

RESEARCH ARTICLE

Genetic diversity and relationship between cultivated, weedy and wild rye species as revealed by chloroplast and mitochondrial DNA non-coding regions analysis

Lidia Skuza^{1,2*}, Izabela Szućko^{1,2}, Ewa Filip^{1,2}, Tomasz Strzała³

1 Department of Molecular Biology and Cytology, The Institute for Research on Biodiversity, Faculty of Biology, University of Szczecin, Szczecin, Poland, **2** The Centre for Molecular Biology and Biotechnology, Faculty of Biology, University of Szczecin, Szczecin, Poland, **3** Department of Genetics, Faculty of Biology and Animal Science, Wrocław University of Environmental and Life Sciences, Wrocław, Poland

* lidia.skuza@usz.edu.pl



OPEN ACCESS

Citation: Skuza L, Szućko I, Filip E, Strzała T (2019) Genetic diversity and relationship between cultivated, weedy and wild rye species as revealed by chloroplast and mitochondrial DNA non-coding regions analysis. PLoS ONE 14(2): e0213023. <https://doi.org/10.1371/journal.pone.0213023>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: September 25, 2018

Accepted: February 13, 2019

Published: February 27, 2019

Copyright: © 2019 Skuza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by the State Committee for Scientific Research grant No. N N310 435498 "Degree of Relatedness Within the Genus *Secale* Using non-coding Chloroplast and Mitochondrial Sequences, and Nuclear rDNA IGS Regions".

Competing interests: The authors have declared that no competing interests exist.

Abstract

The genus *Secale* is small but very diverse. Despite the high economic importance, phylogenetic relationships of rye species have not been fully determined, and they are extremely important for the process of breeding of new cultivars that can be enriched with functional traits derived from wild rye species. The study analyzed the degree of relationship of 35 accessions of the genus *Secale*, representing 13 most often distinguished species and subspecies, originating from various seed collections in the world, based on the analysis of non-coding regions of the chloroplast (cpDNA) and mitochondrial genome (mtDNA), widely used in phylogenetic and population plant studies, because of a higher rate of evolution than the coding regions. There was no clear genetic structure between different species and subspecies, which may indicated the introgression between these taxa. The obtained data confirmed that *S. vavilovii* was very similar to *S. cereale*, which confirmed the assumption that they might share a common ancestor. The results also confirmed the divergence of *S. sylvestre* from other species and subspecies of rye. Areas that may be useful molecular markers in studies on closely related species of the genus *Secale* were also indicated.

Introduction

Rye is a difficult object of genetic and breeding studies. The reason is the open-pollination, self-incompatibility and the relationship between heterozygosity and productivity, which arises as a result of inter-chromosomal gene interactions [1]. Rye has the largest genome ~ 7.9 Gbp among all diploid Triticeae [2], which is in 90% consisting of repetitive sequences. The genus *Secale* is also very diverse—it includes annual, perennial, self-pollinating and open-pollinating species of various morphologies. The classification system of the American Germplasm Resources Information Network (GRIN, <http://www.arsgrin.gov>) currently includes four species to the genus *Secale*: annual *S. cereale* L., annual *S. sylvestre* Host and *S. vavilovii* Grossh

and perennial *S. strictum* (Presl.) Presl. (syn. *S. montanum*) [3,4]. All species within the genus *Secale* are diploid with 14 chromosomes, they cross with each other easily, and crossbreeding results in partially fertile hybrids [5,6]. Only *Secale sylvestre* has different characteristics (low crossability with other species, the lowest amount of t-heterochromatin [5] and the smallest genome (7.23 pg)[7] and is probably the most distant species [5,8], as indicated by many research results [9–15]. Rye (*Secale cereale* L.) is also an important and rich source of valuable genes encoding, e.g., high protein content, resistance to diseases as well as morphological and biochemical traits for wheat and a synthetic wheat–rye hybrid triticale (\times *Triticosecale* Wittmack) improvement.

Rye has an advantage over other cereals in temperate climates because of its excellent tolerance to low temperatures, high levels of soil alumina and its ability to produce acceptable grain yields under drought stress conditions [16]. Rye grows on sandy or acidic soils or on poorly prepared terrain. It is widely adapted, but grows best in cold temperature areas. The colder climate in Central and Eastern Europe, which are the main rye growing regions in the world (<http://www.fao.org/faostat>), favors the high frost tolerance of rye and its ability to grow on poor soils [17]. Unfortunately, despite the economic importance, phylogenetic relationships within the genus *Secale* remain unclear. The reason for this may be the postulated multi-paths in the evolution of rye [11]. There is therefore a need to search for different verification methods for the proposed classification system and phylogenetic relationships.

The differences in the classification of the genus *Secale* may result from the use of various experimental methods, as described in detail previously [18]. Establishing phylogenetic relationships is extremely important for the process of breeding new cultivars, which can be enriched with functional traits derived from wild rye species, e.g., resistance to downy mildew and brown rust, resistance to lodging and pre-harvest sprouting and cms [19,20]. Phylogenetic and population studies often use genetic information contained in chloroplast (cpDNA) and mitochondrial DNA (mtDNA), as they contain two necessary sets of genes in plants [21–23]. Firstly, they encode many key proteins for basic cell bioenergy processes. Secondly, they encode many components necessary for the proper expression of their own genes. Considering the essential importance of these two sets of organelle genes, one can expect them to change very slowly during evolution [21–23]. In addition, organelle DNA is a useful tool in the search for species-specific molecular markers. However, phylogenetic studies of the genus *Secale* were almost exclusively based on the analysis of the nuclear and chloroplast genomes [18,24–34]. So far, two works have been published that include mtDNA: Isik et al. [33] and Skuza et al. [34]. None of the studies included the analysis of highly-variable non-coding sequences in organelle genomes, which due to the rapid evolution and accumulation of a higher number of deletions/insertions or substitutions than coding regions are considered very good markers when testing at the intraspecific level [35].

Recent studies have successfully proven that some parts of chloroplast genomes were more effective in illuminating phylogeny of land plants than cpDNA sequences frequently used in phylogenetic studies [36–39]. It is then interesting whether the non-coding regions of the chloroplast (cpDNA) and mitochondrial genomes (mtDNA) will confirm this dependence, all the more so as the mitochondrial genome of rye has not been sequenced so far, and complete data only on the *S. cereale* chloroplast genome sequence has been published in 2014 [40].

The current work analyzed the relationship of 35 accessions of the genus *Secale*, representing 13 most often distinguished species and subspecies, originating from different seed collections in the world. Twelve universal primers were selected for analysis [41–44], amplifying the non-coding regions of the chloroplast (cpDNA) and mitochondrial genomes (mtDNA), broadly used in phylogenetic and plant population studies [21–23]. The use of such consensus primers that are homologous to most coding regions, but amplify variable non-coding regions are very useful for phylogenetic and population studies [43–45].

The research was aimed at understanding the genetic diversity within the genus *Secale*, as well as the possibility of revising existing taxonomic classifications.

Materials and methods

The plant material consisted of 35 accessions of the genus *Secale*, 13 cultivated and non-cultivated species and subspecies of rye, obtained from several world collections (Center for Biological Diversity Conservation in Powsin–Warsaw, Poland; United States Department of Agriculture–Agricultural Research Service, USA; Nordic Genetic Resource Center, Sweden). The list of species, along with the accession numbers for each sample, is given in [S1 Table](#).

DNA extraction, PCR amplification, and DNA sequencing

Genomic DNA was extracted from coleoptiles (0.2 g per sample) of etiolated 6- to 7-day-old seedlings using a FastDNA Green SPIN Kit (DNAeasy Plant Mini Kit–Wizard Genomic DNA Promega), according to the manufacturer's instructions. The isolated DNA was subjected to quantitative and qualitative evaluation by measuring the absorbance at 230, 260, and 280 nm using a NanoDrop 2000c spectrophotometer (Thermo Scientific, Madison, USA). DNA isolation was performed at least three times.

PCR amplification of chloroplast (cp) non-coding (intron) DNA region. For the amplification of the cpDNA 6 non-coding (intron) regions, genomic DNA was isolated as described previously and the concentration was adjusted to ~50 ng/μl. Primer sequence list (5'-3') for the amplification of different cpDNA regions is provided in [S2 Table](#). PCR amplification was performed three times in a 25-μl reaction mixture containing approximately 50–150 ng genomic DNA template, 2.0–3.0 mM MgCl₂, 0.2–1.0 mM of each dNTP, 0.1–1 μM of each primer, 0.1 mg BSA/ml, and 1 U *Taq* DNA polymerase. Specifics of reaction conditions and components for the amplification of each region are given in [S4](#) and [S6](#) Tables.

PCR amplification of mitochondrial (mt) non-coding (intron) DNA region. For the amplification of the mtDNA 6 non-coding (intron) regions, genomic DNA was isolated as described previously and the concentration was adjusted to ~50 ng/μl. Primer sequence list (5'-3') for the amplification of different cpDNA regions is given in [S3 Table](#). PCR amplification was performed three times in a 25-μl reaction mixture containing approximately 75–250 ng genomic DNA template, 2.5–4.0 mM MgCl₂, 0.1–0.2 mM of each dNTP, 0.2–0.55 μM of each primer, 0.05 mg BSA/ml, and 1–1.5 U *Taq* DNA polymerase. Specifics of reaction conditions and components for the amplification of each analyzed region are presented in [S5](#) and [S7](#) Tables.

Electrophoresis. The products were separated on a 1.5% (m/v) agarose gel in a 1× TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM Na₂EDTA, pH 8.3). Ethidium bromide was added to the gel to a final concentration of 0.1 μg/ml. Electrophoresis was carried out in a Sub-Cell Model 192 PowerPac HV electrophoresis system (Bio-Rad) in a 1× TBE buffer at 100 V for approximately 1.5 h. Gel images were captured with a Gel Doc XR system (Bio-Rad). The bands were scored and analyzed with the Quantity One software (Bio-Rad). The size of the products was determined by comparison with the DNA ladder (MassRuler DNA Ladder Mix, Fermentas).

Sequence and alignment for data analysis. The majority of PCR products were purified and directly sequenced at Macrogen (Seoul, Korea) and Genomed (Warsaw, Poland). Sequencing was performed two times in both forward and reverse directions. Pairwise alignments were made using the sequences obtained from the forward and reverse primers. Multiple sequence alignment was performed using ClustalW. Genetic diversity was estimated based on the identification of unique haplotypes using DnaSP 5.10. [46]. This software was also used

to calculate haplotype (Hd) and nucleotide (π_n). The sequences reported in this paper have been deposited in the NCBI Genbank nucleotide sequence database with the accession numbers MH893827-MH894176 (S1 Table).

