinformatic analysis

Raw FASTQ files were processed according to the USEARCH v.11 pipeline (Edgar, 2010) but with some differences. Unlike the pipeline, we used the dereplicated sequences instead of constructing "zero-radius operational taxonomic units" (ZOTUs). The reason was that this algorithm tries to reduce sequencing errors by deleting reads that differ by one or several substitutions from the assumed biological sequences. As the studied species were closely related, most of their ITS1 variants differed by only one nucleotide, so this correction procedure might unintentionally reduce the true diversity.

Given the above, the data were analyzed as follows. At first, merged forward and reverse reads of each sample were pooled together, quality filtered and dereplicated. Then, the chimera detection was performed and the resulted reads were used to construct the table of variant across samples. Finally, the variants were normalized by total sum scaling, the final list is presented in Electronic Appendix as fasta-file (see appendix on the journal website). Variants with frequencies less than 0.5 % were discarded from further analysis, the remaining sequences were numbered consecutively (v1, v2, v3, etc.).

Genetic distances calculation and visualization

Visualization of observed differences between samples according to the numbers of shared variants of different ITS1 variants and their frequencies was done by the means of Principal Coordinate Analysis (PCoA). The pairwise distances between samples A and B were calculated according to the following formula:

$$Distance = \frac{\sum_{i=1}^{n} \quad (1 - \frac{B_i}{A_i})}{n},$$

where n – number of ITS1 variants represented in samples A and/or B, A_i and B_i – frequencies of variant *i* in samples A and B, A*i* = max(A*i*, B*i*) and B*i* = min(A*i*, B*i*). Thus, if ITS1 variant *i* was present only in one sample, it rendered distance 1; if it was present in both samples with the same frequency, then it rendered 0; if it was present in both samples with different frequencies, it rendered the value between 0 and 1. Graphical presentation of distances between samples was done using python *skbio* module (Rideout et al., 2023).

Sequence alignment was performed using MAFFT v.7 (Katoh et al., 2013) and then visualized in MegaX (Kumar et al., 2018). The maximum likelihood phylogenetic tree was constructed in IQ-TREE v.1.6.1 (Nguyen et al., 2015) and visualized in iTOL (Letunic, Bork, 2021).

The variant network was built in TCS v.2.1 (Clement et al., 2000) and visualized in TCS BU (Múrias dos Santos et al., 2016). The graph nodes stand for different variants of ITS1 region, and graph links stand for single-nucleotide substitution. Indels are not considered, therefore variants which differed by indels only collapse as one node. The frequences of different variants are also not considered.

Results

The sequences of ITS1 region were successfully acquired for each of the 20 samples, varying in length from 196 to 333 bp. 13709–36075 reads were obtained from each sample (averaged 28651). The number of reads from oldest specimen (*T. hollandicum*) was 29671, which is near to the average value, showing that 50-year old specimens can be successfully used in such study.

In total, we observed 52 variants of ITS1 region (with frequencies not less than 0.5 %, at least in one sample). Data matrix with their frequencies in all samples is shown in Table 2. The number of different variants per sample was 3–21. Minimal diversity was observed in section *Taraxacum*, maximal diversity – in section *Palustria* (*T. hollandicum*).

The most homogenous ITS1 region was observed in 6-sample group from the section *Taraxacum,* further termed 'core section *Taraxacum*'. The dominant variant in this group (v1 in Table 2) was 12×–17× more common than any other variant. 'Core' included three specimens sampled as *T. alatum*, one *T. pectinatiforme*, one *T. trilobatum* and one *T. ingens*. Of them, *T. pectinatiforme* and *T. trilobatum* had the most homogenous ITS1 region (only three variants of ITS1 region in each: v1, v4, v16). Two other samples of *T. ingens* were also rather similar to the 'core', their dominant ITS1 region variant was 7× more common than any of the others. Other samples of sect. *Taraxacum* (two were sampled as *T. trilobatum* and three – as *T. planum*)

had less homogenous ITS1 region, and in all other sections sampled this heterogeneity was even more pronounced, in the same time corresponding (with few exceptions) to the preliminary taxonomic determinations.

