Molecular phylogenetic analysis of wild Tulipa species (Liliaceae) present in Kosovo, based on plastid and nuclear DNA sequences

Avni Hajdari[1], Bledar Pulaj[1], Corinna Schmiderer[2], Xhavit Mala[3], Kimete Lluga -Rizani[1], and Behxhet Mustafa[1]

[1]University of Prishtina
[2]University of Veterinary Medicine Vienna
[3]Sharri National Park

April 28, 2020

Abstract

Species of the genus Tulipa L. (Liliaceae) are characterized with a high variability of vegetative and floral characters, which makes the taxonomy of this genus difficult. In Kosovo the genus Tulipa is represented by eight taxa, which sometimes have been synonymized, erroneously identified, or misclassified. To investigate the phylogenetic relationships of Tulipa species originated from Kosovo, ITS and trnL-trnF DNA sequences were used. In total 55 sequences (29 ITS and 26 trnL-trnF), obtained from 14 taxa were analysed. Forty one sequences were newly generated from eight taxa collected from wild population in Kosovo and 14 sequences were obtained from GenBank. Neighbor-Joining, Maximum Parsimony and Maximum Likelihood trees from independent (ITS and trnL-trnF) and combined (ITS + trnL-trnF) datasets were conducted in PAUP. Based on sequence analyses, our sequences of Tulipa species grouped into two main clades, belonging to the subgenera Eriostemones and Tulipa, respectively. There is not sufficient genetic evidence to distinguish species of the T. scardica complex (T. scardica, T. serbica, T. albanica, T. kosovarica and T. luanica) as independent taxa. Despite the lower resolution of the trnL-trnF than the ITS dataset, both loci do not support the separation of taxa of the T. scardica complex as independent species.

Introduction

Species of the genus *Tulipa* L. (Liliaceae) have great economic, horticultural, esthetical, ecological, conservational and taxonomic importance (Veldkamp and Zonneveld, 2012). They are bulbous monocots, characterized by a high variability of vegetative and floral characters which were traditionally used to characterize species. Furthermore, the vegetative and floral characters appear to be plastic (polythetic), sometimes even within populations of a species (Christenhusz et al., 2013; Zonneveld and de Groot, 2012). Because of that, the taxonomy of this genus is considered to be difficult, despite the existence of a large body of literature (Eker et al., 2014; Zonneveld, 2009; Zonneveld and de Groot, 2012). Taking that into consideration, the classification of *Tulipa* has been revised several times (Turktas et al., 2013). The total number of *Tulipa* species still is not exactly defined and ranges from 40-150 species according to various researchers (Eker et al., 2014). In the *World Checklist of Selected Plant Families*(Govaerts, 2019), 516 names have been listed for *Tulipa* , but only 102 taxa have been accepted, while in the *Plant List* ("The Plant List," 2013) 499 names have been listed for *Tulipa* and 120 taxa have been accepted. According to Christenhusz *et al* . (2013) only 76 species are accepted. The number of *Tulipa* species native to the Balkan Peninsula is much less, varying from 15 (Hayek 1933) to 22 (Govaerts 2010). In Kosovo the genus *Tulipa* is represented by eight taxa (six species and two subspecies), belonging to the two subgenera *Eriostemones* and *Tulipa,* respectively. The subgenus *Eriostemones* is represented by *T. sylvestris* , which is represented by two subspecies, *Tulipa sylvestris* subsp.*australis* (Link) Pamp (accepted subsp.) and *Tulipa sylvestris* subsp. *sylvestris* only accepted by the *World Checklist of Selected Plant Families* (Govaerts, 2019). In Kosovo the subgenus *Tulipa* is

represented by several species: *Tulipa gesneriana* L., Sp. Pl.: 306 (1753) (Millaku et al., 2018) has a hybrid origin derived from T. *agenensis, T. armena, T. suaveolens* and others (Govaerts, 2019). *Tulipa scardica* Bornm. is distributed in Southern Kosovo and Macedonia (Mayer and Micevski, 1970). In Kosovo, it is present near the village Krivenik, close to the border of Macedonia. It is synonymized as *Tulipa gesneriana* L. ("The Plant List," 2013; Zonneveld, 2016), accepted as a species by the *World Checklist of Selected Plant Families*(Govaerts, 2019), but not accepted by Flora Eurepea (Tutin et al. 1980). *Tulipa serbica* Tatic & Krivošej is distributed on serpentine soil in the South of Serbia (community Knjaževac: Mt. Rogozna near Donja Kamenica) and Northern Kosovo (Beli Laz hill, near Ibar river) (Tatić and Krivošej, 1997). *Tulipa kosovarica* Kit Tan, Shuka & Krasniqi is distributed in Kosovo, in the serpentine area of Mirusha region at the foot of Mt. Kozniku, between Mrasori and Llapçevë villages (Shuka et al. 2012), as well as in the localities Guriç, Llapushnik, Qafë – Prush and Devë (Millaku et al., 2018). *Tulipa luanica* Millaku is distributed on limestone substrate on Mt. Pashtriku which is located in the district of Prizren, Southern Kosovo near the border with Albania (Millaku and Elezaj, 2015). *Tulipa albanica* Kit Tan & Shuka was the first time described as a new species from Albania (Kukësi district: from Kolshi to Surroj village, on serpentine slopes) (Shuka et al., 2010), but was recently found in the Kosovar village of Deva too (Millaku et al., 2018). *Tulipa scardica, T. serbica, T. albanica, T. kosovarica* and *T. luanica* belong to the *T. scardica*complex (Christenhusz et al., 2013) and share many similar morphological features (Shuka et al., 2010; Shuka and Tan Kit, 2012; Tatić and Krivošej, 1997). Because of their similarities, these species sometimes have been synonymized, erroneously identified or misclassified.

To study the taxonomy and relationships of these species, mainly their morphological characteristics and their geographical distribution has been used. Additionally, karyological analyses for *T. albanica,*(Shuka et al., 2010) and *T. luanica* (Millaku and Elezaj, 2015), as well as nuclear genome size (DNA 2C-values) for *T. albanica,*(Osmani, 2018; Shuka et al., 2010), *T. scardica* (Zonneveld, 2009), *T. kosovarica* and *T. luanica* (Osmani, 2018) have been used. However, DNA content and cytogenetic analyses are not known for all of the species present in Kosovo in order to provide information about species relationships.

DNA barcoding is a techniques, which has emerged in the last decades as powerful tool in plant systematics and became important as an inexpensive and reliable technique for phylogenetic studies (Kress, 2017). Molecular phylogenetic analysis using sequences from nuclear ribosomal DNA (nrDNA) and chloroplast DNA (cpDNA) was successfully used for phylogenetic analysis of the genus *Tulipa* . Thus, *Tulipa* DNA sequences from the ITS region (Christenhusz et al., 2013; Fay et al., 2006; Turktas et al., 2013; Yanagisawa et al., 2012) and the *trnL-trnF* region (Peterson et al., 2008) were used for phylogenetic analyses.

This work aimed to determine the phylogenetic relationships of wild-growing *Tulipa* species of Kosovo, employing the plastid*trnL-trnF* and nuclear ITS region. To the best of our knowledge, up to now no studies addressing this DNA barcodes for the Kosovar *Tulipa* species were reported.

Results

The ITS sequences (ITS1, complete 5.8S rDNA gene, ITS2 and a small part of 26S rDNA gene) in the dataset of *Tulipa* species ranged from 644 to 657 bp. The alignment of the ingroup included 45 ambiguous positions, including the outgroup another 7 positions were ambiguous. 52 positions were potentially informative. Parsimony analysis of ITS sequences included 132 potentially informative indels, MP tree lengths were 213 bp, with RI of 0.945, CI of 0.855, composite index 0.897 and 60.1 % G + C content (Table 2). The sequence lengths of ITS1 were 229-233 bp, 5.8S rDNA were 162-166 bp, ITS2 were 225-231 bp and 26S rDNA (partial) were 26 bp, while the alignments were found to be 238 bp for ITS1, 166 bp for 5.8S rDNA, 233 bp for ITS2 and 26 bp for 26S rDNA. The average G + C content for ITS1 was 59.5%, while for ITS2 it was 60.8%, respectively. Tulip samples showed an average of 141 and 143 conserved sites for ITS1 and ITS2, respectively. ITS1 contained 24 and ITS2 contained 27 potentially parsimony informative sites, respectively. The partial sequences of the 26S rDNA gene included only 26 alignment characters without indels. The average G + C content in 5.8S rDNA was 57.0% and 76.9% in 26S rDNA. The number of potentially parsimony informative sites was 1 for the 5.8S rDNA gene and none for the 26S rDNA gene, respectively.