Phylogenetic analysis

To infer phylogenetic relationship between analyzed species, we used concatenated sequences of the chloroplast genome (cpDNA) (alignment length– 6362 bp) and mitochondrial sequences (mtDNA) (alignment length– 4139 bp) for 37 individuals (including two sequences used as an outgroup). The dataset was analyzed as a partition consisting of two separate types of sequences (cpDNA and mtDNA) to allow different substitution models for both kinds of data.

Both partitions were first aligned separately with the MUSCLE algorithm [47] implemented in the Seaview software [48]. Next, sequences were cleaved to obtain equal alignment blocks and concatenated together (cpDNA and mtDNA) in one file.

Phylogenetic analyses were performed using two approaches: maximum likelihood (ML) and Bayesian inference (BI). ML analyses were carried out with PhyML [49,50] and RaxML 7.7.1 [51,52]. In PhyML, smart model selection function [53] was used with the Akaike Information Criterion along with five random starting trees and SPR type tree improvement. The bootstrap analysis (1000 replications) was used to test tree topology. In RaxML, the GTR +I +G substitution model was used as the best-fit model selected using jModelTest 2.1.10 [54] for both partitions. Hundred replications were used in the bootstrap analysis to test tree topology. MrBayes 3.2.6 [55] was used for the Bayesian tree, with a separate mixed +I +G [56] model for each partition, and the analysis was run with two independent runs (four chains each) starting from random trees. Trees were sampled every 100th generation of the Markov chain step for 25,000,000 generations with 25% burn-in. Final tree set consisted of trees probed when the average standard deviation between robots was stabilized significantly below 0.01.

All trees were visualized with FigTree [57] and compared with each other.

Distance between and within clades. We estimated distances between and within clades revealed on the tree, as we were interested in determining the relationships between species. The analyses were conducted using the 2-parameter Kimura model [58] with gamma distribution (shape parameter = 0.05) of the rate variation among sites and deletion of all positions with gaps, using MEGA X [59].

Results

Characteristics of cpDNA non-coding sequences

The size polymorphism of the PCR products obtained was not detected within the analyzed regions. Due to the difficulty in amplifying two regions *trnL* (UAA) 3' exon-*trnF* (GAA) and *trnK* [tRNA-Lys (UUU) exon 1] -*trnK* [tRNA-Lys (UUU) exon2], further analyses were based on the results obtained from the four regions (Table 1).

A total of 3,031 bp from 140 sequences were analyzed. The longest identified noncoding region in the chloroplast genome was *trnD* [tRNA-Asp (GUC)]-*trnT* [tRNA-Thr (GGU)]. Four regions: *atpB-rbcL*, *trnT*(UGU)-*trnL*(UAA)5' exon, *trnL*(UAA) intron and *trnD* [tRNA-Asp(GUC)]—*trnT*[tRNA-Thr(GGU)] showed differences on both intraspecific and interspecific levels. The most variable regions (*trnT* (UGU)-*trnL* (UAA) 5' exon and *trnD* [tRNA-Asp(GUC)]—*trnT*[tRNA-Thr(GGU)]) were selected, where the number of haplotypes was 12 and 16, respectively (Table 1).

Substitutions were the main source of variation, despite the fact that the studied regions also showed indels (Table 1). The number of identified indels ranged from 16 to 29 (*trnD*

Table 1. Characterization of nucleotide variability of analyzed plastid regions (cpDNA) in the type of *Secale*.

Nucleotide variability	^a Regions of cpDNA				Total	Mean
	1	2	3	4		
Length [bp]	815	600	653	963	3031	758
No. of Conserved sites [C]	671	248	490	889	3054	575
No. of Variable sites [V]	114	336	159	68	677	271
No. of Parsim-info [Pi]	7	31	59	42	139	81
No. of InDel site	17	16	19	29	81	32.40
Max. InDel size	14	1	2	1	-	2,61
No. of Substitutions sites	106	142	118	40	406	102
No. of Haplotype	8	12	5	16	41	8.40
Haplotype diversity [Hd]	0.408	0.565	0.341	0.765	-	0.520
Nucleotides diversity [Pi]	0.011	0.034	0.010	0.006	-	0.015

^aAnalyzed plastid regions (cpDNA): 1-*atpB-rbcL* intergenic spacer; 2-*trnT*(UGU)-*trnL*(UAA)5' exon intergenic spacer; 3- *trnL*(UAA) intron intergenic spacer; 4- *trnD* [tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] intergenic spacer.

<https://doi.org/10.1371/journal.pone.0213023.t001>

[tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] (Table 1). The size of indels ranged from 1 bp in regions *trnD*[tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] and *trnT*(UGU)-*trnL*(UAA)5' exon to 14 bp in the *atpB-rbcL* region (Table 1).

On average, 8.4 haplotypes were identified in the four analyzed non-coding regions (Table 1). The highest number of haplotypes was determined in two regions: *trnD*[tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] (16) and *trnT*(UGU)-*trnL*(UAA)5' exon (12). The most frequently occurring haplotype was haplotype 1 in the non-coding *atpB-rbcL* region (S8 Table). The haplotypes diversity ranged from 0.341 to 0.765 in the *trnD*[tRNA-Asp(GUC)]-*trnT* [tRNA-Thr(GGU)] region, whereas the nucleotide diversity ranged from 0.006 to 0.034 in the *trnT*(UGU)-*trnL*(UAA)5' exon region.

The BI, PhyML, and RaxML trees of all analyzed non-coding sequences of chloroplast DNA presented a similar topology with minor differences (Fig 1). The tree was not very informative and the topology of the tree was mainly polytomous, and one node connected all the analyzed sequences.

A few subgroups could be noticed within this group, formed mainly by *S. strictum* and *S. sylvestre* species. *S. africanum* (RSA1) formed a separate group (Fig 1). *S. vavilovii* was dispersed within the entire second group; only *S. vavilovii* Afghanistan and Russia showed a relative similarity. *S. cereale* species also did not form groups consistent with the current classification.

Noteworthy are the results obtained for *S. strictum* ssp. *africanum*. Two samples originating from South Africa were analyzed. The results clearly suggested that the origin of these seeds was not homogeneous—the second sample (RSA1) formed a separate subgroup (Fig 1).

Sequence similarity, depending on the origin, was not consistent. Only *S. sylvestre* species (Poland, Bulgaria, Russia) formed a monophyletic clade.

Characteristics of mtDNA non-coding sequences

Products were obtained for all analyzed mtDNA regions, as a result of the amplification (Table 2). The size polymorphism of the PCR products obtained was not detected within the analyzed regions.

A total of 4,171 bp from 210 sequences were analyzed. The longest identified non-coding region in the mitochondrial genome was: *nad4/1-2* (1642 bp) and it showed the highest polymorphism within the mtDNA noncoding regions of genus *Secale* (Table 2).

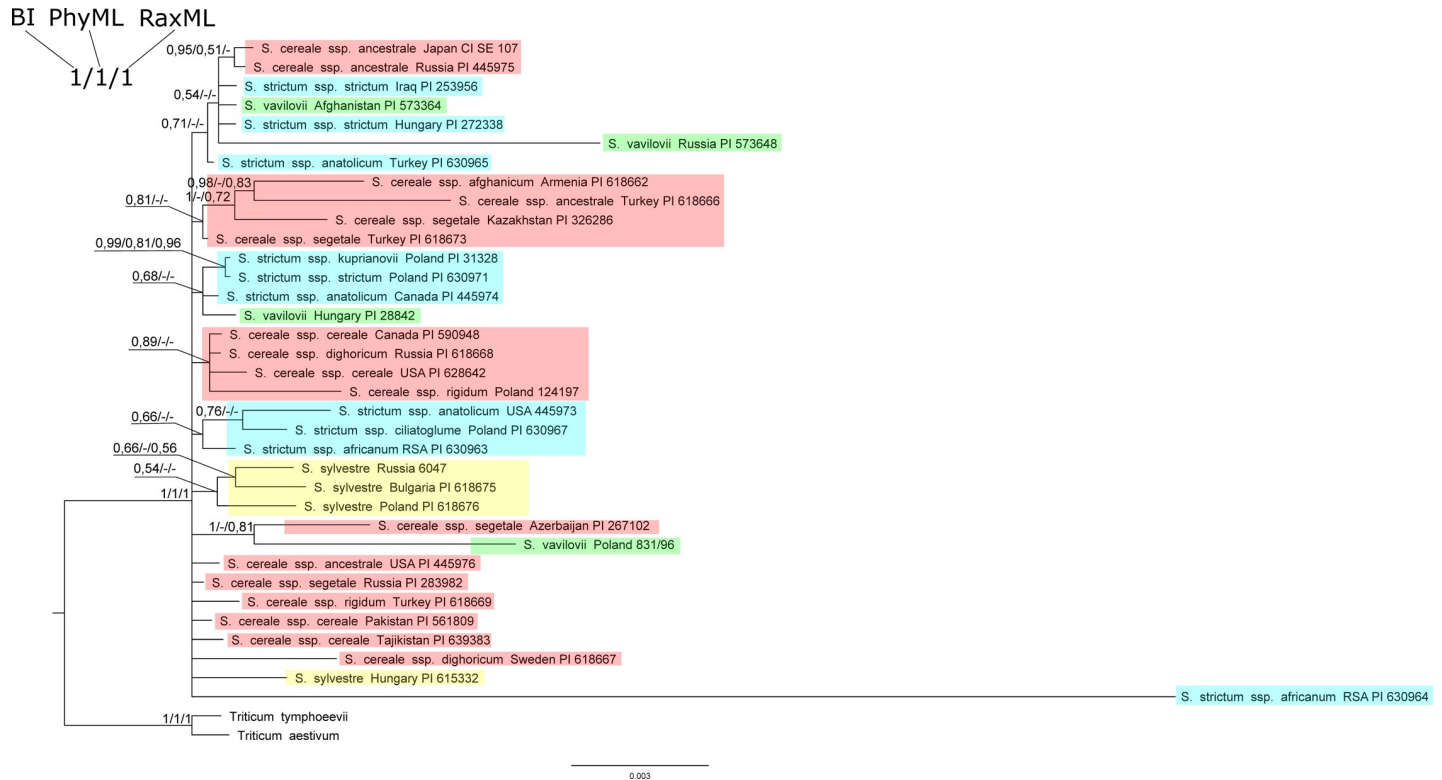


Fig 1. Bayesian tree based on the cpDNA noncoding sequences. Bootstrap values (RaxML and PhyML) and posterior probability values are presented along nodes (values below 0.5 are not presented).

<https://doi.org/10.1371/journal.pone.0213023.g001>

Two regions, *nad4/1-2* and *rps12-nad3(1)*, showed variation at both intraspecific and inter-specific levels. The remaining regions were conserved.