PCoA (Fig. 1) based on frequencies illustrates ± satisfactory grouping of the studied specimens according to the taxonomy. The most distinct sample was that of *T. hollandicum* (sect. *Palustria*). Two of its three most common ITS1 region variants (v10 and v24, see Table 2) and 14 minor ones (v25, v30–v33, v38, v41–v42, v45–v50) were found exclusively in this taxon, indicating its separate phylogenetic position. Three specimens sampled as *T. lojoense* (sect. *Borea*) had two species-specific subdominant variants (v3 and v6), and several minor ones (v17, v19, v27). Three specimens of *T. perattenuatum* (sect. *Boreigena*) also had variants present in this species only (v7, v28, v37). In the same time, two ITS1 region variants (v2 and v8) proved to be specific to *T. perattenuatum* + *T. lojoense*, indicating their possible relationship. Very small distances were observed between three sampled specimens of *T. perattenuatum*, and the same was true for three samples of *T. lojoense* (Fig. 1); both constituted very monomorphic groups, despite large geographic distance between some of the samples (Fig. 2).

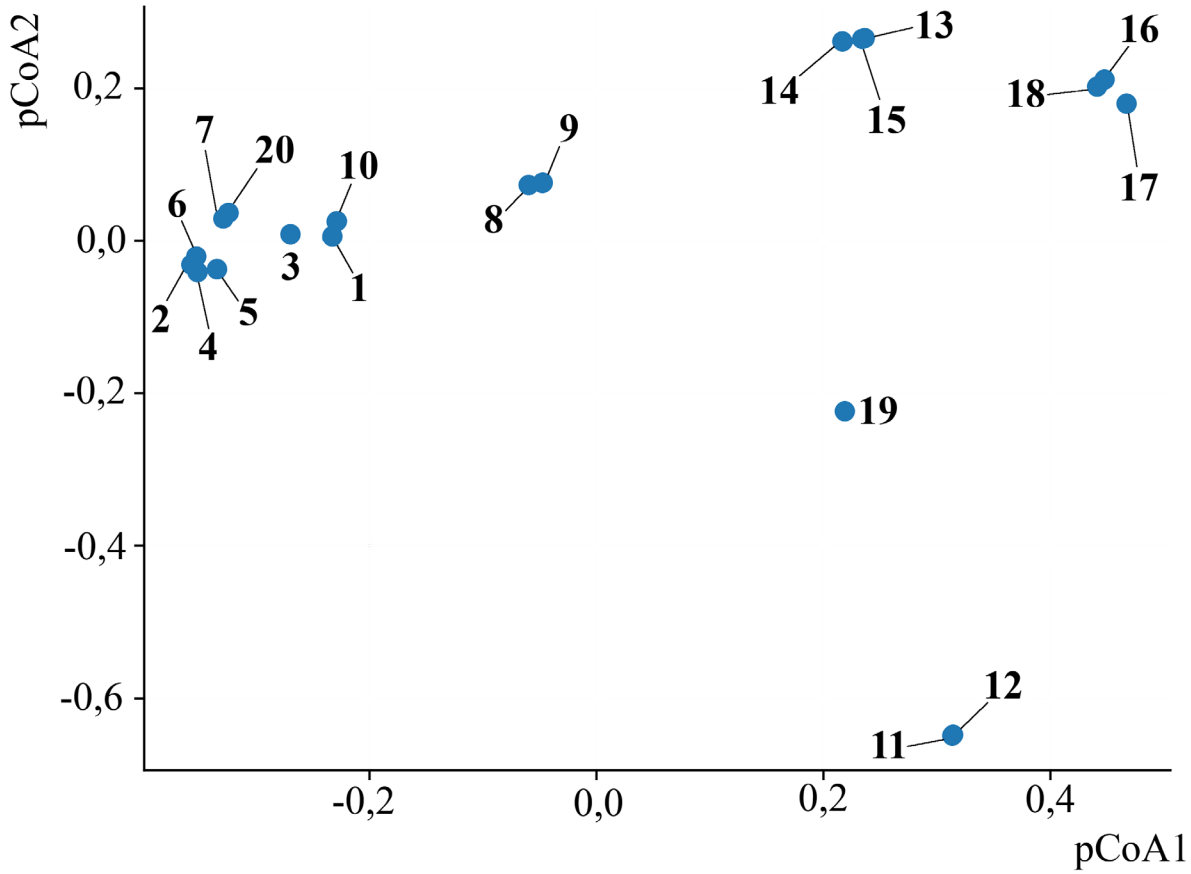

Fig. 1. PCoA of the studied specimens. Specimen numbering according to Table 1.