The *trnL-trnF* sequences of *Tulipa* species in the dataset ranged from 720 to 775 bp in length. The complete alignment (including outgroup) contained 86 ambiguous positions, within the in-group 51 positions were ambiguous. Analysed sequences showed ten potentially informative characters, 47 potentially informative indels, MP tree length was 58 bp, CI was 0.769, RI 0.928, composite index 0.897 and the G + C content was 31.2% (Table 2). The *trnL-trnF* (*trnL*631-692 bp, *trnF* 57-64 bp and IGS 25 bp for each sequence, respectively) alignment included 717 bp for *trnL,* 64 bp for*trnF* and 25 bp for the IGS. G + C content was 30.8% for*trnL,* 44.5% for *trnF* and 11.8% for the IGS, respectively. Ten potentially informative characters were counted for this region.

The combined ITS + *trnL-trnF* sequences ranged from 1377 to 1430 pb in length. The complete alignment including the outgroup showed 93 ambiguous positions, within the ingroup 91 positions were ambiguous. The alignment included 1229 conserved sites, 241 variable sites, 62 potentially informative characters and 179 potentially informative indels (Table 2). The average G+C content was 44.5%, the CI was 0.840, RI was 0.940, and the composite index was 0.890.

Table 2. Data set and parsimony-based tree characteristics for ITS and*trnL-trnF* analyses.

Phylogenetic analysis

Phylogenetic trees generated from the different datasets (separated ITS and *trnL–trnF* and combined ITS + *trnL–trnF* datasets) and using different methods (NJ, MP and ML), mostly revealed similar topologies. The phylogenetic trees obtained from ITS sequences provided better resolutions compared to those generated from *trnL-trnF*sequences. ITS and *trnL-trnF* sequences were congruent (ILD test revealed p = 0.82), while generated trees from the combined ITS +*trnL–trnF* dataset were more similar to trees obtained from ITS sequences, but showed a stronger support.

ITS region

The phylogenetic analyses based on 29 ITS sequences are shown in Figure 1A (NJ), 1B (MP), and 1C (ML). The generated trees show that the analysed *Tulipa* taxa divided into two main clades. The first clade includes specimens of the subgenus *Eriostemones* (*T. sylvestris* including both subspecies) strongly supported by NJ (100% BS), MP (100% BS) and ML (95%), while the second clade includes members of the subgenus *Tulipa* (*T. albanica, T. kosovarica, T. luanica, T. scardica, T. serbica, T. ulophylla, T. tschimganica, T. suaveolens, T. julia* and *T. gesneriana*) 100% supported by all three calculation methods. In the first clade, the newly sequenced samples of *T. sylvestris* , *T. sylvestris*subsp. *sylvestris* and *T. sylvestris* subsp. *australis* separated from *T. sylvestris* subsp. *sylvestris*provided from the GenBank. In the second clade, the newly sequenced species of *T. albanica, T. kosovarica, T. luanica, T. scardica*and *T. serbica* are grouped with already published sequences of*T. ulophylla, T. tschimganica, T. suaveolens, T. julia,* and*T. gesneriana* . Within the *Tulipa* clade, in all trees the most distinct species from the newly generate sequences was *T. tschimganica* (section *Spiranthera*), followed by *T. julia, T. ulophylla* (both section *Tulipanum*) and *T. suaveolens*and *T. gesneriana* , which were the closest related to newly sequenced species (all section *Tulipa*). The separation of*T. suaveolens* and *T. gesneriana* from newly sequenced species (*T. albanica, T. kosovarica, T. luanica, T. scardica, T. serbica*) in general was strongly supported by NJ (97%) and moderately supported by MP (86%) and ML (78%). Species specific grouping of the newly sequenced samples was not supported in all trees, except *T. albanica* which was moderately supported only in NJ (82%). A clade including all samples of *T. albanica* and one sample of *T. kosovarica* (T6), *T. luanica* (T13) and *T. serbica* (T19) respectively, was weakly supported in NJ (less than 54%) and MP (less than 64%).

Figure 1. Phylogenetic tree constructed from ITS dataset of the*Tulipa* taxa.

trnL-trnF **region**

Plastid data obtained from *trnL-trnF* sequences shown in Figure 2A (NJ), 2B (MP) and 2C (ML) resemble those found in the ITS sequences analysis, but in general with lower resolution. The trees obtained from analysed sequences are divided into two major clades too. The first clade consists of members of the subgenus *Eriostemones* (*T. sylvestris* including both subspecies), which is moderately supported by NJ (86%) and ML

(74%) but not supported by MP (<50%). In the *Eriostemones* clade, the published sequence of *T. sylvestris* subsp. *sylvestris*showed sequence differences to the newly obtained sequences of the species. The second clade consists of members of the subgenus*Tulipa,* belonging to section *Tulipa* (*T. albanica, T. kosovarica, T. luanica, T. scardica, T. serbica, T. suaveolens* and*T. gesneriana*), section *Tulipanum* (*T. julia* and*T. ulophylla*) and section *Spiranthera* (*T. tschimganica*). The subgenus *Tulipa* is moderately supported by NJ (87%), MP (89%) and ML (89%). Furthermore, the specimens of subgenus *Tulipa* divides into two subclades, the first subclade consists of *T. kosovarica, T. Luanica* and *T. serbica,* which was moderately supported by NJ (86%) but not supported by MP and ML, while the second subclade consisting of *T. albanica, T. scardica, T. julia, T. ulophylla, T. Suaveolens* and *T. tschimganica* was moderately supported by NJ (84%) and ML (76%), but not supported by MP. Within the first subclade, in all phylogenetic trees *T. kosovarika* (samples T6, T7 and T8) were further divided with weak support by NJ (63%), ML (62%) and MP (67%).

Figure 2. Phylogenetic tree constructed from *trnL-trnF*dataset of the *Tulipa* taxa.

Combined ITS + *trnL-trnF* dataset

The phylogenetic trees obtained from the combined *ITS* and*trnL-trnF* dataset are shown in 3A (NJ), 3B (MP) and 3C (ML). The generated trees show that the analysed *Tulipa* species are divided into two main clades too. The first clade (specimens of subgenus*Eriostemones*) is strongly supported in all trees (99-100%) and the second clade (specimens of subgenus *Tulipa*) is strongly supported (97-100%) in all trees too. The generated trees from the combined dataset are more similar with trees obtained from ITS sequences only, but in general with stronger support. In the *Tulipa* clade, the species *T. tschimganica* (section *Spiranthera*), was the most distant species from the newly sequenced species (*T. albanica, T. kosovarica, T. luanica, T. scardica* and *T. serbica*) moderately to strongly supported in all trees, followed by *T. julia* and *T. ulophylla* (section *Tulipanum*) moderately supported by NJ (76%) and by MP (76%), but not supported by ML (<50%). The specimen of *T. suaveolens* was the closest related to the newly sequenced species, as all of these species belongs to section *Tulipa* . Within the newly sequenced species *T. albanica* separated from *T. scardica, T. serbica, T. kosovarica*and *T. luanica* within the NJ (87%) and MP (85%) trees. The other specimens showed slightly intra-specific variation (*T. kosovarica* samples T6 and T7, weakly supported by NJ (65%), MP (64%) and ML (65%)) but no species specific grouping.