The number of haplotypes in the variable *nad4/1-2* and *rps12-nad3(2)* regions was 12 and 2, respectively (Table 2). Although the studied regions also contained indels, substitutions were the main source of variation. The number of identified indels ranged from 0 to 40 (in

Table 2. Characterization of nucleotide variability of analyzed mitochondrial regions (mtDNA) in the type of Secale.

Nucleotide variability	^a Regions of mtDNA						Total	Mean
	1	2	3	4	5	6		
Length [bp]	1483	1642	626	134	120	166	4171	695
No. of Conserved sites [C]	1438	1088	626	63	120	166	3546	591
No. of Variable sites [V]	0	509	0	71	0	0	580	97
No. of Parsim-info [Pi]	0	106	0	71	0	0	177	30
No. of InDel site	0	40	0	5	0	0	45	8
Max. InDel size	0	5	0	1	0	0	-	4.5
No. of Substitutions sites	0	222	0	70	0	0	292	49
No. of Haplotype	1	12	1	2	1	1	18	3
Haplotype diversity [Hd]	0.000	0.867	0.000	0.329	0.000	0.000	-	0.199
Nucleotides diversity [Pi]	0.000	0.029	0.000	0.186	0.000	0.000	-	0.036

^aAnalyzed mitochondrial regions mtDNA: 1- *nad1*exon B-*nad1*exon C intron intergenic spacer; 2- *nad4/1-2* intergenic spacer; 3- *nad4L-orf25* intergenic spacer; 4- *rps12-1/nad3-2* intergenic spacer; 5- *rps12-2/nad3-1* intergenic spacer; 6- *rrn5/rrn18-1* intergenic spacer.

<https://doi.org/10.1371/journal.pone.0213023.t002>

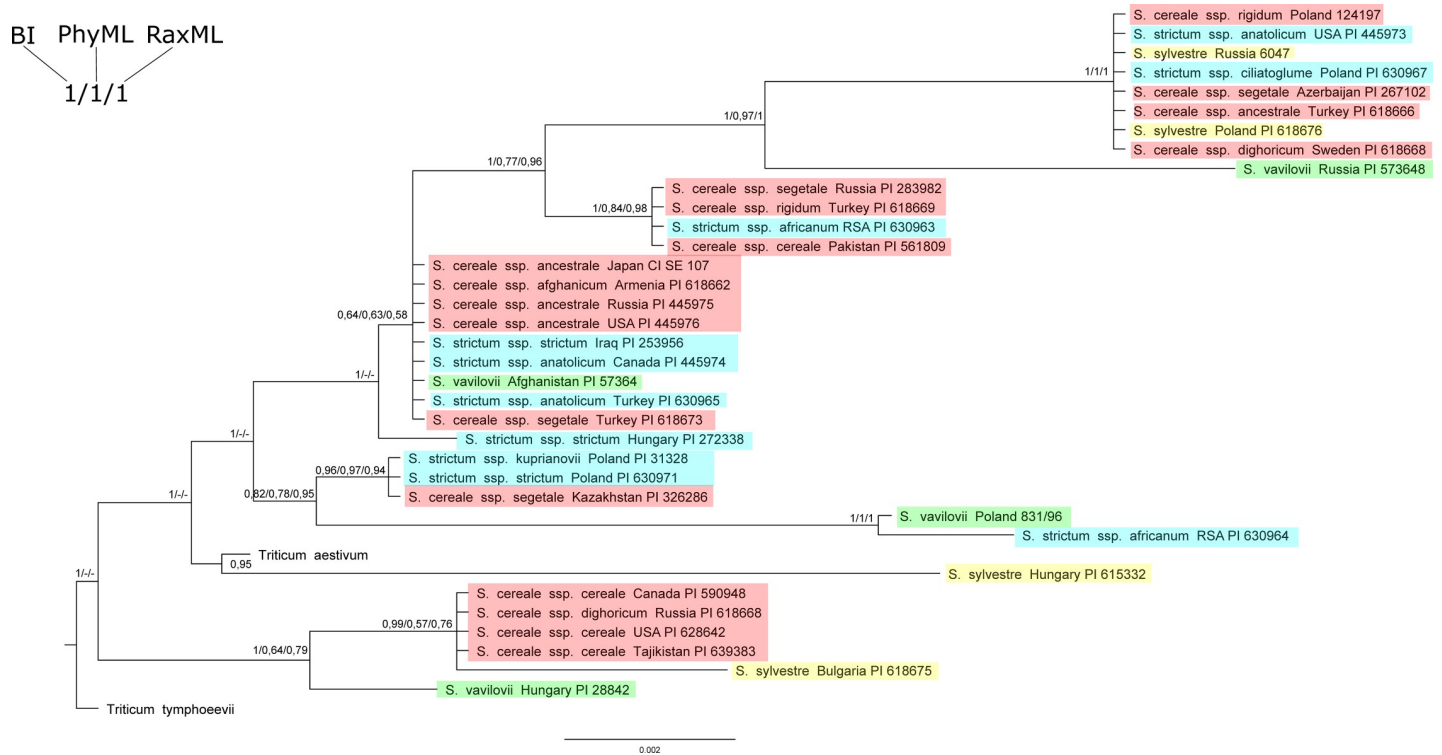


Fig 2. Bayesian tree based on the mpDNA noncoding sequences. Bootstrap values (RaxML and PhyML) and posterior probability values are presented along nodes (values below 0.5 are not presented).

<https://doi.org/10.1371/journal.pone.0213023.g002>

nad4/1-2) (Table 2). The size of indels ranged from 1 bp in the *rps12-nad3(2)* regions to 5 bp in the *nad4/1-2* region (Table 2).

On average, 3 haplotypes were identified among the six analyzed sequences (Table 2). Most haplotypes (12) were detected in the *nad4/1-2* region. The most frequently occurring was haplotype 1 in the non-coding *rps12-nad3(2)* region (S9 Table). The haplotypes diversity ranged from 0 (conserved regions) to 0.867 in the *nad4/1-2* region, while the nucleotides diversity ranged from 0 to 0.186 in *rps12-nad3(2)* (Table 2).

The BI, PhyML, and RaxML trees of all analyzed non-coding sequences of mitochondrial DNA presented a similar topology with slight difference (Fig 2). The tree was polytomous, but only at the branch termini. The main nodes branched out dichotomously, and polytomy at the branch termini resulted from the fact that the sequences that formed particular clades were the same.

The obtained tree presents the division of rye species into two groups (Fig 2). A separate group consisted of the following species: *S. cereale* ssp. *cereale* (Canada, USA, Tajikistan), *S. cereale* ssp. *dighoricum* (Russia) forming a common subgroup with *S. sylvestre* (Bulgaria) and *S. vavilovii* (Hungary).

In the second group, *S. sylvestre* (Hungary) was the least similar to the others, followed by *S. vavilovii* (Poland) and *S. strictum* ssp. *africanum* (RSA1), forming one subgroup and showing similarity to the following species subgroup: *S. strictum* ssp. *kuprijanovii* (Poland), *S. strictum* ssp. *strictum* (Poland) and *S. cereale* ssp. *segetale* (Kazakhstan).

There were no clear species and geographic specific patterns (e.g. monophyly) on the tree, as resolved clades were consisted of different species from different geographic locations. Two analyzed samples of *S. strictum* ssp. *africanum* were located in two distant similarity groups, as in the cpDNA analyses.

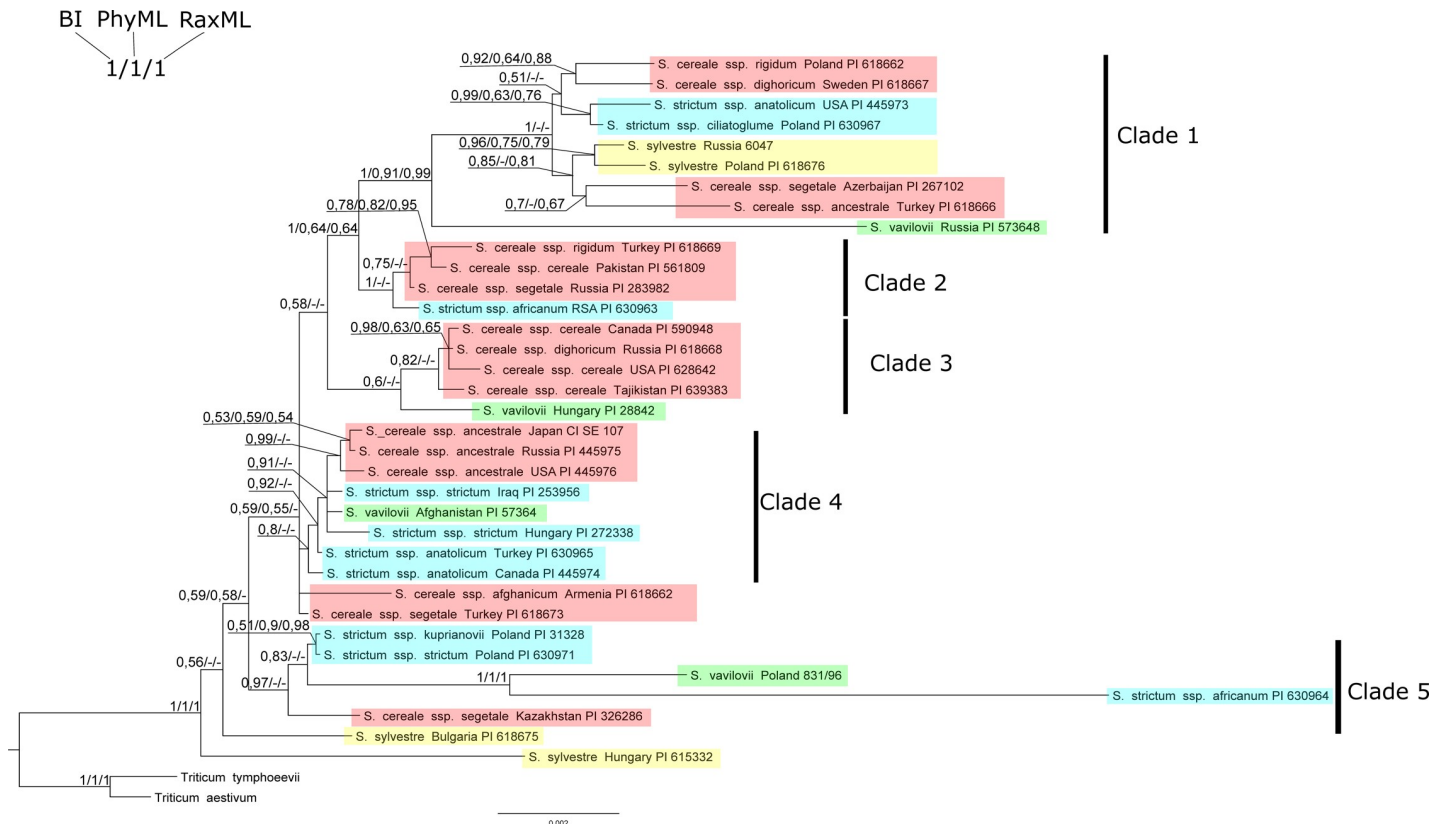


Fig 3. Bayesian phylogenetic tree of 35 *Secale sp.* individuals and two *Triticum sp.* representatives used as an outgroup. Bootstrap values (RaxML and PhyML) and posterior probability values are presented along nodes (values below 0.5 are not shown).