As stated earlier, the studied specimens from the section *Taraxacum* proved to be more variable: some belong to the 'core' with very monomorphic ITS1 region, others fall into 2–3 separate groups. Of them, the most distant from the 'core' were two specimens sampled as *T. planum*, they were growing several meters from one another and proved to be substantially different from the third sample of '*T. planum*', collected distantly in the Republic of Karelia, evidencing that they may belong to different microspecies (the first possible case of misidentification). Two of three specimens sampled as *T. trilobatum* constituted another group, again distinct from the third specimen of this species (the second possible case of misidentification). The *T. planum* specimen from the Republic of Karelia had prominent similarities with the *T. trilobatum* group, indicating their possible relationship.

The differences between nucleotide sequences of different ITS1 region variants screened in the studied specimens were not large. Mostly 1–2 single-nucleotide substitutions were observed (Fig. 3). However, it is worth mentioning that we came across several minor variants with two relatively large deletions. One deletion was 36 bp long and was present in many variants with low frequencies (below 0.5), except for variant v51, which reached frequency 0.57 in one accession of *T. planum* (Table 2). Variants with such deletion were found (in small quantities) in all of the taxa analyzed, without prominent correspondence with taxonomy or other trend. The second deletion, 133 bp long, also as minor variant, was found only in two accessions of *T. planum* (no. 11 and 12) only.

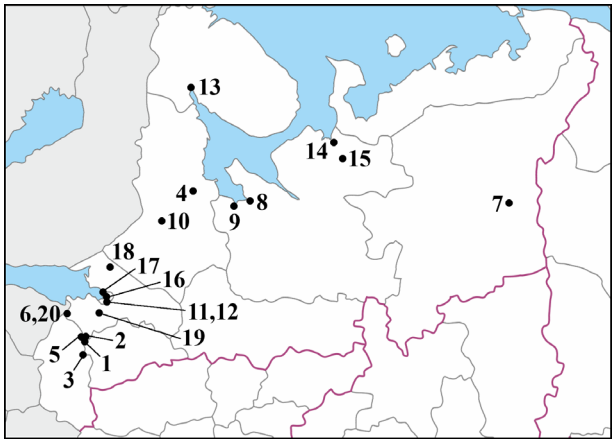

Fig. 2. Geographic localities of sampling. Specimen numbering according to Table 1.