Figure 3. Phylogenetic tree constructed from ITS +*trnL-trnF* data of the *Tulipa* taxa.

Discussion

In this study, we carried out molecular phylogenetic analyses using ITS and *trnL-trnF* alignments, as well as the combination of these datasets to evaluate intra-generic relationships of *Tulipa*species growing wild in Kosovo. Our data revealed the feasibility of ITS and *trnL-trnF* sequences for the phylogeny of *Tulipa*species, confirming previous findings for successfully use of ITS (Christenhusz et al., 2013; Fay et al., 2006; Turktas et al., 2013; Yanagisawa et al., 2012) and *trnL-trnF* sequences (Peterson et al., 2008). Phylogenetic trees obtained from ITS sequences provided better resolution compared with those generated from *trnL-trnF*sequences, which are in accordance with previous reports (Peterson et al., 2008; Sang et al., 2015; Turktas et al., 2013). The generated trees from the combined ITS + *trnL-trnF* dataset showed more similarities with trees obtained from ITS sequences, but in general with stronger support.

The analysed sequences of *Tulipa* species grouped into two main clades, one composed by specimens of the subgenus *Eriostemones (T. sylvestris*) and the second composed by specimens of the subgenus*Tulipa (T. albanica, T. kosovarica, T. luanica, T. scardica, T. serbica, T. ulophylla, T. tschimganica, T. suaveolens, T. julia* and*T. gesneriana*). The subgenera *Eriostemones* and*Tulipa* showed strong support for the monophyly generated by all applied methods, which agreed with previous findings (Christenhusz et al., 2013; Turktas et al., 2013). All obtained phylogenetic trees based on ITS, as well as combined ITS and *trnL-trnF* datasets, grouped the analyzed species of subgenus *Tulipa* , including section*Spiranther* (*T. tschimganica*), section *Tulipanum*(*T. julia and T. ulophylla*) and section *Tulipa* (*T. albanica, T. kosovarica, T. luanica, T. scardica* and *T. serbica*) confirming the previous classification of those species by Christenhusz*et al.* (2013). Phylogenetic analyses based on *trnL-trnF*sequences fully congruent with those provided by ITS.

Within the *Tulipa* clade, generated by the ITS and combined ITS + *trnL-trnF* datasets, the most distinct species from the newly generated sequence was *T. tschimganica,* which belongs to section *Spiranthera* , followed by *T. julia* and *T. ulophylla* (section *Tulipanum*), while *T. suaveolens* and *T. gesneriana* were the closest to the newly sequenced species as all of these species belongs to section *Tulipa* et al. 2012; Zonneveld, 2016). Furthermore, grouping of the species *T. scardica, T. serbica, T. albanica, T. kosovarica* and *T. luanica* in one subclade (section *Tulipa*), clear evidenced a close relationship between those taxa and confirmed that these species belong to a complex of species (*T. scardica*complex) distributed in the southern Balkans (Christenhusz et al., 2013), which cannot be clearly resolved by molecular methods. However, in most of the cases intra-specific variation was detected within the *T. scardica* complex indicating the presence of polymorphism in the gene pool. The polymorphism within populations of the *Tulipa*species was reported previously too (Booy and Raamsdonk, 1998; Christenhusz et al., 2013; Mayer Ernest, Micevski, 1970; Zonneveld, 2009).

Our phylogenetic analyses did not show species-specific resolution between newly sequenced specimens of the *T. scardica* complex (*T. albanica, T. scardica, T. serbica, T. kosovarica* and *T. luanica*), except *T. albanica* , which was moderately supported (Figure 1A) or moderately to strongly supported (Figure 3A).

Between those species, except molecular similarities the morphological similarities exist too. Because of their similarities *T. scardica*sometimes have been synonymized as *T. gesneriana* (Christenhusz et al., 2013; "The Plant List," 2013; Zonneveld, 2016), not accepted by Flora Eurepea (Tutin et al. 1980), but accepted as a species by the *World Checklist of Selected Plant Families* (Govaerts, 2019). Our findings based on ITS sequences confirm that *T. gesneriana* was erroneously used as a synonym for *T. scardica. Tulipa scardica*was the first species described as a new species in the *T. scardica* complex (Bornnullejr, 1923), (Mayer Ernest, Micevski, 1970). Individuals of this species show great variation in several morphological characters such as leaf form, flower colour, length of filaments and anthers in different areas of its distribution (Mayer Ernest, Micevski, 1970). *Tulipa serbica,* also belonging to this complex (Christenhusz et al., 2013), was the first time recorded at Mt Rogozna (Pavlovic 1962) and described as *T. scardica* , but was later revised and described as the new species *T. serbica* (Tatić and Krivošej, 1997). Both species (*T. scardica* and *T. serbica*) are considered to be closely related to each other. *Tulipa serbica* differs from *T. scardica* in its paler, unspotted periapt segments, pale (not blackish) staminal filaments, dull violet (not yellowish) and acute anthers (Tatić and Krivošej, 1997). Our phylogenetic results generated from ITS and ITS + *trnL-trnF*sequences did not support the separation of *T. serbica* as independent species, while results obtained by the *trnL-trnF*sequences weakly supported this opinion. Based on those findings *T. serbica* could not be confirmed as independent species. Specimens of *T. kosovarica* collected for the first time along Mrasori river (Mirusha region) at the foot of Mt Kozniku in 2010, were described in that time as *T. scardica* (Shuka et al., 2010). In 2012 the material was revised and described as *T. kosovarica*(Shuka et al. 2012). Later, this species was recorded in some other location such as Guriç, Llapushnik, Qafë - Prush, Devë (Millaku et al., 2018). *Tulipa kosovarica* shares morphological similarities with *T. scardica, T . serbica* and *T. albanica* (Shuka et al., 2010). It differs from *T. scardica* by its white or whitish perianth base that is sometimes masked by obtrullate patches of maroon and violet, while *T. albanica* differs from this species by its combination of yellow perianth bases without black blotches (Shuka et al. 2012). Phylogenetic analyses obtained by the ITS,*trnL-trnF* and ITS + *trnL-trnF* dataset did not show the divergence of *T. kosovarica* sequences from other taxa of the *T. scardica* complex. Thus phylogenetic results did not support the separation of the *T. kosovarica* as an independent species. *Tulipa luanica* is the most recent species described as member of the *T. scardica* complex (Millaku and Elezaj, 2015). According to Millaku and Elezaj (2015) *T. luanica* shares many morphological characters with *T. gesneriana, T. albanica* , *T .kosovarica* and *T . serbica* , but also differs in several characters, including the substrate (*T. luanica* grows exclusively on limestone, while other species grow only on serpentine). Based on our sequence analyses there is no genetic difference between *T. luanica* and *T. scardica* . *T. albanica,* another species of the *T. scardica* complex (Christenhusz et al., 2013). *Tulipa luanica* Millaku was recorded as a new species in Northeast Albania for the first time, but it was recently found in Kosovo too (Millaku et al., 2018). *Tulipa albanica* shows great variation in several morphological characters, for example its campanulate flowers exist in two colour forms, yellow to golden-yellow or carmine-scarlet turning deep reddish maroon,

with a dominance of the golden-yellow flowers (Shuka et al., 2010). Furthermore some individuals have an intermediate colour of yellow to reddish maroon. *Tulipa albanica*shares many morphological similarities with *T. scardica, T. serbica, T . kosovarica* and *T. luanica,* but it differs from them by its combination of yellow perianth bases without black blotches, yellow filaments and violet-purple pollen (Shuka et al. 2012, Millaku and Elezaj, 2015). Our phylogenetic results based on ITS,*trnL-trnF* and ITS + *trnL-trnF* datasets showed weak to moderate supported divergences between *T. albanica* and other taxa of the *T. scardica* complex, what does not necessarily confirm *T. albanica* as independent species, but indicate the presence of polymorphism in the gene pool of this taxon.

Our phylogenetic analyses showed that the unidentified *Tulipa species* (sample T9, Table 1) obtained from herbarium material of the Herbarium of the University Prishtina, belongs to a taxon of the *T. scardica* complex.