<https://doi.org/10.1371/journal.pone.0213023.g003>

The pooled analysis of sequences of non-coding regions of chloroplast and mitochondrial DNA—Phylogenetic analysis

The BI, PhyML, and RaxML trees of all analyzed non-coding sequences of chloroplast and mitochondrial DNA presented a similar topology with slight differences (Fig 3). There were no or almost no clear systematic or geographical specific patterns (e.g. monophyly) on the tree, and the distance between representatives of some species was similar (or larger) than between representatives of different species. Lack of visible monophyly suggested that the species division did not fully reflect the reality.

We were able to distinguish five distinct genetic clades on the tree that were used for distance analyses (Fig 3). The first clade consisted of 4 *S. cereale* subspecies, 2 subspecies of *S. strictum*, *S. sylvestre* (Russia and Poland) and *S. vavilovii* (Russia). The second clade was formed of three *S. cereale* and *S. strictum ssp. africanum* (RSA) subspecies. The third clade mainly grouped *S. cereale* and *S. vavilovii* subspecies (Hungary). The fourth clade included subspecies of *S. cereale*, *S. strictum ssp. strictum* (Iraq, Hungary), *S. s. ssp. anatolicum* (Turkey and Canada) and *S. vavilovii* (Afghanistan). The last highlighted clade contained *S. vavilovii* (Poland) and *S. s. ssp. africanum* (RSA1).

S. sylvestre (Hungary) and *S. sylvestre* (Bulgaria) formed a separate group. The other two accessions of this species (Russia and Poland) were similar, but they were in a separate clade 1. Most of the accessions of the cultivated species were located in clade 3, with the exception of *S. c. ssp. cereale* from Pakistan, which was located in clade 2 and showed a high similarity to *S. c.*

Table 3. Estimates of average evolutionary divergence over sequence pairs within clades. The number of base substitutions per site from averaging over all sequence pairs within each group are shown (MPD) along with standard error estimate(s) (SE).

	MPD	SE
Clade 1	0.003959	0.000301
Clade 2	0.000493	0.000157
Clade 3	0.00057	0.000173
Clade 4	0.000296	9.88E-05
Clade 5	0.011245	0.000853
Average	0.003313	0.000317

<https://doi.org/10.1371/journal.pone.0213023.t003>

ssp. *rigidum* from Turkey. *S. c.* ssp. *cereale* (Canada, Russia, USA) was also genetically similar to *S. c.* ssp. *dighoricum* (Russia).

Similarly, 3 accessions of *S. c.* ssp. *ancestrale* (Japan, Russia, USA) grouped together in clade 4, but showed similarity to *S. s.* ssp. *strictum* (Iraq, Hungary) and *S. vavilovii* (Afghanistan). The remaining *S. c.* ssp. *ancestrale* (Turkey) was similar to *S. c.* ssp. *segetale* (Azerbaijan) (clade 1). *S. vavilovii* was dispersed in four clades, but it was always distant from other species (Fig 3).

Distance between and within clades

The average pairwise distance between sequences within clades was estimated at 0.003313 with a range from 0.000296 for clade 4 to 0.011245 for clade 5 (Table 3). On the other hand, the average distance between different clades was 0.0055 ranging from 0.00108, for clade 3-clade 4 distance, to 0.0126 for clade 1-clade 5 distance (Table 4). When compared, distances between clades were on average higher than within them. Nevertheless, the lowest distance within the clades was nearly four times lower than the same value between clades and the highest distance within clades was only slightly different than between them.

Discussion

The genetic information contained in both chloroplast (cpDNA) and mitochondrial DNA (mtDNA) is often analyzed in phylogenetic and population studies. However, for most species, including rye, the data is incomplete. So far, rye mtDNA has not been fully sequenced, while in 2014, complete data on the *S. cereale* chloroplast genome sequence has been published [40]. Only in 2007, Isik et al. [33] analyzed the organelle genomes of *S. cereale* cultivars originated from different geographical regions using the PCR-RFLP method. The study used 7 cpDNA fragments and 4 mtDNA fragments for both coding and non-coding regions. Each amplified sequence was digested with 13 different restriction enzymes. The study of Isik et al. [33] proved that the mitochondrial genome, as compared to the chloroplast genome, showed a higher level of organelle polymorphism between analyzed rye cultivars. Our research [34], consisting in the analysis of mtDNA of seven species and subspecies of rye based on the RFLP method,

Table 4. Estimates of evolutionary divergence over sequence pairs between clades. The number of base substitutions per site from averaging over all sequence pairs between groups are shown. Standard error estimate(s) are shown above the diagonal.

Clade 1		0.0006099	0.0006728	0.0007307
Clade 2	0.0057124		0.0003744	0.0003348
Clade 3	0.0059956	0.0012996		0.0003018
Clade 4	0.0063539	0.0012154	0.0010806	
Clade 5	0.0126718	0.0073360	0.0071080	0.0062579

<https://doi.org/10.1371/journal.pone.0213023.t004>

showed the division of the analyzed taxa into two groups. The first included *Secale cereale* ssp. *segetale* and *Secale sylvestre*, and the second group comprised the remaining of the analyzed species (*Secale strictum*, *Secale strictum* ssp. *kuprijanovii* and *Secale vavilovii*, *Secale cereale* and *Secale strictum* ssp. *africanum*), which did not overlap with the existing classification system.

In 2016, Hagenblad et al. [25] analyzed the genetic diversity of 76 accessions of wild, feral and cultivated rye based on SNP polymorphisms. They performed an analysis of five chloroplast SSRs, derived from *Lolium* and wheat. Discriminant analysis of principal components (DAPC) of cpSSR data indicated very large genetic variation within the genus *Secale* and did not reflect taxonomic groups, except for *S. strictum* and *S. africanum*, which formed a separate cluster.

In the latest study on phylogenetic analysis of the genus *Secale* [24] the following collection of plants representing 11 species and subspecies of rye was studied using SSR and EST nuclear sequences. These results indicated high genetic diversity of *S. strictum* in comparison to other *Secale* species. The existence of two separate clusters of different species and subspecies was also found, and their division was not consistent with taxonomic affiliation, but was based on geographical origin (samples from Asia and outside of Asia). A clear separation between *S. sylvestre* and the rest of the genus was also revealed.

The analyses in the current study have included commonly used regions as well as those that are much less frequently applied in phylogenetic analyses, since no relationship analyses have been performed so far within the genus *Secale* based on non-coding sequences of organelle genomes. The amplification was carried out using universal primer pairs that allowed the amplification of non-coding regions dividing the two coding fragments in most plants [42,43,45]. The use of primers that are homologous to most coding regions, but amplify variable non-coding regions is very useful for phylogenetic and population studies [42,60,61].

Evaluation of sequence variability of cpDNA non-coding regions in rye species

The cpDNA regions are widely used as markers in phylogenetic and phylogeographic studies, however, little is known about their usefulness for studying the relationships between closely related species, especially in monocotyledonous plants [61,62]. The slow rate of cpDNA-specific evolution hinders taxonomic analyses at lower levels, especially at the population level. In turn, non-coding regions are characterized by a higher rate of evolution than the coding regions—for example, the *trnL-trnF* region evolved 1.93–1.72 times faster than *rbcl* in certain genera of *Gramineae* [35].

The most commonly used non-coding regions, such as the *trnL*, *trnL-trnF* and *trnK* intron/*matK* have proved useful in phylogenetic analyses of some plant groups [63, 64], but often showed too low resolution in other groups, at least in some clades of angiosperms [65–67]. To obtain additional data and a better resolution for phylogenetic studies, sequences from these popular regions are often used in combination with other cpDNA sequences or with mtDNA and nDNA sequences [68–72], because additional information is often necessary to provide a satisfactory hypothesis regarding phylogenesis.

Studies on cpDNA clearly indicate that the utility in phylogenetic analyses of different cpDNA non-coding regions within a given taxonomic group can vary enormously [68, 72–77], and the selection of the appropriate cpDNA region is often difficult due to the lack of information about the rate of evolution between different non-coding cpDNA regions. This was confirmed by the results obtained in the present work, because they indicate a high variability of the studied regions of the chloroplast genome in the majority of taxa (Table 1). The most useful for the analysis of closely related taxa turned out to be *trnT* (UGU)-*trnL* (UAA)

5' exon and *trnD*[tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] regions, especially when used in combination. These areas should serve as useful molecular markers for studies of closely related species, at least at inter-specific level in the genus *Secale*.

cpDNA shows a much more stable structure in case of intramolecular rearrangement than the mitochondrial genome of plants. However, the rate of plastid genome substitution is 3–4 times higher than that of plant mtDNA [78]. Most of the variability observed in cpDNA non-coding regions concerns insertion-deletion (indel) mutations, but, as stated by other authors, they should be treated with caution as they may indicate heteroplasmy [79]. However, indels were analyzed in the following studies due to the fact that they were found to be common and often phylogenetically informative [35,80,81]. In our studies the phylogenetic indeles context were not analysed because algorithms describing substitution models are not able to model indel-type changes and they are removed by tree-creating programs. A total of 81 indels have been identified—an average of 1.03 per sample (Table 1). This number is not very high compared to the results presented by other authors. Shaw et al. [82], analyzing 10 groups representing three different angiosperm lines: *Atropa* vs. *Nicotiana* (asteridy); *Lotus* vs. *Medicago* (rosids); and *Saccharum* vs. *Oryza* (monocotyledons), characterized a total of 1260 indels—an average of 3.0 per sample. It can be concluded that the appropriate combination of regions alongside the described indels as an additional factor, strengthens the usefulness of cpDNA for phylogenetic studies.

This study identified from 5 to 16 haplotypes per DNA fragment, and the most frequent was haplotype 1 in the non-coding *atpB-rbcL* region (S8 Table). The *atpB-rbcL* region, located between the gene encoding the large ribulose-1,5-bisphosphate carboxylase/oxygenase (*rbcL*) subunit and the gene encoding the β subunit of chloroplast ATP synthase (*atpB*) has different size in different taxonomic groups. According to a study conducted by Chiang [42], its size in the studied species ranged from 524 bp to about 1000 bp [42]. In *S. cereale*, it has a length of 774 bp [40]. The results obtained in the present work are comparable, since the size of this region in rye species is 815 bp (Table 1). In addition, it is characterized by a moderate variability. The number of haplotypes for this region is 8, which considering the 13 analyzed taxa is quite high. The occurrence of deletions and insertions as well as numerous nucleotide substitutions is a common phenomenon in the analyzed region. In addition, *atpB-rbcL* has been shown to be AT rich. The majority of non-coding regions rich in these base pairs show a low number of functions [83].