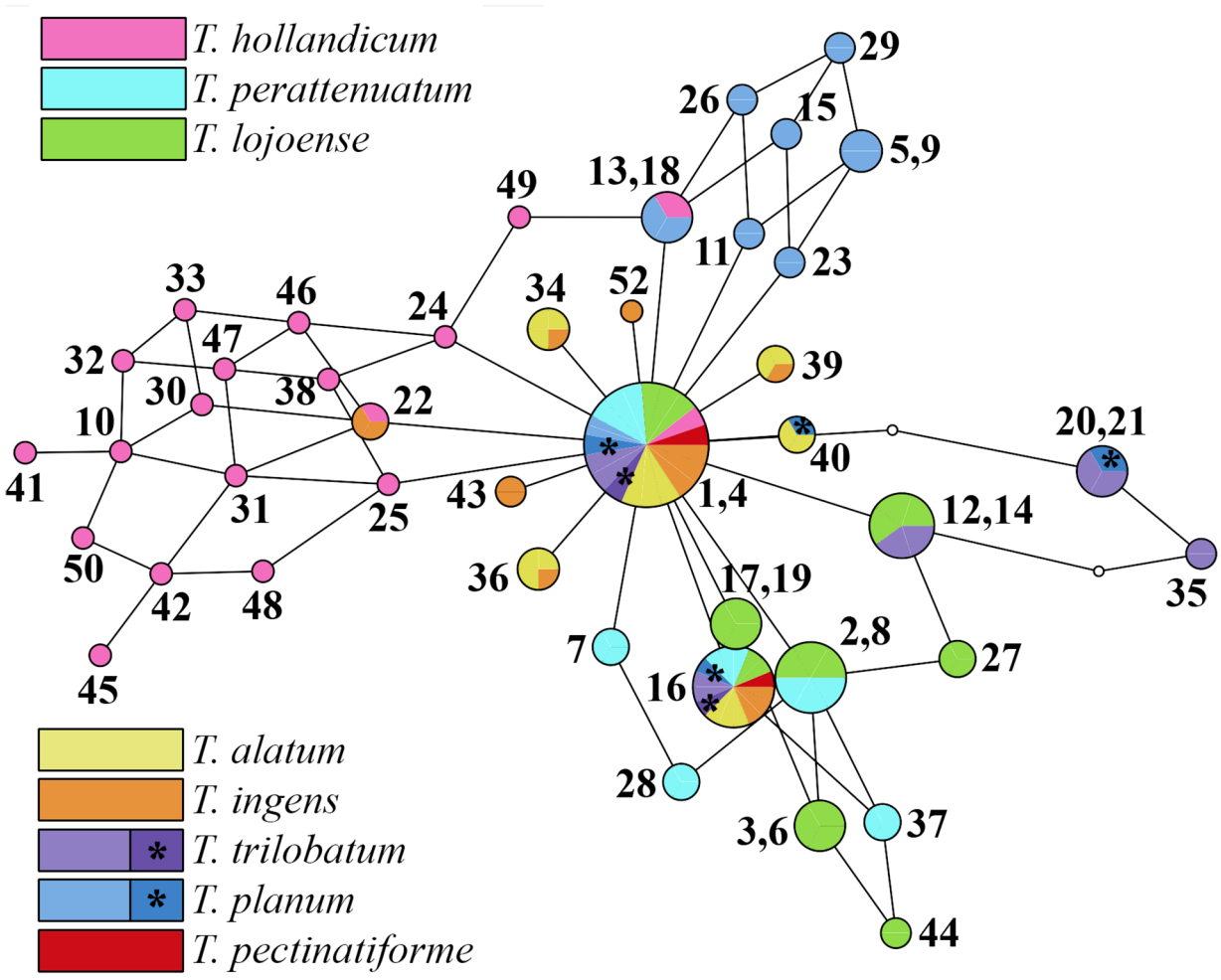

Fig. 3. Network of rDNA variants (v1–v52, excluding v51) of studied *Taraxacum* specimens. Numbers of ITS1 region variants correspond to Table 2. Asterisks (*) denote variants of the specimens which were possibly misdetermined.

Table 2

Variants of ITS1 region of rDNA repeats of the studied *Taraxacum* samples

	T. ingens			***T. alatum***			***T. trilobatum***			***T. planum***			***T. perattenuatum***			***T. lojoense***			***T. hollandicum***	***T. pectinatiforme***
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
v1	47.57	50.32	48.76	49.36	41.94	52.87	49.39	37.54	33.02	44.53	2.06	2.52	23.79	22.60	22.30	16.83	16.80	16.06	6.94	52.00
v2	-	-	-	-	-	-	-	-	-	-	-	-	18.48	21.77	21.17	8.22	7.37	4.60	-	-
v3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	15.01	16.26	14.75	-	-
v4	3.00	3.24	2.94	3.20	3.21	3.04	3.34	2.67	2.53	2.94	-	-	1.77	1.84	1.67	1.11	1.19	1.27	0.57	3.14
v5	-	-	-	-	-	-	-	-	-	-	20.21	20.84	-	-	-	-	-	-	-	-
v6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9.38	9.94	13.32	-	-
v7	-	-	-	-	-	-	-	-	-	-	-	-	11.86	8.44	9.58	-	-	-	-	-
v8	-	-	-	-	-	-	-	-	-	-	-	-	6.24	7.75	7.71	2.65	2.58	2.46	-	-
v9	-	-	-	-	-	-	-	-	-	-	14.31	12.81	-	-	-	-	-	-	-	-
v10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	25.77	-
v11	-	-	-	-	-	-	-	-	-	-	11.99	12.96	-	-	-	-	-	-	-	-
v12	-	-	-	-	-	-	-	6.64	6.41	-	-	-	-	-	-	3.66	3.38	2.70	-	-
v13	-	-	-	-	-	-	-	-	-	-	10.22	10.33	-	-	-	-	-	-	2.01	-
v14	-	-	-	-	-	-	-	5.69	7.13	-	-	-	-	-	-	2.71	2.82	3.19	-	-
v15	-	-	-	-	-	-	-	-	-	-	10.39	10.85	-	-	-	-	-	-	-	-
v16	1.40	1.68	1.66	1.68	1.73	2.03	1.76	1.25	1.32	1.43	-	-	0.68	0.83	0.78	0.51	-	0.58	-	1.55
v17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	8.01	7.02	4.22	-	-
v18	-	-	-	-	-	-	-	-	-	-	8.96	7.41	-	-	-	-	-	-	2.40	-
v19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6.40	6.10	5.54	-	-
v20	-	-	-	-	-	-	-	5.62	4.71	7.08	-	-	-	-	-	-	-	-	-	-
v21	-	-	-	-	-	-	-	4.57	5.11	5.48	-	-	-	-	-	-	-	-	-	-
v22	6.82	-	6.73	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.43	-
v23	-	-	-	-	-	-	-	-	-	-	6.06	6.53	-	-	-	-	-	-	-	-
v24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	11.16	-
v25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	7.65	-
v26	-	-	-	-	-	-	-	-	-	-	3.46	3.62	-	-	-	-	-	-	-	-
v27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2.78	2.52	1.69	-	-
v28	-	-	-	-	-	-	-	-	-	-	-	-	2.32	1.82	2.28	-	-	-	-	-
v29	-	-	-	-	-	-	-	-	-	-	2.64	2.92	-	-	-	-	-	-	-	-
v30	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.42	-