In the morphological analysis, the flower colour was one of the main characters used to discriminate the species of the *T. scardica*complex*,* but the flower colour appeared to be very variable within one species (Eker et al., 2014; Mayer Ernest, Micevski, 1970; Millaku and Elezaj, 2015; Shuka et al., 2010; Shuka et al. 2012; Zonneveld, 2009). Hence, it seems not to be very suitable for the classifications of *Tulipa* species (Christenhusz et al., 2013). For example, the flower colour even within one population of *T. albanica* was reported to be form yellow/golden-yellow to carmine-scarlet turning deep reddish maroon (Shuka et al., 2010), including individuals with intermediates colour. Later individuals with intermediate flower colour were reported and explained as natural hybrids of different species in sympatric distribution areas (Millaku et al. 2018). For example, individuals with mixed characters like half of the perigon base in yellow colour (an inherited trait from *T. gesneriana*) and the other half of the perigon base in white colour (an inherited trait from *T. kosovarica*) were recorded, or intermediate individuals with yellow perigon base (as it is in *T. gesneriana*), while the rest of the perigon was pink (like to *T. luanica*) (Millaku et al., 2018). According to Raamsdonk and Vries (1995) the flower colour within species may differ in two aspects, first the blotch and the blotch margins may show differences in size and in colour intensity and secondly, within some species anthocyanidins are lacking in certain accessions resulting in yellow or very light colours. Experiments based on selection of accessions obtained from natural provenances, as well as mutation experiments with radiation showed that blotch margin and flower colour can easily be influenced (Christenhusz et al., 2013).

Genome size analyses (2C) of *Tulipa* species revealed that 2C of *T. albanica* was 54.15 pg (Shuka et al., 2010) or 43.86 pg (Osmani, 2018), 45.71 pg for *T. kosovarica* and 47.49 pg for *T. luanica* (Osmani 2018) and 69 pg for *T. scardica*(Zonneveld, 2009). The incongruent results for *T. albanica*reported by Shuka et al. (2010) and Osmani (2018) were explained by the origin of the plant material by Osmani (2018). Osmani (2018) used leaves collected from plants in blooming time, from wild populations, while Shuka et al. (2010) used adult leaves germinated from seeds collected from natural populations, what seems to be an unconvincing explanation. Such differences of genome sizes within species could be correlated with differences in the habitat (Jakob et al., 2004), plant phenotype (Beaulieu et al., 2005), or probably caused by technical artefacts (Obermayer and Greilhuber, 2005). However, DNA content and cytogenetic analyses were not carried out in all of the species present in Kosovo to provide information about relationships of species within *T. scardica* complex.

Conclusions

Based on the presented sequence analyses, our sequences of *Tulipa*species grouped into two main clades, belonging to the subgenera*Eriostemones* or *Tulipa* . There is not sufficient genetic evidence to confirm species of the *T. scardica* complex (*T. scardica, T. serbica, T. albanica, T. kosovarica* and *T. luanica*) as independent taxa. Despite the lower resolution of the*trnL-trnF* dataset, it does not support the separation of the taxa related to the *T. scardica* complex as independent species too. Genetic differences between taxa of the *T. scardica* complex show intra-specific variation, indicating the presence of polymorphism in the gene pool. Further analysis with more extensive sampling, using additional markers, as well as the determination of nuclear genome size (used DNA 2C-values) will be necessary for better understanding of the natural variability within the taxa of the *T. scardica* complex.

Materials and methods

Plant material

Eight taxa (six species and two subspecies) of the genus *Tulipa*were collected from their wild populations in April to May 2017, 2018 and 2019. All *Tulip* species were collected in Kosovo, except *Tulipa albanica* , which was collected in Albania. One unidentified plant specimen of *Tulipa* sp. (sample T9, Table 1) was obtained from herbarised material provided by the Herbarium of the University Prishtina. *Tulipa kosovarica* (locations Goriç and Koznik) and *T. luanica* (locations Pashtrik and Qafë Prush) were collected from two different localities each. Plant specimens were herbarised and part of young leaves were dried in silica gel for DNA extraction. The voucher specimens were deposited at the Herbarium of the University Prishtina, Kosovo and the Emory University Herbarium, Atlanta, USA. Detailed sample information is given in Table 1.

Table 1 . Basic characteristics of the collection sites, voucher information, and GenBank accession numbers of the *Tulipa*specimens used for this study

DNA extraction, polymerase chain reaction (PCR) and sequencing

Total genomic DNA was extracted from silica gel-dried materials or herbarium specimens using the DNeasy® Plant Mini Kit (Qiagen Hilden, Germany) according to the manufacturer's instructions. The DNA quality was checked using agarose gel electrophoresis with 1.0% agarose gels containing 0.4 x PeqGreen (VWR, Erlangen, Germany) for 40 min at 120 V, which was documented using microDOC system with UV transilluminator (Cleaver Scientific LTD, Rugby, Warwickshire, UK) using 312 nm wavelength.

Extracted DNA was 1:50 diluted with deionized water and then used for PCR. The nuclear internal transcribed spacer region (ITS) and the chloroplast *trnL-trnF* intergenic spacer were amplified and then sequenced from 23 samples of six species and two subspecies. For a 15 µL PCR reaction, 1 µL of diluted genomic DNA (equivalent to approx. 1–50 ng) was added to 14 µL master mix containing 1 × PCR buffer B, 2.5 mM $MgCl_2$, 130 µM dNTP mix, 0.6 U Taq HOT FIREPol DNA Polymerase (all reagents from Solis Biodyne, Tartu, Estonia) and 300 nM forward (ITS5 [5'-GGAAGGAGAAGTCGTAACAAGG-3'; White at al., 1990] or c [5'-CGAAATCGGTAGACGCTACG-3'; Taberlet et al., 1991]) and reverse primers (ITS4 [(5'-TCCTTCCGCTTATTGATATGC-3; White et al., 1990] or f [5'-ATTTGAACTGGTGACACGAG-3'; Taberlet et al., 1991]) (Sigma Aldrich, Taufkirchen, Germany). The PCRs were performed in a MIC qPCR cycler (Biomloceular systems, Upper Coomera, Australia). PCR amplifications were performed with an initial denaturation step at 95 °C for 14:30 min, followed by 40 cycles at 95/58/72 °C for 30/30/90 s, and a final elongation step of 7 min at 72 °C. The amplified PCR fragments (2 µL of PCR products) were checked using electrophoresis in 1% agarose gels (low melting point agarose, Sigma Aldrich, Taufkirchen, Germany), using similar conditions as described above for genomic DNA.

Exonuclease I from *E. coli* 20 U/µl (EXO I) and Thermosensitive Alkaline Phosphatase 1 U/µl (FastAP) (Thermo Fisher Scientific Baltics, Vilnius, Lithuania) were premixed in the ratio 1:4 and stored in the freezer. 13 µL PCR products were mixed with 1.3 µL EXO I and FastAP mixture and incubated at 37 °C for 15 min and 15 min in 85 °C. Purified PCR products were diluted with distilled water and admixed with sequencing primers according to the requirements of the sequencing company. Sequencing was performed by Microsynth Austria (Vienna, Austria) using Applied Biosystems 3730xl 96 capillary DNA analyzer (Thermo Fisher Scientific). Every Sequence was manually edited with CHROMAS vers. 2.6.6 (Technelysium, South Brisbane, Australia) and aligned with MEGA X software (Kumar et al. 2018). Edited sequences were subjected to BLAST searches for preliminary analysis (Altschul et al., 1990). Newly generated sequences were submitted to the National Center for Biotechnology Information (NCBI). GenBank accession numbers for all sequences are given in Table 1.