The next analyzed region—*trnT* (UGU)-*trnL* (UAA) 5' exon is located between the tRNA genes. It shows a high frequency of insertions or deletions, depending on the species, which makes it possible to use them as genetic markers. The sequence length of this region in the current work was 600 bp (Table 1), similarly in *S. cereale* it was 615 bp [40]. In general, this area was characterized by a variable length, from 298 bp to 700 bp, as demonstrated in studies in certain bryophytes, gymnosperms and angiosperms [44]. Exceptionally, its size in rice was 770 bp, in tobacco—710 bp, and it was much shorter in *Marchantia polymorpha*—188 bp [84]. This region proved to be a very good marker for studying closely related *Secale* species, due to its high variability.

The region of the *trnL*(UAA) intron is well known and its sequences have already been used to identify phylogenetic relationships between closely related species or to identify plants [85]. There is a view expressed repeatedly in the literature that among the regions of the chloroplast genome, the *trnL*(UAA) intron, belonging to type I introns, is less variable due to its catalytic properties and secondary structure, and therefore, may be more useful at higher taxonomic levels [44, 85–88]. This was confirmed by the results obtained in the presented work. The *trnL*(UAA) intron size in most of the species studied was from 254–767 bp. It was of medium size (594 bp) in *S. cereale* [40], and the amplified size in the studied *Secale* taxa was 653 bp (Table 1) and corresponded to the size of other species.

In contrast, the length of the *trnD* [tRNA-Asp (GUC)]—*trnT* [tRNA-Thr (GGU)] region in the analyzed rye species and subspecies was 963 bp (Table 1), similarly as in *S. cereale*—948 bp [40]. It was slightly shorter than described in wheat (1047 bp) and maize (1185 bp) or couch grass (1200 bp), and also significantly shorter than the corresponding region in barley (1978 bp) [43]. This region, next to *trnT*(UGU)-*trnL*(UAA) 5' exon, was characterized by the highest variability in rye species among all analyzed cpDNA regions and could be successfully used in studies of closely related species.

Evaluation of sequence variability of mtDNA non-coding regions in rye species

The mitochondrial genomes of plants are characterized by the presence of a relatively large number of group II introns compared to mtDNA of fungi and bacteria [89,90]. Group II introns described so far in plants have been located in the genes encoding proteins [91]. There was one exception noted—the presence of group II intron in the *trnA* gene in several species of the genus *Citrus* and in the wheat genome [92,93]. However, the *trnA* gene is of chloroplast origin and also has an intron in cpDNA [94]. Approximately 1/3 of group II introns is located in the plastid tRNA.

Several plant genera, including *Peperomia* and *Marchantia* also contains group I introns, located in the *coxI* gene [94]. However, there is no correlation between phylogenesis and the presence of this intron, which indicates that it was introduced by horizontal gene transfer, and a fungal species was probably the donor.

In this study, a total of 45 indels have been identified (Table 2) and this number is comparable to the results of other authors in mtDNA of plants. For example, Christensen et al. identified 35 indels in *A. thaliana* mtDNA [95], and Ossowski et al. [96] reported that the *A. thaliana* mitochondrial genome was characterized by a much higher content of indels in comparison to the nuclear genome.

Unfortunately, there is no sequence data on rye mtDNA, however, the mtDNA sequence in winter wheat (*Triticum aestivum* cv. Chinese Yumai) is almost identical to the spring wheat sequence (*T. aestivum* cv. Chinese Spring) [97]. Only 10 indels were identified between two independently acquired sequences, and all variants were found in non-coding regions. A 4-bp indel served as a convenient marker for discriminating the cultivar Chinese Yumai from Chinese Spring [97].

On average, 3 haplotypes were identified among the six analyzed mtDNA sequences (Table 2). Most haplotypes (12) were detected in the *nad4/1-2* region. The most frequently occurring was haplotype 1 in the non-coding *rps12-nad3(2)* region (S9 Table).

Mitochondrial primers amplifying the intergenic *nad1B-nad1C* region are located in exon b and c [43]. This region has a length of about 1483 bp in the analyzed rye species and subspecies (Table 2). Its size is comparable with the corresponding region in the couch grass (1600 bp) [60], and pedunculate oak (1550 bp) [43] and is shorter than that described in the *Arabidopsis* [41]. *Marchantia polymorpha* is one of the species that do not have an intron in the *nad1* gene [94]. The results obtained in this work confirmed the highly conserved nature of this intron group. The *nad1* intron region may nevertheless serve as a useful molecular marker in population studies, because it contains size variants that reveal the population structure in the entire range, e.g., in *Pinus ponderosa* and *P. flexilis* [98–100]. Unfortunately, no such differences were found in rye, and the obtained sequences were conserved (Table 2).

The intron located within subunit 4 of the *nad4* gene is considered to be a slowly evolving mitochondrial marker, whose evolution occurs 23 times slower than that of the ITS rDNA sequence [101]. Therefore, it is considered more useful in solving the “deeper” phylogenetic

relationships. It was successfully used in the phylogenetic analyses of the Brassicaceae family [102]. It was found in this work that this sequence could describe phylogenetic relationships within almost the whole family. Similarly, based on the presented studies, it was confirmed that sequences of this region proved to be the most informative among all tested mtDNA sequences. Twelve haplotypes were described among 35 analyzed taxa, and the haplotypes diversity was the highest (0.867) (Table 2). The length of this intron in *Secale* species is 1642 bp and is significantly shorter than that described in *Arabidopsis* (2103 bp) [41]. The reason why the *nad4/1-2* intron is shorter is not known due to the lack of information on these sequences in rye in the NCBI Genbank. The sequences of this intron, in addition to the *rps12-nad3(2)* intergenic sequence, showed variation at both intraspecific and interspecific levels (Table 2).

Another analyzed intergenic region, *nad4L-orf25*, in rye (626 bp) is of size corresponding to those sequences in *Arabidopsis* (671 bp) and slightly shorter than in wheat (899 bp) [99]. These sequences were conserved in the taxa analyzed (Table 2). This result was confirmed by the data on sugar beet [103], indicating conserved nature of this region in angiosperms. The gene encoding the 4L subunit of NADH dehydrogenase (*nad4L*) in the mitochondrial genome of *Arabidopsis thaliana* is located between exon c of the *nad5* open reading frame and *orf25* [104], in wheat between *nad7* and *rps19* [99].

Two different combinations of primers were used to amplify the intergenic sequences of the *rps12-nad3* region: the first amplified only the intergenic region, the second—the intergenic region and the *nad3* gene. The results obtained in this study indicated the variability of the analyzed intergenic sequences, while the first region proved conserved in *Secale* species and subspecies (Table 2). The length of the *rps12-nad3* intron in the analyzed taxa was 134 bp, and it was much shorter than the intron of *Arabidopsis* (700 bp). In contrast, its size in wheat is 45 bp [99]. The second primer pair amplified the 120-bp region, comparable to the corresponding region in *Arabidopsis* (139 bp).

The *rrn5-rrn18* ribosomal region in the analyzed forms was 166 bp long, slightly shorter than that described in *Arabidopsis* (273 bp) and comparable to the corresponding wheat region (115 bp). This region was also conserved among the analyzed *Secale* species and sub-species (Table 2).

Evaluation of genetic diversity in the genus *Secale*

Trees obtained using different methods (BI, ML) for all analyzed sequences of non-coding chloroplast and mitochondrial DNA and a collective tree, present a similar topology. There are no visible systematic or geographical relationships on the trees (Figs 1–3). Clades separated on the trees combine various rye subspecies from different locations. The results obtained in this work confirmed the reports of other authors about the lack of monophyly of *Secale* sp. subspecies resulting from similar values of inter- and intraspecific distance [105]. In contrast to the results presented by Santos [26], our analyses showed that sequences from the same geographical regions did not group in any of the phylogenetic trees.

The existence of 5 clades encompassing various taxa may indicate the introgression between them (Fig 3), as previously described [6,106,107]. Only *S. sylvestre* Hungary and Bulgaria, formed a separate group, which confirmed the latest results on the basis of the SSR and nuclear EST analysis [24] and the results based on the DArTseq analysis [27]. Many authors [3,4,12,108] consider *S. sylvestre* as the oldest species, from which all other species evolved. Similarly, in our research, it was the most diverse species of all *Secale* taxa (Fig 3). Hammer [109] argued that *S. sylvestre* was developing separately, and its evolution could have begun very early. Unfortunately, the remaining analyzed *S. sylvestre* taxa from Poland and Russia grouped in other clades (Clade 1) with *S. cereale* and *S. strictum* species and subspecies (Fig 3).

Similar results were obtained by Ren et al. [110] based on the analysis of microsatellite sequences. They showed similarity of *S. sylvestre* to *S. strictum* ssp. *africanum* and *anatolicum*. Skuza et al. [18] in turn classified *S. sylvestre* with *S. cereale* ssp. *segetale* based on mtDNA analysis.

S. vavilovii is an annual species, similar to *S. cereale*, *S. sylvestre* and *S. segetale*, however, according to Hammer [109], it has evolved with multiannual species: *S. africanum*, *S. strictum* and *S. kuprijanovii*. Research conducted by Jones and Flavell [28] hypothesized the common origin of *S. vavilovii*, *africanum* and *cereale* from the *S. strictum* line. The results obtained in the current work did not confirm this theory. *S. vavilovii* was dispersed and was in the similarity group with both *cereale* and *strictum* (Fig 3). The result of the analysis was partly consistent with the classification of Frederiksen and Petersen [111], who identified only three species within the genus *Secale*: *S. sylvestre*, *S. strictum* and *S. cereale* and included *S. vavilovii* to *S. cereale*. Similarly, Kobyljanski [112] classified *S. vavilovii* as a subspecies of *S. cereale*. Bolibok-Bągoszewska et al. [113] also postulated to classify these species together. The placement of *S. vavilovii* and *S. cereale* in one section is supported by the fact that seven bivalens are observed in the crosses of these two species, which indicates a complete structural similarity between these species [106]. Our result was also confirmed by the earlier data of Al-Beyroutiouva et al. [27], who argued that *S. vavilovii* could not be considered as a separate species, but only as a subspecies of *S. cereale*. Maraci et al. [24] reached similar conclusions, showing a greater similarity between *S. vavilovii* and *S. c. ssp. cereale* and *segetale* than within the *cereale* species.