Table 2 (continued)

	T. ingens			*T. alatum*			*T. trilobatum*			*T. planum*			*T. perattenuatum*			*T. lojoense*			*T. hollandicum*	*T. pectinatiforme*
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
v31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.14	-
v32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4.11	-
v33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.37	-
v34	-	0.85	-	0.72	0.79	0.69	-	-	-	-	-	-	-	-	-	-	-	-	-	-
v35	-	-	-	-	-	-	-	1.13	0.98	-	-	-	-	-	-	-	-	-	-	-
v36	-	0.51	-	0.50	0.51	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-
v37	-	-	-	-	-	-	-	-	-	-	-	-	0.56	0.71	0.64	-	-	-	-	-
v38	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.87	-
v39	-	0.58	-	0.61	0.55	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
v40	-	-	-	0.50	0.50	-	-	-	-	0.60	-	-	-	-	-	-	-	-	-	-
v41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.44	-
v42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.36	-
v43	0.55	-	0.59	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
v44	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.52	0.51	-	-
v45	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.97	-
v46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.90	-
v47	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.80	-
v48	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.77	-
v49	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.76	-
v50	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.72	-
v51	-	-	-	-	-	-	-	-	-	-	0.57	-	-	-	-	-	-	-	-	-
v52	0.51	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Discussion

Targeted NGS of ITS1 region as a source of phylogenetic information

Specimens sampled as separate sections were substantially different from each other and usually had many (not less than three) specific variants of ITS1 region, often at high frequencies. Thus the screened polymorphism didn't contradict to the preliminary sectional-level determinations, evidencing that the targeted NGS of ITS1 region is a prospective method for the taxonomic and phylogenetic studies in *Taraxacum*. However, the data acquired are not directly applicable for sectional-level taxonomic and phylogenetic issues due to poor sampling, and a larger screening of *Taraxacum* from North and North-West European Russia is planned for that purpose in the future.

The utility of targeted NGS of ITS1 region for intrasectional classification and phylogeny was studied only in the case of the section *Taraxacum*, which is the largest and very complicated group in the genus (e. g. Kirschner et al., 2022). Among 13 specimens sampled as 5 different species, not less than three groups have been demonstrated. Such strong intrasectional polymorphism may be associated with heterogeneity of this group. Ribosomal DNA diversity in apomicts is generally explained by inherited variation from past hybridization events rather than accumulated with apomixis (Wang et al., 2023). As the whole genus is known for complex hybridity, it is possible that parts of the section *Taraxacum* have somewhat different hybrid origin and possibly can be classified even in several different sections.

It is worth mentioning in this respect that the two well-known segregates of the section *Taraxacum*, sections *Hamata* H. Øllg. and *Borea*, were relatively recently described (Øllgaard, 1983; Richards, 1985). Our data confirm correctness of such splitting for the section *Borea*, which proved to be substantially different in present analysis from all other specimens. Moreover, our results argue that there may be more groups that ought to be segregated from the section *Taraxacum* like the section *Borea*. Study of the heterogeneity of the section *Taraxacum* may be a prioritized aim of the further sampling.