Phylogenetic analyses

In total 55 sequences obtained from 14 taxa were analysed, 41 of them were newly generated sequences provided from eight *Tulipa* taxa (six species and two subspecies) collected from wild populations in Kosovo and

14 sequences were obtained from GenBank (Table 1). The ITS sequences for *Tulipa ulophylla* (HF952978), *T. tschimganica*(HF952976), *T. sylvestris* subsp. *sylvestris* (HF952974),*T. suaveolens* (MK334468), *T. julia* (HF952964), *T. gesneriana* (MK335217, MK335224) and the *trnL-trnF* sequences for*T. ulophylla* (HF953003), *T. tschimganica* (HF953001),*T. sylvestris* subsp. *sylvestris* (HF952999), *T. suaveolens* (HF952998), *T. julia* (HF952989) were obtained from GenBank. The trees were rooted using *Lilium martagon* (obtained from GenBank: *trnL-trnF* KF850988 and ITS KX865057) as outgroup.

ITS and *trnL-trnF* sequence of most of taxa were amplified and then sequenced from three specimens for each species, while the *T. kosovarica* (locality Goriç) and *T. luanica* (locality Qafë Prush) were amplified and sequenced successfully from two specimens per species. Because of the failure of the amplification of some specimens (ITS T8 and T10; trnL-trnF T9, T16 and T18), some species were represented by only one or two sequences.

The sequences were aligned using MEGA X software (Kumar et al. 2018). The datasets were loaded into PAUP 4.0 (Swofford, 2002) and then analysed as separate (ITS and *trnL-trnF*) and combined (ITS +*trnL–trnF*) datasets. For ITS analyses, in total 29 sequences were aligned to determine sequence statistics, 21 of them were newly generated and eight were obtained from GenBank, while for*trnL-trnF* statistical analyses included 26 sequences (20 newly generated and six obtained from GenBank) (Table 2). The combined (ITS +*trnL–trnF*) dataset included 24 sequences obtained from the taxa which were present in both datasets and were used for compatibility testing (ILD test). The ILD test revealed no incongruence between ITS and *trnL-trnF* sequences ($p = 0.82$). Phylogenetic trees were constructed using Neighbor-Joining (NJ), Maximum Parsimony (MP) and Maximum Likelihood (ML) methods, conducted in PAUP 4.0 (Swofford, 2002). In the phylogenetic trees, bootstrap scores <50% were not taken in consideration, scores between 50% and 74% were defined as weak support, scores between 75% and 89% as moderate support and scores >90% BS as strong support.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal of Molecular Biology. 1990: 215: 403–410.

Beaulieu JM, Moles AT, Leitch IJ, Bennett MD, Dickie JB, Knight CA, et al. Correlated evolution of genome size and seed mass. 2005;422–37.

Booy G, Raamsdonk LWD Van. Variation in the enzyme esterase within and between Tulipa species ; usefulness for the analysis of genetic relationships at different taxonomical levels. Biochem Syst Ecol. 1998;26(2):199-224

Bornnullejr J. Uber eine neue Tulpe der Flora Mazedoniens. Feddes Repert. 1923; 19:199—200.

Christenhusz MJM, Govaerts R, David JC, Hall T, Borland K, Roberts PS, et al. Tiptoe through the tulips - cultural history, molecular phylogenetics and classification of *Tulipa* (Liliaceae). Bot J Linn Soc. 2013; 172(3):280–328.

Eker I Il, Babac MTI, Koyuncu M. Revision of the genus *Tulipa* L. (Liliaceae) in Turkey. Vol. 157, Phytotaxa. 2014. 1–112 p.

Fay MF, Chase MW, Ronsted N, Devey DS, Pillon... Y. Phylogenetics of Liliales: summarized evidence from combined analyses of five plastid and one mitochondrial loci. Aliso [Internet]. 2006;22(November 2015):559–65.

Govaerts R. World Checklist of Selected Plant Families: Royal Botanic Gardens, Kew [Internet]. Access. 2019 [cited 2019 Aug 13]. Available from: http://wcsp.science.kew.org/qsearch.do

Govaerts, R., Dransfield, J., Zona, S., Hodel, D.R. & Henderson, A. World Checklist of Arecaceae. Facilitated by the Royal Botanic Gardens, Kew. Published on the Internet; http://wcsp.science.kew.org/ Retrieved 13 August 2019.

Hayek A. Prodromus Florae peninsulae Balcanicae, Band 3 (Monocotyledoneae). Repertorium specierum Novarum Regni Vegetabilis, 1933; 30:472

Jakob SS, Meister A, Blattner FR. The Considerable Genome Size Variation of Hordeum Species (Poaceae) Is Linked to Phylogeny, Life Form , Ecology , and Speciation Rates. Mol Biol Evol. 2004; 21 (5), 860-9

Kress WJ. Plant DNA barcodes: Applications today and in the future. J Syst Evol. 2017; 55(4):291–307.

Kumar S., Stecher G., Li M., Knyaz C., and Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Molecular Biology and Evolution. 2018. 35:1547-1549.

Mayer Ernest, Micevski K. Zur Taxonomie und Chorologie von *Tulipa scardica* Bornm. Feddes Repertorium. 1970; 591–8.

Millaku F, Elezaj I. *Tulipa luanica* (Liliaceae), a New Species from Southern Kosovo . Ann Bot Fenn. 2015;52(5–6):315–20.

Millaku F, Elezaj ISA, Berisha N. Sympatric area and ecology of some *Tulipa* species in the west Balkan Peninsula. Thaiszia J Bot. 2018;28(1):35–47.

Obermayer R, Greilhuber J. Does genome size in Dasypyrum villosum vary with fruit colour? Hredity (Edinb), 2005:91–5.

Osmani M. Disa adaptime citogjenetike dhe biokimike te tre llojeve te gjinise *Tulipa* (*T. albanica, T. kosovarica* dhe *T. luanica*) ne habitatet e tyre natyrore. PhD dissertation. University of Prishtina; 2018.

Pavlovic, Z. Karakteristicni elementi serpentinske flore Srbije. Glasn. Prir. Muz. Beogradu, 1962; 8: 3-20.

Peterson A, Levichev IG, Peterson J. Systematics of *Gagea* and *Lloydia* (Liliaceae) and infrageneric classification of Gagea based on molecular and morphological data. Mol Phylogenet Evol. 2008; 46(2):446–65.

Raamsdonk LWD Van, Vries T DE. Species relationships and taxonomy in *Tulipa* subg . *Tulipa* (Liliaceae). 1995;195:13–44.

Sang T, Crawford DJ, Stuessy TF. Chloroplast DNA Phylogeny, Reticulate Evolution, and Biogeography of *Paeonia* (Paeoniaceae). Am J Bot. 1997; 84 (9),1120–36.

Shuka L, Tan K, Siljak-Yakovlev S. Tulipa albanica (Liliaceae), a new species from northeastern Albania. Phytotaxa. 2010;10(1):17.

Shuka Lulezim, Tan Kit KE. *Tulipa kosovarica* (Liliaceae), a new species of tulip from Kosovo. Phytotaxa. 2012; 9:1–9.

Swofford, D. L. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). 2002. Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet P., Gielly L., Pautou G., Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 1991; 17:1105–1109.

Tatic B, Krivosej Z. *Tulipa serbica* (Liliaceae), a new species from Serbia. Bocconea. 1997;5(2):733–6.

The Plant List. Choice Rev Online. 2013; [cited 2019 Aug 15]. Available from: http://www.theplantlist.org/.

Turktas M, Metin OK, Başţuğ B, Ertuğrul F, Saraç YI, Kaya E. Molecular phylogenetic analysis of *Tulipa* (Liliaceae) based on noncoding plastid and nuclear DNA sequences with an emphasis on Turkey. Bot J Linn Soc. 2013; 172(3):270–9.

Tutin TG, Heywood VH, Burges NA, Valentine DH. Flora Europaea. Flora Europaea. 1980. 381–382 p.

Veldkamp JF, Zonneveld BJM. The infrageneric nomenclature of *Tulipa* (Liliaceae). Plant Syst Evol. 2012; 298(1):87–92.

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, and T. J. White [eds.], PCR protocols: A guide to methods and applications, 315–322. Academic Press, San Diego, California, USA.