The previous results of other authors indicated that the separate and phylogenetic position of *S. strictum* subspecies was unquestionable [12]. However, the assumption regarding the origin of *strictum* species is inconsistent with our results, as well as with the results of tandem sequence repeats and AFLP and SSR data [12,13,15]. The *strictum* species group is heterogeneous and shows similarity to *S. cereale* subspecies, similarly as in the work of Ren [113]. This is consistent with the hypothesis that cultivated rye evolved from *S. strictum* [113–117]. Genetic diversity in the evolutionary process was lower in the *strictum* group than between perennial and annual forms and species. In addition, it has been shown that the perennial forms are morphologically similar and cross easily to form hybrids [3,118].

The taxonomic position of *S. strictum* ssp. *ciliatoglume* is also often undermined. Bolibok-Bągoszewska et al. [113] argued to classify it together with *S. vavilovii* in the *S. cereale* species group. However, the results obtained in the present work did not confirm this theory, because *S. strictum* ssp. *ciliatoglume* showed a clear similarity to *S. strictum* ssp. *anatolicum* (Figs 1 and 3). Broda et al. [118] also classified *S. s. ssp. ciliatoglume* in a separate group to *S. vavilovii*. Al-Beyroutiouva et al. [27] postulated to analyze this taxon more accurately due to the fact that *S. s. ssp. ciliatoglume* was not associated with other *strictum* subspecies.

Hammer [109] grouped *S. strictum* species into one group with *S. vavilovii* and *S. kuprijanovii*, which was not confirmed in the present study. On the other hand, a large similarity of *S. vavilovii* (Poland) to *S. s. ssp. africanum* (RSA1) could be observed (Figs 2 and 3). The results obtained in the current study confirmed the similarity of *S. s. strictum* to *S. c. ssp. kuprijanovii* (Figs 2 and 3), consistent with the adopted classification. Noteworthy is the heterogeneity of the sequence of two analyzed samples of *S. strictum* ssp. *africanum*. The results clearly suggested that the origin of these seeds was not homogeneous—the second sample (RSA1) formed a separate subgroup.

S. africanum species has been most often classified to *S. strictum*. Hammer [109] considered *S. africanum* and *S. strictum* species to be very similar to each other, which was confirmed in this work (Fig 3). However, the latest data based on SNP analyses confirmed our results, classifying *S. strictum* and *africanum* as separate species [25]. *S. africanum* occurs only in southern Africa, but it can cross with other *Secale* species [13]. The high genetic similarity between *S.*

africanum and *S. strictum* species strengthens the hypothesis that the location of *S. africanum* away from other *Secale* taxa should be explained by human activities, and not as a remnant of the originally much larger *Secale* distribution area [111].

S. cereale species also did not form groups consistent with the current classification. The dispersion of *cereale* species was also described by Hagenblad et al. [25], who found that genetic grouping of *cereale* species was more dependent on geographical origin than the taxonomic classification. Similarly, Al-Beyroutioua et al. [27] reported insufficient divergence of *cereale* subspecies.

Conclusions

The use of non-coding sequences of chloroplast and mitochondrial DNA provided new data on genetic diversity within the genus *Secale*. Such sequences have so far not been used in any analysis of phylogenetic relationships in rye. The results obtained in this study clearly indicated disproportions in the available information regarding various non-coding cpDNA regions used in phylogenetic studies, and some of them—due to high variability—can be successfully used in the analyses of closely related species.

The analysis of genetic diversity and phylogenetic relationships of the genus *Secale* shows the lack of monophyly of *Secale* sp. subspecies resulting from similar inter- and intraspecific distance, which indicates the continuation of gene flow between species. The reason for this may be the relatively recent evolution of rye species that did not allow the full formation of isolation mechanisms.

The lack of a clear genetic diversity between *S. cereale* and *S. vavilovii* confirmed the assumption that they might share a common ancestor. Furthermore, it allows to conclude—as mentioned by many authors—that *S. vavilovii* should be classified as a subspecies of *S. cereale*. Moreover, the results obtained in this work confirmed the divergence of *S. sylvestre* from other species and subspecies of rye.

The results also indicate regions that may be useful molecular markers in studies on closely related species of the genus *Secale*. These include the non-coding regions of chloroplast DNA: *trnT* (UGU)-*trnL*(UAA)5' exon and *trnD*[tRNA-Asp(GUC)]-*trnT*[tRNA-Thr(GGU)] and non-coding regions of mitochondrial DNA: *nad4/1-2* and *rps12-nad3(2)*. However, the chloroplast genome is more useful for this kind of analysis. We presume that the analysis of the complete chloroplast genome sequences in species of rye will be useful and cost-effective for evolutionary and phylogenetic studies.

Supporting information

S1 Table. The list of plant species, origin, accession numbers, NCBI Genebank accession numbers, type and life cycle.

(DOCX)

S2 Table. Primer list of sequences (5'-3') for amplification different regions of cpDNA and their references used in this study.

(DOCX)

S3 Table. Primer list of sequences (5'-3') for amplification different regions of mtDNA and their references used in this study.

(DOCX)

S4 Table. PCR reagents for amplification different regions the cpDNA.

(DOCX)

S5 Table. PCR reagents for amplification different regions the mtDNA.
(DOCX)

S6 Table. Thermocycling conditions for a PCR amplification of cpDNA noncoding (intron) regions.
(DOCX)

S7 Table. Thermocycling conditions for a PCR amplification of mtDNA noncoding (intron) regions.
(DOCX)

S8 Table. The type and number of haplotypes obtained for the noncoding cpDNA sequences.
(DOCX)

S9 Table. The type and number of haplotypes obtained for the noncoding mtDNA sequences.
(DOCX)

Author Contributions

Conceptualization: Lidia Skuza.

Data curation: Tomasz Strzała.

Investigation: Lidia Skuza, Izabela Szućko, Ewa Filip.

Writing – original draft: Lidia Skuza.

Writing – review & editing: Lidia Skuza, Izabela Szućko, Ewa Filip.

References

1. Bujak H. Analysis of the inheritance of quantitative traits of rye. Bulletin of Plant Breeding Acclimatization Institute IHAR. 2000; 216(1): 43–54.
2. Bartoš J, Paux E, Kofler R, Havránková M, Kopecký D, Suchánková P, et al. first survey of the rye (*Secale cereale*) genome composition through BAC end sequencing of the short arm of chromosome 1R. BMC Plant Biology. 2008; 8: 1–12. <https://doi.org/10.1186/1471-2229-8-1>
3. Spencer HA, Hawkes JG. On the origin of cultivated rye. Biol J Linn Soc. 1980; 13: 299–313.
4. De Bustos A, Jouve N. Phylogenetic relationships of the genus *Secale* based on the characterization of rDNA ITS sequences. PI Syst Evol. 2002; 235:147–154.
5. Singh R, Röbbelen G. Identification by Giemsa technique of the translocations separating cultivated rye from three wild species of *Secale*. Chromosoma. 1977; 59: 217–225
6. Stutz HC. A cytogenetic analysis of the hybrid *Secale cereale* L. x *Secale montanum* Guss. and its progeny. Genetics. 1957; 42, 199. PMID: [17247691](https://pubmed.ncbi.nlm.nih.gov/17247691/)
7. Bennett MD, Smith J. B. Nuclear DNA amounts in angiosperms. Phil. Trans. Roy. Soc. Lond. B. 1976; 274: 227–274
8. Bennett MD, Gustafson JP, Smith JB. Variation in nuclear DNA in the genus *Secale*. Chromosoma. 1977; 2:61, 149–176.
9. Murai K, Naiyu X, Tsunewaki K. Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. III. Chloroplast DNA variation and interspecific relationships in the genus *Secale*. Japan. J. Genet. 1989; 64: 36–47.
10. Petersen G, Doebley F. Chloroplast DNA variation In the genus *Secale* (Poaceae). Plant Syst Evol. 1993; 187: 115–125.
11. Cuadrado A, Jouve N. Evolutionary trends of different repetitive DNA sequences during speciation in the genus *Secale*. J Hered. 2002; 93(5): 339–45. PMID: [12547922](https://pubmed.ncbi.nlm.nih.gov/12547922/)

12. Shang HY, Wei YM, Wang XR, Zheng YL. Genetic diversity and phylogenetic relationships in the rye genus *Secale* L. (rye) based on *Secale cereale* microsatellite markers. *Genet Mol Biol*. 2006; 29(4): 685–691.
13. Zhou J, Yang Z, Li G, Liu C, Tang Z, Zhang Y, et al. Diversified chromosomal distribution of tandemly repeated sequences revealed evolutionary trends in *Secale* (Poaceae). *Plant Syst Evol*. 2010; 287: 49–56.
14. Chikmawati T, Skovmand B, Gustafson JP. Phylogenetic relationships among *Secale* species revealed by amplified fragment length polymorphisms. *Genome*. 2005; 48(5): 792–801. <https://doi.org/10.1139/g05-043> PMID: 16391685
15. Fowler DB, Limin AE. Exploitable genetic-variability for cold tolerance in commercially grown cereals. *Can J Plant Sci*. 1987; 67: 278.
16. Madej LJ. Worldwide trends in rye growing and breeding. *Vortr Pflanzenzucht*. 1996; 35:1–6.
17. Schreiber M, Himmelbach A, Börner A, Mascher M. Genetic diversity and relationship between domesticated rye and its wild relatives as revealed through genotyping by-sequencing. *Evol Appl*. 2019; 12:66–77.
18. Skuza L, Filip E, Szućko I. Intergenic spacer length variability in cultivated, weedy and wild rye species. *Open Life Science*. 2015; 10(1): 175–181.
19. Rzepka-Plevneš D. Utility properties of hybrids *S. cereale* × *S. vavilovii* Gross. in terms of their suitability in growing rye varieties resistant to sprouting. Part I. *Bulletin of Plant Breeding Acclimatization Institute IHAR*. 1993; 37/ 5(6): 69–79.
20. Rzepka-Plevneš D, Tomczak P. Utility properties of hybrids *S. cereale* × *S. vavilovii* Gross. in terms of their suitability in growing rye varieties resistant to sprouting. Part II. *Bulletin of Plant Breeding Acclimatization Institute IHAR*. 1993; 37/ 5(6): 81–91.
21. Doebley J, Bothmer von R, Larson S. Chloroplast DNA variation and the phylogeny of *Hordeum* (Poaceae). *Am J Bot*. 1992; 79: 576–584.
22. Aoki S, Ito M. Molecular phylogeny of *Nicotiana* (*Solanaceae*) based on the nucleotide sequence of the *matK* gene. *Plant Biol*. 2000; 2: 316–324.
23. Matyas G, Sperisen C. Chloroplast DNA polymorphisms provide evidence for postglacial re-colonisation of oaks (*Quercus* spp.) across the Swiss Alps. *Theor Appl Genet*. 2001; 102: 12–20.
24. Maraci Ö, Özkan H, Bilgin R. Phylogeny and genetic structure in the genus *Secale*. *PLoS One*. 2018; 13(7):e0200825. Available from: <https://doi.org/10.1371/journal.pone.0200825> PMID: 30024916
25. Hagenblad J, Oliveira HR, Forsberg NEG, Leino MW. Geographical distribution of genetic diversity in *Secale* landrace and wild accessions. *BMC Plant Biol*. 2016; 16(1):23.
26. Santos E, Matos M, Silva P, Figueiras AM, Benito C, Pinto-Carnide O. Molecular diversity and genetic relationships in *Secale*. *J Genet*. 2016; 95: 273–28128. PMID: 27350669
27. Al-Beyroutiouva M, Sabo M, Sleziaik P, Dušínský R, Birčák E, Hauptvogel P, et al. Evolutionary relationships in the genus *Secale* revealed by DArTseq DNA polymorphism. *Plant Syst Evol*. 2016; 302(8): 1083–1091.
28. Jones JDG, Flavell RB. The structure, amount and chromosomal localization of defined repeated DNA sequences in species of the genus *Secale* *Chromosoma* (Berl). 1982; 86:13–41.
29. Cuadrado A, Schwarzacher T. The chromosomal organization of simple sequence repeats in wheat and rye genomes. *Chromosoma*. 1998; 107: 587–594. PMID: 9933412
30. Fuchs JK, Ohne M, Schubert I. Assignment of link-age groups to pea chromosomes after karyotyping and gene mapping by fluorescent in situ hybridization. *Chromosoma*. 1998; 107: 272–276. PMID: 9745054
31. Cuadrado A, Jouve N. Distribution of highly repeated DNA sequences in species of the genus *Secale*. *Genome*. 1997; 40: 309–317. PMID: 9202411
32. Cuadrado A, Jouve N. Fluorescent in situ hybridization and C-banding analyses of highly repetitive DNA sequences in heterochromatin of rye (*Secale monatanum* Guss.) and wheat incorporating *S. monatanum* chromosome segments. *Genome* 1995; 38: 795–802. PMID: 7672610
33. Isik Z, Parmaksiz I, Cotuh C, Geylan-Su YS, Cebeci O, Beecher B, Budak H. Organellar genome analysis of rye (*Secale cereale*) representing diverse geographic regions. *Genome*. 2007; 50(8): 724–734. <https://doi.org/10.1139/g07-052> PMID: 17893732
34. Skuza L, Rogalska SM, Bocianowski J. RFLP Analysis Of Mitochondrial DNA In The Genus *Secale*. *Acta Biol Cracoviensia Series Botanica*. 2007; 49(1): 77–87.
35. Gielly L, Taberlet P. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus *rbcl* sequences. *Mol Biol Evol*. 1994; 11(5): 769–77. <https://doi.org/10.1093/oxfordjournals.molbev.a040157> PMID: 7968490