In other words, some of the 'informal groups' in the section *Taraxacum*, yet claimed to be only morphogroups for practical purpose for use in determination keys by I. Uhlemann (Uhlemann, 2003; Uhlemann et al., 2016) and J. Richards (2021), may also exist as real phylogenetic entities. However, one association of the specimens assembled in the current study, 'core' section *Taraxacum*, includes taxa from different 'informal groups'. This may be either due to the limited resolution abilities of the method applied, or may be alternatively explained by another factor, the concerted DNA evolution in older taxa (see further). Záveská Drábková et al. (2009: 83) explained total identity of ITS region in different agamosperm *Taraxacum* species only as an evidence of the shared parentage, which may be not so simple in reality.

Concerted evolution of rDNA repeats in *Taraxacum*

It is a widely discussed topic that apomicts have lower frequencies of concerted evolution due to a reduced rate of meiotic crossovers (Machackova et al., 2020; Wang et al., 2023). This was confirmed in *Taraxacum* in an explicit study of intra- and between-individual diversity of ITS repeats by Záveská Drábková et al. (2009) and Machackova et al. (2020) and was similarly shown in some earlier studies of the genus (Kirschner et al., 2003, 2015).

Our results tell that in the apomictic 'core' section *Taraxacum*, which in our study included 6–8 from 13 samples of this section, ITS region may be substantially homogenized by concerted evolution. Here, dominant ITS1 variant was (7×)12×–17× more common than any other variant. The 'core' consisted of taxa from 2–3 morphogroups (as outlined by I. Uhlemann and J. Richards). Therefore, the 'core' may be not necessarily an association of closely related taxa but may comprise taxa from various lineages, brought together by their more advanced age. It is noteworthy that *T. alatum* and *T. pectinatiforme* are considered rather well-defined taxa with wide distribution, and *T. alatum* was even included in the sampling by Kirschner et al. (2016), as a taxon consistently recognized by many specialists. Good morphological differentiation of *T. alatum* and *T. pectinatiforme* may correlate with their advanced age, which, in its turn, contributes to stronger homogeneity of rDNA repeats. If so, homogenized rDNA repeats are a possible indicator of older, better stabilized species, at least in section *Taraxacum*. However, for more robust decisions in this subject, deeper sampling is needed.

Somewhat different results observed by Záveská Drábková et al. (2009) and Machackova et al. (2020) may be due to largely different taxonomic groups they focused on, which may have somewhat different evolutional patterns. Unlike our study, they investigated mainly (or only) sections other than

Taraxacum. Uhlemann et al. (2009: 41) reported invariable rDNA repeats that didn't need cloning for standard Sanger sequencing in section *Taraxacum*, which corresponds to the results presented here.

Other issues

The phylogenetic tree of obtained sequences (not shown) was uninformative, due to very small differences between phylogenetically informative variants, not enabling to resolve their relative phylogenetic position to each other.

The present study gives evidence that 2 of 20 samples in our study may have been misidentified, with high probability: one of them was sampled as '*T. trilobatum*', another as '*T. planum*'. This species pair (along with species pair *T. alatum* / *T. ingens*) was included in the analysis having in mind that misdetermination is possible (see 'selection of specimens for analysis').

In this study, we observed some minor ITS1 variants with large deletions, 36 or 133 bp. As far as no variants with similar deletions were found in Genbank, and because such large deletions should have important impact on functionality, we treat such variants as pseudogenes. Putative pseudogenes are discovered here for the first time in rDNA units of *Taraxacum*; earlier (Záveská Drábková et al., 2009; Machackova et al., 2020) they were missed probably due to limited genome screening done without second-generation sequencing approaches.

In relation to the results presented in the current contribution, the desirable next step is large-scale screening of *Taraxacum* with an emphasis on the groups present in North and North-West Russia.

Conclusions

1. The intragenomic polymorphism of ITS1 region in rDNA repeats of *Taraxacum* is a prospective source of the data for phylogenetic reconstructions. It may be used at the level of sections and species aggregates, and for disentangling recent hybridization events between distantly related taxa.

2. Section *Taraxacum* may be not monomorphic and possibly consists of separate groups of species.

3. Concerted evolution of nrDNA in apomictic polyploids from the section *Taraxacum* was found in more than one half of the studied specimens from this section and may be a common phenomenon. Hypothetically, homogenized rDNA repeats in this group may serve as an indicator of older, better stabilized taxa.

4. Putative pseudogenized rDNA repeats with large deletions were observed in low frequencies, which is the first case of their identification in *Taraxacum*.