Yanagisawa R, Kuhara T, Nishikawa T, Sochacki D, Marasek-Ciolakowska A, Okazaki K. Phylogenetic analysis of wild and garden tulips using sequences of chloroplast DNA. Acta Hortic. 2012; 953:103–10.

Zonneveld BJM, de Groot JJ. *Tulipa kolbintsevii* Zonn., a new species from Eastern Kazakhstan. Plant Syst Evol. 2012; 298(7):1293–6.

Zonneveld BJM. The systematic value of nuclear genome size for " all " species of Tulipa L . (Liliaceae). Plant Syst Evol. 2009; 281(1):217-245

Zonneveld BJM. The systematic value of nuclear genome size for "all" species of *Tulipa* L. (Liliaceae). Plant Syst Evol. 2009; 281(1–2):217–45.

Table 1 . Basic characteristics of the collection sites, voucher information, and GenBank accession numbers of the *Tulipa* samples used for this study

Potential species	Sequence_-ID	Collection locality	Country	Longitude	Latitude	Altitude	ITS accession number	trnL-trnF accession number	Un Pr- ish her acc no
T. albanica	T._albanica_T1 (yellow flower)	Surroj	Albania	42° 2.744'N	20° 20.037'E	622	MN336199	MN446897	000
T. albanica	T._-albanica_-T2 (reddish maroon flower)	Surroj	Albania	42° 2.744'N	20° 20.037'E	622	MN336200	MN446898	000
T. albanica	T._-albanica_-T3 (reddish maroon /yellow flower)	Surroj	Albania	42° 2.744'N	20° 20.037'E	622	MN336201	MN446899	000
T. kosovarica	T._-kosovarica_-T4	Goriç	Kosovo	42° 26.689'N	20° 45.337'E	659	MN336202	MN446900	000
T. kosovarica	T._-kosovarica_-T5	Goriç	Kosovo	42° 26.689'N	20° 45.337'E	659	MN336203	MN446901	000

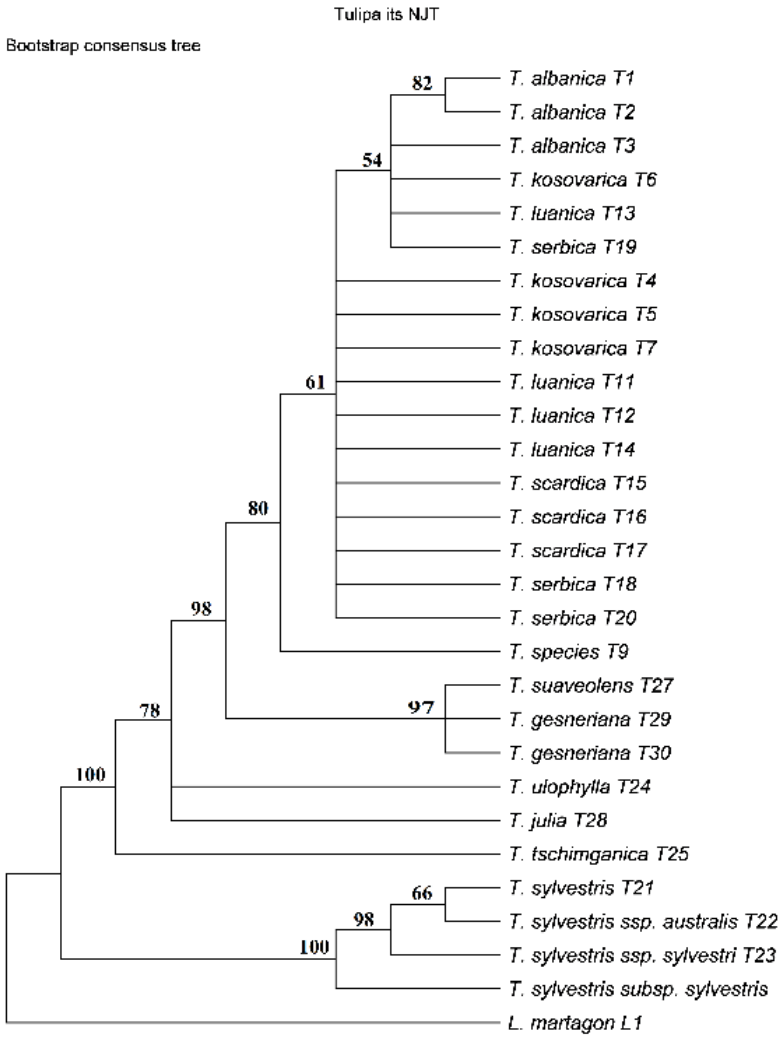
Potential species	Sequence_-ID	Collection locality	Country	Longitude	Latitude	Altitude	ITS accession number	trnL-trnF accession number	Un Pr- ish her acc no
T. kosovarica	T._-kosovarica_-T6	Koznik	Kosovo	42° 30.334'N	20° 33.987'E	425	MN336204	MN446902	000
T. kosovarica	T._-kosovarica_-T7	Koznik	Kosovo	42° 30.334'N	20° 33.987'E	425	MN336205	MN446903	000
T. kosovarica	T._-kosovarica_-T8	Koznik	Kosovo	42° 30.334'N	20° 33.987'E	425	///	MN446904	000
T. species	T._-species_-T9	Krojmir	Kosovo	///	///	///	MN336206	///	000
T. luanica	T._luanica_-T10	Pashtrik	Kosovo	42° 14.966'N	20° 30.399'E	1041	///	MN446905	000
T. luanica	T._luanica_-T11	Pashtrik	Kosovo	42° 14.966'N	20° 30.399'E	1041	MN336207	MN446906	000
T. luanica	T._luanica_-T12	Pashtrik	Kosovo	42° 14.966'N	20° 30.399'E	1041	MN336208	MN446907	000
T. luanica	T._luanica_-T13	Qafë Prush	Kosovo	42° 18.275'N	20° 23.529'E	580	MN336209	MN446908	000
T. luanica	T._luanica_-T14	Qafë Prush	Kosovo	42° 18.275'N	20° 23.529'E	580	MN336210	MN446909	000
T. scardica	T._-scardica_-T15	Krivenik	Kosovo	42° 6.254'N	21° 14.958'E	575	MN336211	MN446910	000
T. scardica	T._-scardica_-T16	Krivenik	Kosovo	42° 6.254'N	21° 14.958'E	575	MN336212	///	000
T. scardica	T._-scardica_-T17	Krivenik	Kosovo	42° 6.254'N	21° 14.958'E	575	MN336213	///	000
T. serbica	T._serbica_-T18	Serboc	Kosovo	42° 58.067'N	20° 49.757'E	596	MN336214	MN446911	000
T. serbica	T._serbica_-T19	Serboc	Kosovo	42° 58.067'N	20° 49.757'E	596	MN336215	MN446912	000