36. Suo Z, Zhang C, Zheng Y, He L, Jin X, Hou B, et al. Revealing genetic diversity of tree peonies at micro-evolution level with hyper- variable chloroplast markers and floral traits. 2012; *Plant Cell Rep.* 31, 2199–2213. <https://doi.org/10.1007/s00299-012-1330-0> PMID: 22961193
37. Dong W, Xu C, Li D, Jin X, Li R, Lu Q, et al. Comparative analysis of the complete chloroplast genome sequences in psammophytic *Haloxylon* species (Amaranthaceae). *PeerJ* 4:e2699. 10.7717/peerj.2699 <https://doi.org/10.7717/peerj.2699> PMID: 27867769
38. Wang M, Xie X, Yan B, Yan X, Luo J, Liu Y, et al. The completed chloroplast genome of *Ostrya trichocarpa*. *Conserv. Genet. Resour.* 2013; 1–3. 10.1007/s12686-017-0869-z
39. Xu C, Dong WP, Li WQ, Lu YZ, Xie XM, Jin XB, et al. Comparative analysis of six *Lagerstroemia* complete chloroplast genomes. *Front. Plant Sci.* 2017; 8:15. 10.3389/fpls.2017.00015 <https://doi.org/10.3389/fpls.2017.00015> PMID: 28154574
40. Middelton CP, Senerchia N, Stein N, Akhunov ED, Keller B, Wicker T, et al. Sequencing of Chloroplast Genomes from Wheat, Barley, Rye and Their Relatives Provides a Detailed Insight into the Evolution of the Triticeae Tribe. *PlosOne.* 2014; 9(3): e85761. Available from: doi.org/10.1371/journal.pone.0085761
41. Duminiel J, Pemonge MH, Petit RJ. A set of 35 consensus primer pairs amplifying genes and introns of plant mitochondrial DNA. *Mol Ecol Notes.* 2002; 2:428–430.
42. Chiang TY, Schaal BA, Peng C. Universal primers for amplification and sequencing a noncoding spacer between the *atpB* and *rbcl* genes of chloroplast DNA. *Bot Bull Acad Sin.* 1998; 39:245–250.
43. Demesure B, Sodzi N, Petit RJ. A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Mol Ecol.* 1995; 4: 129–131. PMID: 7711952
44. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of the three noncoding regions of chloroplast DNA. *Plant Mol Biol.* 1991; 17: 1105–1109. PMID: 1932684
45. Dumolin-Lapegue S, Pemonge MH, Petit RJ. An enlarged set of consensus primers for the study of organelle DNA in plants. *Mol Ecol.* 1997; 6(4): 393–7. PMID: 9131816
46. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol Biol Evol.* 2017; 34:3299–3302. <https://doi.org/10.1093/molbev/msx248> PMID: 29029172
47. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32:1792–7. Available from: <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
48. Gouy M, Guindon S, Gascuel O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 2010; 27: 221–224. Available from: <https://doi.org/10.1093/molbev/msp259> PMID: 19854763
49. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 2003; 52: 696–704. Available from: [doi:10.1080/10635150390235520](https://doi.org/10.1080/10635150390235520) PMID: 14530136
50. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010; 59: 307–321. Available from: <https://doi.org/10.1093/sysbio/syq010> PMID: 20525638
51. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014; 30: 1312–1313. Available from: <https://doi.org/10.1093/bioinformatics/btu033> PMID: 24451623
52. Stamatakis A, Ludwig T, Meier H. RAxML: A Parallel Program for Phylogenetic Tree Inference. *Poster Abstr Proc 2nd Eur Conf Comput Biol.* 2003; 325–326.
53. Lefort V, Longueville JE, Gascuel O. SMS: Smart Model Selection in PhyML. *Mol Biol Evol.* 2017; 34: 2422–2424. Available from: <https://doi.org/10.1093/molbev/msx149> PMID: 28472384
54. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods.* 2012; 772–772. Available from: <https://doi.org/10.1038/nmeth.2109>
55. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 2012; 61: 539–42. Available from: <https://doi.org/10.1093/sysbio/sys029> PMID: 22357727
56. Huelsenbeck JP, Larget B, Alfaro ME. Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Mol Biol Evol.* 2004; 21: 1123–1133. Available from: <https://doi.org/10.1093/molbev/msh123> PMID: 15034130
57. Rambaut A. FigTree v1.4.3 [Internet]. Molecular evolution, phylogenetics and epidemiology. 2016; Available from: <http://tree.bio.ed.ac.uk/software/figtree/>
58. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol.* 1980; Available from: <https://doi.org/10.1007/BF01731581>

59. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018; 35: 1547–1549. Available from: <https://doi.org/10.1093/molbev/msy096> PMID: 29722887
60. Sun G, Interspecific polymorphism at non-coding regions of chloroplast, mitochondrial DNA and rRNA IGS region in *Elymus* species. *Hereditas.* 2002; 137:119–124. PMID: 12627837
61. Scarcelli N, Barnaud A, Eiserhardt W, Treier UA, Seveno M, d'Anfray A, et al. A Set of 100 Chloroplast DNA Primer Pairs to Study Population Genetics and Phylogeny in Monocotyledons. *PLoS One.* 2011; 6(5): e19954. Available from: <https://doi.org/10.1371/journal.pone.0019954> PMID: 21637837
62. Patwardhan A, Ray S, Roy A. Molecular Markers in Phylogenetic Studies-A Review. *J Phylogen Evolution Biol.* 20014; 2(2): 1000131
63. Ge S, Li A, Lu BR, Zhang SZ, Hong DY. A phylogeny of the rice tribe Oryzaceae (Poaceae) based on *matK* sequence data. *Am J Bot.* 2002; 89:1967–1972. <https://doi.org/10.3732/ajb.89.12.1967> PMID: 21665625
64. Bellstedt DU, Linder HP, Harley EH. Phylogenetic relationships in *Disa* based on noncoding *trnL-trnF* chloroplast sequences: evidence of numerous repeat regions. *Am J Bot.* 2001; 88: 2088–2100. Available from: <https://doi.org/10.2307/3558434> PMID: 21669640
65. Klak C, Hedderson TA, Linder HP. A molecular systematic study of the *Lampranthus* group (Aizoaceae) based on the chloroplast *trnT-trnF* and nuclear ITS and 5S NTS sequence data. *Systematic Botany.* 2003; 28: 70–85.
66. Muellner AN, Samuel R, Johnson SA, Cheek M, Pennington TD, Chase MW. Molecular phylogenetics of Meliaceae (Sapindales) based on nuclear and plastid DNA sequences. *Am J Bot.* 2003; 90: 471–480. <https://doi.org/10.3732/ajb.90.3.471> PMID: 21659140
67. Samuel R, Stuessy TF, Tremetsberger K, Baeza CM, Siljak-Yakovlev S. Phylogenetic relationships among species of *Hypochoeris* (Asteraceae, Cichorieae) based on ITS, plastid *trnL* intron, *trnL-F* spacer, and *matK* sequences. *Am J Bot.* 2003; 90: 496–507. <https://doi.org/10.3732/ajb.90.3.496> PMID: 21659142
68. Sang T, Crawford DJ, Stuessy TF. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *Am J Bot.* 1997; 84(9): 1120–1136.
69. Wang XR, Tsumura Y, Yoshimaru H, Nagasaka K, Szmjdt AE. Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcl*, *matK*, *rpL20-rpS18* spacer, and *trnV* intron sequences. *Am J Bot.* 1999; 86:1742–1753. PMID: 10602767
70. Hardig TM, Soltis PS, Soltis DE. Diversification of the North American shrub genus *Ceanothus* (Rhamnaceae): conflicting phylogenies from nuclear ribosomal DNA and chloroplast DNA. *Am J Bot.* 2000; 87: 108–123. PMID: 10636834
71. Schöenberger J, Conti E. Molecular phylogeny and floral evolution of the Penaeaceae, Oliniaceae, Rhynchochalcaceae, and Alzateaceae (Myrtales). *Am J Bot.* 2003; 90: 293–309. <https://doi.org/10.3732/ajb.90.2.293> PMID: 21659121
72. Yamane K, Yasui Y, Ohnishi O. Intraspecific cpDNA variations of diploid and tetraploid perennial buckwheat, *Fagopyrum cymosum* (Polygonaceae). *Am J Bot.* 2003; 90: 339–346. <https://doi.org/10.3732/ajb.90.3.339> PMID: 21659125
73. Ma R, Yli-Mattila T, Pulli S. Phylogenetic relationships among genotypes of worldwide collection of spring and winter ryes (*Secale cereale* L.) determined by RAPD-PCR markers. *Hereditas.* 2004; 140: 210–221. <https://doi.org/10.1111/j.1601-5223.2004.01844.x> PMID: 15198711
74. Xu DH, Sakai AJ, Kanazawa M, Shimamoto A, Shimamoto Y. Sequence variation of non-coding regions of chloroplast DNA of soybean and related wild species and its implications for the evolution of different chloroplast haplotypes. *Theor Appl Genet.* 2000; 101:724–732.
75. Hartmann SJ, Nason D, Bhattacharya D. Phylogenetic origins of *Lophocereus* (Cactaceae) and the senita cactus-senita moth pollination mutualism. *Am J Bot.* 2002; 89: 1085–1092. <https://doi.org/10.3732/ajb.89.7.1085> PMID: 21665708
76. Hamilton MB, Braverman JM, Soria-Hernanz DF. Patterns and relative rates of nucleotide and insertion/deletion evolution at six chloroplast intergenic regions in New World species of the Lecythidaceae. *Mol Biol Evol.* 2003; 20:1710–1721. <https://doi.org/10.1093/molbev/msg190> PMID: 12832633
77. Sakai M, Kanazawa A, Fujii A, Thseng FS, Abe J, Shimamoto Y. Phylogenetic relationships of the chloroplast genomes in the genus *Glycine* inferred from four intergenic spacer sequences. *Plant Syst Evol.* 2003; 239:29–54.
78. Avise JC. Phylogeography: retrospect and prospect. *J Biogeogr.* 2009; 36:3–15. Available from: doi.org/10.1111/j.1365-2699.2008.02032.x
79. Yamane K, Kawahara T. Intra and interspecific phylogenetic relationships among diploid *Triticum aestivum* species (Poaceae) based on base-pair substitutions, indels and microsatellites in chloroplast