Acknowledgements

The study was supported by the Ministry of Education and Science of the Russian Federation, grant agreement No. 20-04-00561075-12-2021-1056 from 28.09.2021. DNA extraction, amplification and sequencing were carried out using the equipment of the resource center "Genomic Technologies, Proteomics and Cell Biology" of ARRIAM and resource center of Komarov Botanical Institute.

REFERENCES / ЛИТЕРАТУРА

Alsos I. G., Lavergne S., Merkel M. K. F., Boleda M., Lammers Y., Alberti A., Pouchon C., Denoeud F., Pitelkova I., Pușcaș M., Roquet C., Hurdu B.-I., Thuiller W., Zimmermann N. E., Hollingsworth P. M., Coissac E. 2020. The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. *Plants* 9(4): 432. DOI: 10.3390/plants9040432

Clement M., Posada D., Crandall K. A. 2000. TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* 9: 1657–1660.

Doyle J. J., Doyle J. L. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19: 11–15.

Edgar R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461. DOI: 10.1093/bioinformatics/btq461.

Katoh K., Standley D. M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30(4): 772–780. DOI: 10.1093/molbev/mst010

Kirschner J., Oplaat C., Verhoeven K. J. F., Zeisek V., Uhlemann I., Trávníček B., Räsänen J., Wilschut R. A., Štěpánek J. 2016. Identification of oligoclonal agamospermous microspecies: taxonomic specialists versus microsatellites. *Preslia* 88(1): 1–17.

Kirschner J., Štěpánek J., Klimeš L., Dvorský M., Brůna J., Macek M., Kopecký M. 2020. The *Taraxacum* Flora of Ladakh, with notes on the adjacent regions of the West Himalaya. *Phytotaxa* 457(1): 1–409. DOI: 10.11646/phytotaxa.457.1.1

Kirschner J., Štěpánek J., Mes T. H. M., den Nijs J. C. M., Oosterveld P., Štorchová H., Kuperus P. 2003. Principal features of the cpDNA evolution in *Taraxacum* (Asteraceae, Lactuceae): a conflict with taxonomy. *Plant Syst. Evol.* 239: 231–255. DOI: 10.1007/s00606-003-0002-5

Kirschner J., Záveská Drábková L., Štěpánek J., Uhlemann I. 2015. Towards a better understanding of the *Taraxacum* evolution (Compositae – Cichorieae) on the basis of nrDNA of sexually reproducing species. *Plant Syst. Evol.* 301: 1135–1156. DOI: 10.1007/s00606-014-1139-0

Kumar S., Stecher G., Li M., Knyaz C., Tamura K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* 35: 1547–1549. DOI: 10.1093/molbev/msy096

Letunic I., Bork P. 2021. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucl. Acids Res.* 49: W293–W296. DOI: 10.1093/nar/gkab301

Lundevall C. F., Øllgaard H. 1999. The genus *Taraxacum* in the Nordic and Baltic countries: types of all specific, subspecific and varietal taxa, including type locations and sectional belonging. *Preslia* 71(1–2): 43–171.

Machackova P., Majeský L., Hrones M., Bilkova L., Hribova E., Vasut R. J. 2022. New insights into ribosomal DNA variation in apomictic and sexual *Taraxacum* (Asteraceae). *Bot. J. Linn. Soc.* 199(4): 790-815.

Majeský L., Vašut R. J., Kitner M. 2015. Genotypic diversity of apomictic microspecies of the *Taraxacum scanicum* group (*Taraxacum* sect. *Erythrosperma*). *Plant Syst. Evol.* 301: 2105–2124. DOI: 10.1007/s00606-015-1218-x

Mes T. H. M., Kuperus P., Kirschner J., Štěpánek J., Štorchová H., Oosterveld P., den Nijs J. C. M. 2002. Detection of genetically divergent clone mates in apomictic dandelions. *Molec. Ecol.* 11: 253–265.

Múrias dos Santos A., Cabezas M. P., Tavares A. I., Xavier R., Branco M. 2016. tcsBU: A tool to extend TCS network layout and visualization. *Bioinformatics* 32: 627–628.