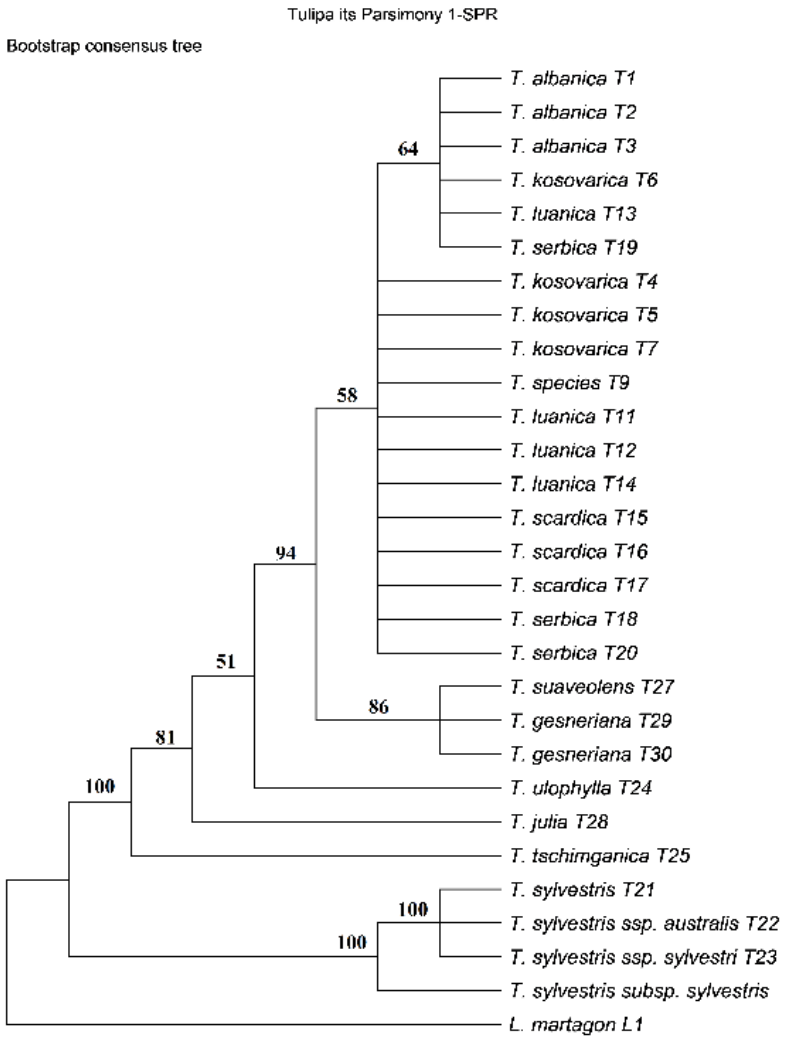
Potential species	Sequence_ID	Collection locality	Country	Longitude	Latitude	Altitude	ITS accession number	trnL-trnF accession number	Un Pr- ish her acc no
T. serbica	T._serbica_-T20	Serboc	Kosovo	42° 58.067'N	20° 49.757'E	596	MN336216	MN446913	000
T. sylvestris	T._-sylvestris_-T21	Goriç	Kosovo	42° 26.747'N	20° 45.293'E	665	MN336217	MN446914	000
T. sylvestris ssp. *australis*	T._-sylvestris_-ssp._-australis_-T22	Devë	Kosovo	42° 19.950'N	20° 20.517'E	700	MN336218	MN446915	000
T. sylvestris ssp. *sylvestris*	T._-sylvestris_-ssp._-sylvestris_-T23	Devë	Kosovo	42° 19.950'N	20° 20.517'E	700	MN336219	MN446916	000
T. ulophylla	T._ulophylla_-T24	///	///	///	///	///	HF952978	HF953003	//
T. tschimganica	T._-tschimganica_-T25	///	///	///	///	///	HF952976	HF953001	//
T. sylvestri ssp. sylvestris	T._-sylvestris_-subsp._-sylvestris_-T26	///	///	///	///	///	HF952974	HF952999	//
T. suaveolens	T._-suaveolens_-T27	///	///	///	///	///	MK33446	HF952998	//
T. julia	T._-julia_-T28	///	///	///	///	///	HF952964	HF952989	//
T. gesneriana	T._gesneriana_-T29	///	///	///	///	///	MK335217	///	//
T. gesneriana	T._gesneriana_-T30	///	///	///	///	///	MK335224	///	//

Potential species	Sequence_-ID	Collection locality	Country	Longitude	Latitude	Altitude	ITS accession number	trnL-trnF accession number	Un Pr- ish her acc no
Lilium martagon	L._-martagon_-L01	///	///	///	///	///	KX865057	KF850988	//

Table 2. Data set and parsimony-based tree characteristics for ITS and*trnL-trnF* analyses.

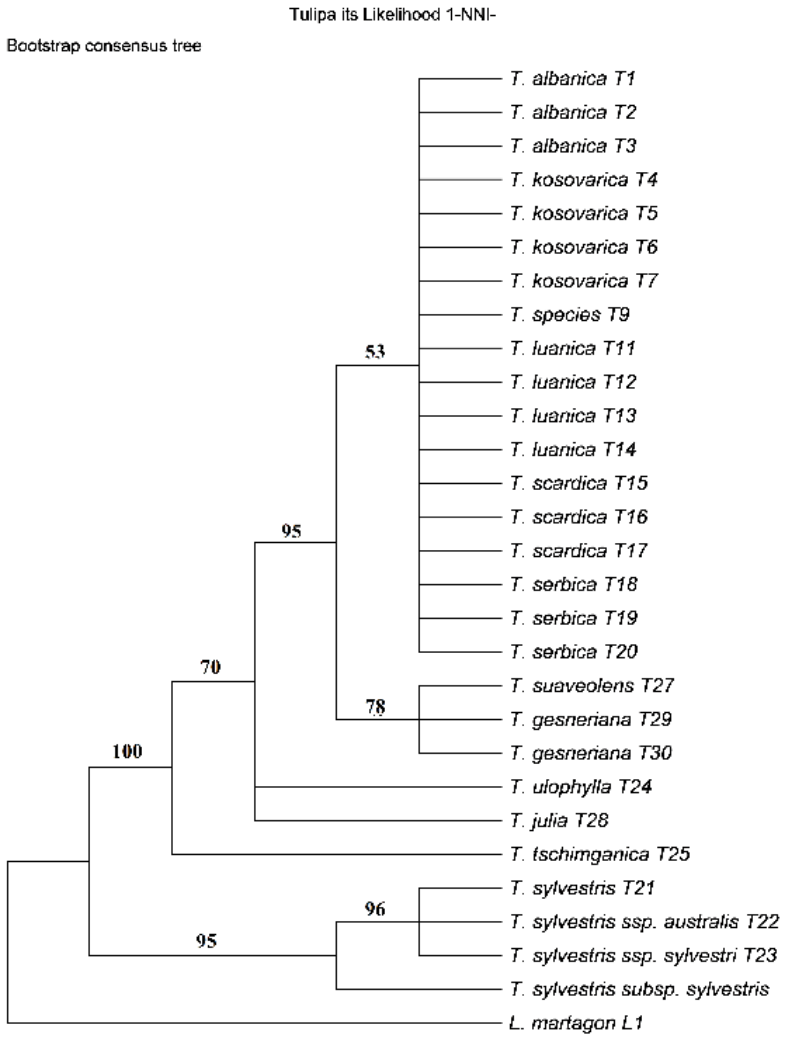
Parameters	ITS	*trnL-trnF*	Combined ITS + *trnL-trnF*
No. taxa	14	12	12
No. sequences	29	26	24
Alignment length (bp)	664	806	1471
Sequence minimum length (bp)	644	720	1377
Sequence maximum length (bp)	657	775	1430
Number of ambiguous positions: ingroup	45	51	91
Number of ambiguous positions: outgroup	7	86	93
Conserved characters	480	749	1229
Variable characters	184	57	241
Potentially informative characters	52	10	62
Number of potentially informative indels	132	47	179
MP tree length	213	60	273
CI (informative characters only) (consistency index)	0.855	0.769	0.840
RI (retention index)	0.945	0.928	0.940
composite index	0.897 (0.808)	0.880 (0.714)	0.890
G + C contents	60.1%	31.2%	44.5%