- noncoding sequences. *Am J Bot.* 2005; 92(11): 1887–1898. Available from: <https://onlinelibrary.wiley.com/doi/epdf/10.3732/ajb.92.11.1887> PMID: 21646106
80. Golenberg EM, Clegg MT, Durbin ML, Doebley J, Ma DP. Evolution of a noncoding region of the chloroplast genome. *Mol Phyl Evol.* 1993; 2: 52–64.
 81. Morton BR, Clegg MT. A chloroplast DNA mutational hotspot and gene conversion in a noncoding region near *rbcl* in the grass family (Poaceae). *Current Genetic.* 1993; 24: 357–365.
 82. Shaw J, Lickey EB, Schilling EE, Small RL. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the Tortoise and the Hare III. *Am J Bot.* 2007; 94(3): 275–288. <https://doi.org/10.3732/ajb.94.3.275> PMID: 21636401
 83. Li WH. *Molecular Evolution*, Sunderland, Sinauer Associates. 1997; 487.
 84. Sugiura M. The chloroplast genome. *Plant Mol Biol.* 1992; 19: 149–168. Available from: <https://doi.org/10.1007/BF00015612> PMID: 1600166
 85. Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, et al. Power and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. *Nucleic Acids Res.* 2007; 35(3): e14 Available from: <https://doi.org/10.1093/nar/gkl938> PMID: 17169982
 86. Michel F, Jacquier A, Dujon B. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure, *Biochimie.* 1982; 64: 867–881. PMID: 6817818
 87. Davies RW, Waring RB, Ray JA, Brown TA, Scazzocchio C. Making ends meet a model for RNA splicing in fungal mitochondria. *Nature.* 1982; 300: 719–724. PMID: 6757759
 88. Shaw J, Lickey EB, Beck JT, Small RL. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot.* 2005; 92(1): 142–66. Available from: <https://doi.org/10.3732/ajb.92.1.142> PMID: 21652394
 89. Bonen I, Vogel J. The ins and outs of group II introns. *Trends Genet.* 2001; 17: 322–331. PMID: 11377794
 90. Lambowitz AN, Zimmerly S. Mobile group II introns. *Annu Rev Genet.* 2004; 38: 1–35. <https://doi.org/10.1146/annurev.genet.38.072902.091600> PMID: 15568970
 91. Brown GG, Colas des Francs-Small C, Ostersetzer-Brian O. Group II intron splicing factors in plant mitochondria. *Front Plant Sci.* 2014; 5: 35. Available from: <https://doi.org/10.3389/fpls.2014.00035> PMID: 24600456
 92. Ohtani K, Yamamoto H, Akimitsu K. Sensitivity to *Alternaria alternaria* toxin in citrus because of altered mitochondrial RNA processing. *Proc Natl Acad Sci USA.* 2002; 99: 2439–2444. <https://doi.org/10.1073/pnas.042448499> PMID: 11842194
 93. Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, et al. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucleic Acid Res.* 2005; 33: 6235–6250. <https://doi.org/10.1093/nar/gki925> PMID: 16260473
 94. Cho Y, Qiu YL, Kuhlman P, Palmer JD. Explosive invasion of plant mitochondria by a group II intron. *Proc Natl Acad Sci USA.* 1998; 95: 14224–14229.
 95. Christensen AC. Plant mitochondrial genome evolution can be explained by DNA repair mechanisms. *Genome Biol Evol.* 2013; 5(6): 1079–86. Available from: <https://doi.org/10.1093/gbe/evt069> PMID: 23645599
 96. Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, et al. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science.* 2010; 327(5961): 92–4. Available from: <https://doi.org/10.1126/science.1180677> PMID: 20044577
 97. Cui P, Liu H, Lin Q, Ding F, Zhuo G, Hu S et al. A complete mitochondrial genome of wheat (*Triticum aestivum* cv. Chinese Yumai), and fast evolving mitochondrial genes in higher plants. *J Genet.* 2009; 88(3): 299–307. PMID: 20086295
 98. Latta RG, Linhart YB, Fleck D, Elliot M. Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine. *Evolution. International Journal of Organic Evolution.* 1998; 52: 61–67.
 99. Mitton JB, Kreiser BR, Latta RG. Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Mol Ecol.* 2000; 9: 91–97. PMID: 10652078
 100. Johansen AD, Latta RG. Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Mol Ecol.* 2003; 12: 293–298. PMID: 12492897
 101. Yang YW, Lai KN, Tai PY, Li WH. Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between *Brassica* and other angiosperm lineages. *J Mol Evol.* 1999; 48: 597–604. PMID: 10198125

102. Couvreur TL, Franzke A, Al-Shehbaz IA, Bakker FT, Koch MA, Mummenhoff K. Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). *Mol Biol Evol.* 2010; 27(1): 55–71. Available from: <https://doi.org/10.1093/molbev/msp202> PMID: [19744998](https://pubmed.ncbi.nlm.nih.gov/19744998/)
103. Kubo T, Yamamoto MP, Mikami T. The *nad4L-orf25* gene cluster is conserved and expressed in sugar beet mitochondria. *Theor Appl Genet.* 2000; 100 (2):214–220.
104. Brandt P, Sünkel S, Unseld M, Brennicke A, Knoop V. The *nad4L* gene is encoded between exon c of *nad5* and *orf25* in the *Arabidopsis* mitochondrial genome. *Mol Gen Genet.* 1992; 236(1): 33–38. PMID: [1494348](https://pubmed.ncbi.nlm.nih.gov/1494348/)
105. Kellogg EA, Appels R. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. *Genetics.* 1995; 140(1): 325–343. PMID: [7635297](https://pubmed.ncbi.nlm.nih.gov/7635297/)
106. Singh RJ. Cross compatibility, meiotic pairing and fertility in 5 *Secale* species and their interspecific hybrids. *Cer Res Com.* 1977; 5 (1): 67–75.
107. Hrishi NJ, Müntzing A. Structural heterozygosity in *Secale kuprijanovii*. *Hereditas.* 1960; 46, 745–752.
108. Reddy P, Appels R, Baum BR. Ribosomal DNA spacer-length variation in *Secale* spp. (*Poaceae*). *Pl Syst Evol.* 1990; 171: 203–220.
109. Hammer K, Skolimowska E, Knupffer L. Vorarbeiten zur Monographischen Darstellung von Wildplanzen-sortimenten *Secale* L., Kulturpflanze. 1987; 33: 135–177.
110. Ren TH, Chen F, Zou YT, Jia YH, Zhang HQ, Yan BJ, et al. Evolutionary trends of microsatellites during the speciation process and phylogenetic relationships within the genus *Secale*. *Genome.* 2011; 54 (4): 316–26. <https://doi.org/10.1139/g10-121> PMID: [21491974](https://pubmed.ncbi.nlm.nih.gov/21491974/)
111. Frederiksen S, Petersen G. A taxonomy revision of *Secale* (Triticeae, Poaceae). *Nordic J Bot.* 1998; 18: 399–420.
112. Kobyljanskij VD. Studies of rye and their relations to aspects of breeding. *Vestn Selsch Nauki.* 1987; 111: 35–41.
113. Bolibok-Bragoszewska H, <http://www.biomedcentral.com/1471-2229/14/184/> - ins1Targońska M, Bolibok L, Kilian A, Rakoczy-Trojanowska M. Genome-wide characterization of genetic diversity and population structure in *Secale*. *BMC Plant Biology* 2014; 14: 184. Available from: <https://doi.org/10.1186/1471-2229-14-184> PMID: [25085433](https://pubmed.ncbi.nlm.nih.gov/25085433/)
114. Vavilov NI. On the origin of cultivated rye. *Bull Appl Bot Genet Plant Breed* 1917; 10:561–90.
115. Vavilov NI. Studies on the origin of cultivated plants. *Bull Appl Bot.* 1926;1–248.
116. Khush GS, Gurdev S, Stebbins GL. Cytogenetic and evolutionary studies in *Secale* L.: Some new data on the ancestry of *S. cereale*. *Am J Bot.* 1961; 48:723–730.
117. Khush GS. Cytogenetic and evolutionary studies in *Secale*. II. Interrelationships of the wild species. *Evolution.* 1962; 16: 484–496.
118. Broda Z, Tomkowiak A, Mikołajczyk S, Weigt D, Górski F, Kurasiak-Popowska F. The genetic polymorphism between the wild species and cultivars of rye *Secale cereale* L. *Acta Agrobotanica.* 2016; 69: 1–10. Available from: <https://doi.org/10.5586/aa.1652>