Nelson J. O., Watase G. J., Warsinger-Pepe N., Yamashita Y. M. 2019. Mechanisms of rDNA copy number maintenance. *Trends Genet.* 35(10): 734–742. DOI: 10.1016/j.tig.2019.07.006

Nguyen L.-T., Schmidt H. A., von Haeseler A., Minh B. Q. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* 32: 268–274. DOI: 10.1093/molbev/msu300

Øllgaard H. 1983. *Hamata*, a new section of *Taraxacum* (Asteraceae). *Plant Syst. Evol.* 141: 199–217.

Richards A. J. 1985. Sectional nomenclature in *Taraxacum* (Asteraceae). *Taxon* 34: 633–644.

Richards A. J. 2021. *Field Handbook to British and Irish Dandelions*. B.S.B.I. Handbook no. 23. Durham: Botanical Society of Britain and Ireland. 302 pp.

Rideout J. R., Caporaso G., Bolyen E., McDonald D., Vázquez Baeza Y., Alastuey J. C., et al. 2023. *Biocore/scikit-bio: scikit-bio 0.5.9: Maintenance release*. URL: https://zenodo.org/records/8209901 (Accessed 25 March 2024).

Ridgway K. P., Duck J. M., Young J. P. 2003. Identification of roots from grass swards using PCR-RFLP and FFLP of the plastid *trnL* (UAA) intron. *BMC Ecol.* 3: 8. DOI: 10.1186/1472-6785-3-8

Salih R. H. M., Majeský L., Schwarzacher T., Gornall R. Heslop-Harrison P. 2017. Complete chloroplast genomes from apomictic *Taraxacum* (Asteraceae): Identity and variation between three microspecies. *PLoS ONE* 12(2): e0168008. DOI: 10.1371/journal.pone.0168008

Sennikov A. N. 2007. The genus *Taraxacum* (Asteraceae) in the Russian part of East Fennoscandia: distribution, new records and comparison of taxonomic concepts. *Memoranda Soc. Fauna Fl. Fenn.* 83: 59–81.

Trávníček B., Kirschner J., Štěpánek J., Vašut R. J. 2010. *Taraxacum* Wiggers – pampeliška (smetánka). In: *Květena České republiky*. Vol. 8. Prague: Academia. Pp. 23–269. [In Czech].

Tzvelev N. N. 1989. *Taraxacum* Wigg. In: *Flora partis europaeae USSR*. T. 8. Leningrad: Nauka. Pp. 61–114. [In Russian] (***Цвелёв Н. Н.*** Taraxacum Wigg. // Флора европейской части СССР. Т. 8. Л.: Наука, 1989. С. 61–114).

Uhlemann I. 2003. Die Gattung *Taraxacum* (Asteraceae) im östlichen Deutschland. *Mitt. Florist. Kart. Sachsen-Anhalt,* Sonderheft: 1–136.

Uhlemann I., Kirschner J., Štěpánek J. 2016. *Taraxacum*. In: Rothmaler – Exkursionsflora von Deutschland. Gefäßpflanzen: Kritischer Ergänzungsband. Aufl. 11. Berlin, Heidelberg: Springer. Pp. 133–184.

Uhlemann I., Ritz C. M., Peñailillo P. 2009. Relationships in *Taraxacum* section *Arctica* s. l. (Asteraceae, *Cichorieae*) and allies based on nrITS. *Feddes Repertorium* 120(1–2): 35–47. DOI: 10.1002/fedr.200811193

Wang W., Zhang X., Garcia S., Leitch A. R., Kovařík A. 2023. Intragenomic rDNA variation – the product of concerted evolution, mutation, or something in between? *Heredity* 131: 179–188. DOI: 10.1038/s41437-023-00634-5

White T. J., Bruns T. D., Lee S. B., Taylor J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In: M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White (eds.). *PCR Protocols: A Guide to Methods and Applications*. New York: Academic Press. Pp. 315–322. DOI: 10.1016/B978-0-12-372180-8.50042-1

Wolanin M., Klichowska E., Jedrzejczyk I., Rewers M., Nobis M. 2023. Taxonomy and distribution of *Taraxacum* sect. *Erythrosperma* (Asteraceae) in Poland. *PhytoKeys* 224: 1–88. DOI: 10.3897/phytokeys.224.99463

Záveská Drábková L., Kirschner J., Štěpánek J., Záveský L., Vlček Č. 2009. Analysis of nrDNA polymorphism in closely related diploid sexual, tetraploid sexual and polyploid agamospermous species. *Plant Syst. Evol.* 278: 67–85. DOI: 10.1007/s00606-008-0134-8