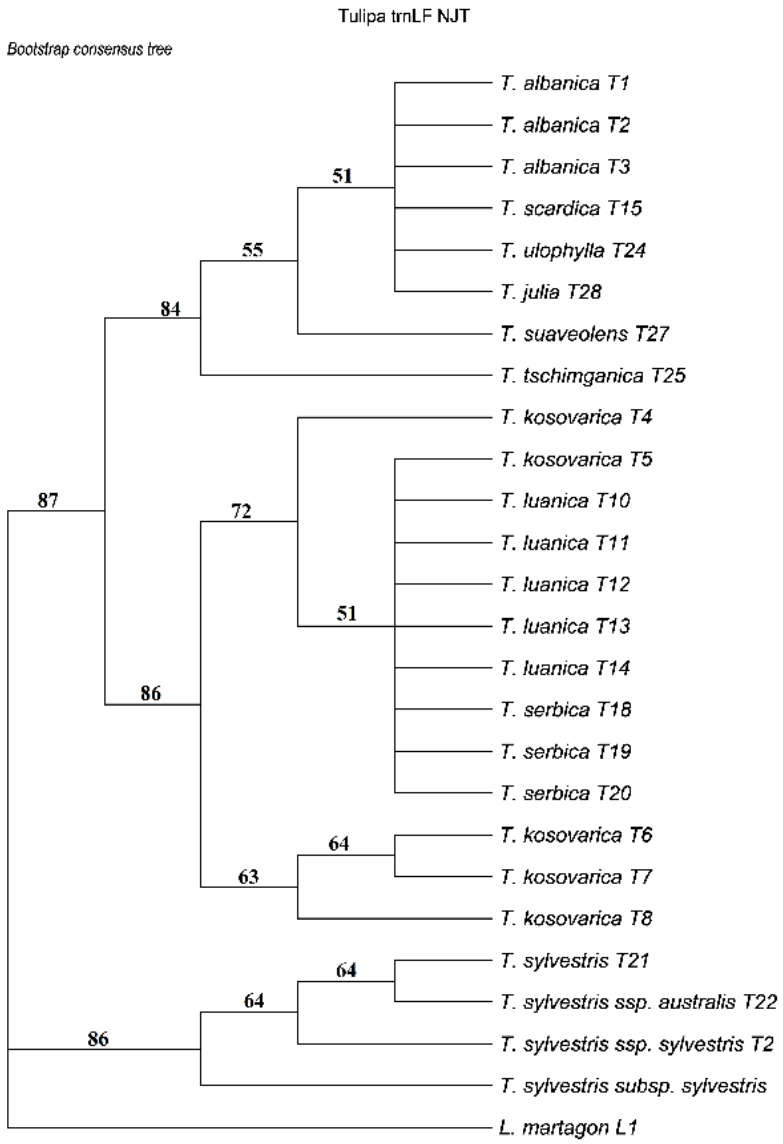
A.

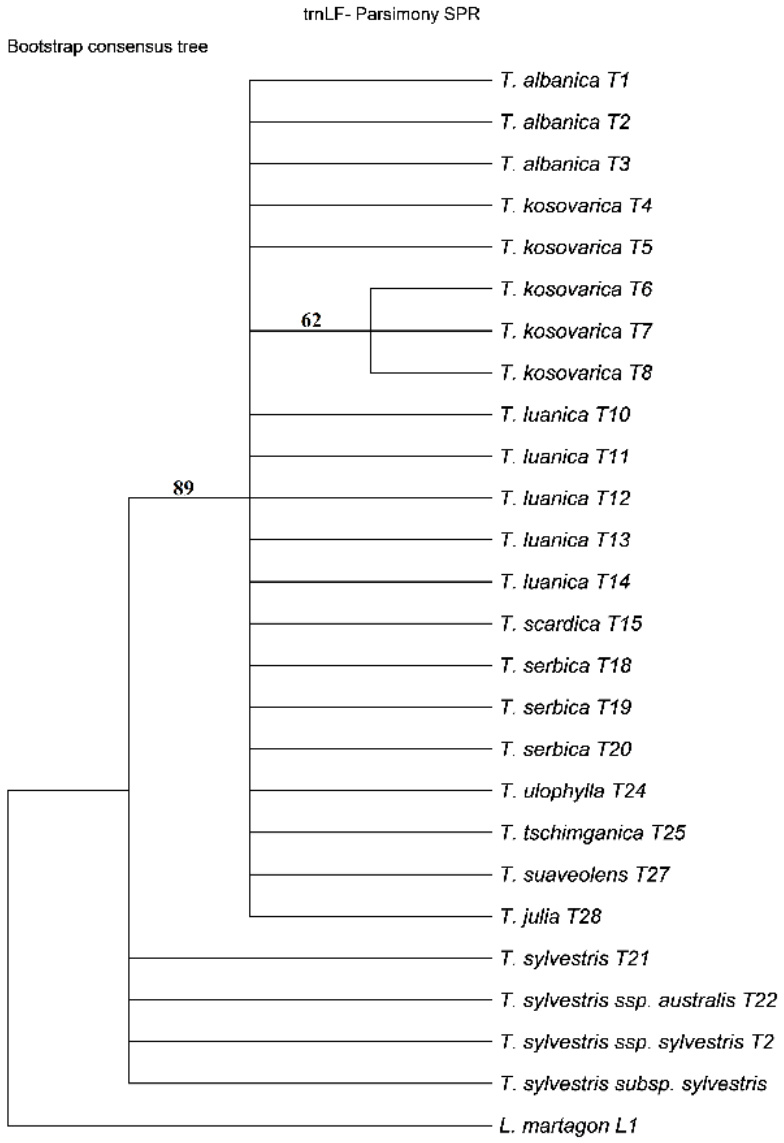
B.



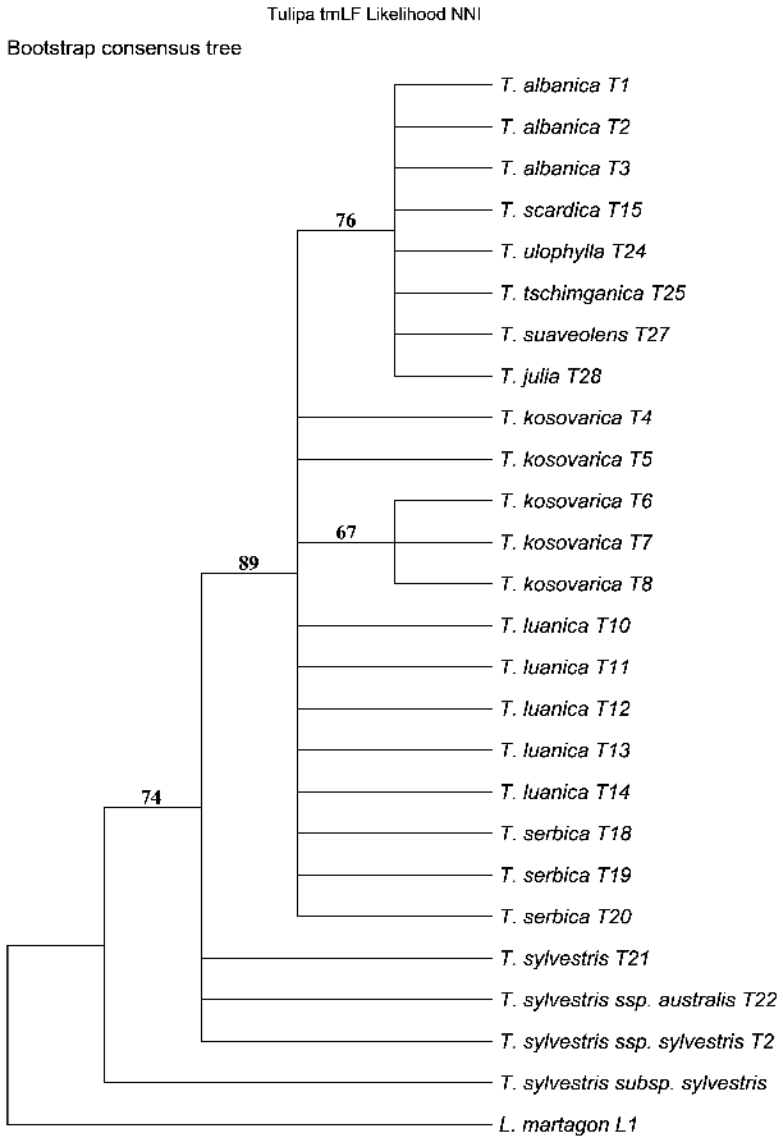
C.

Figure 1. Different phylogenetic trees based on ITS sequences including Bootstrap values based on 1000 replicates. A. Neighbor-joining tree. B. Maximum parsimony tree. C. Maximum likelihood tree.



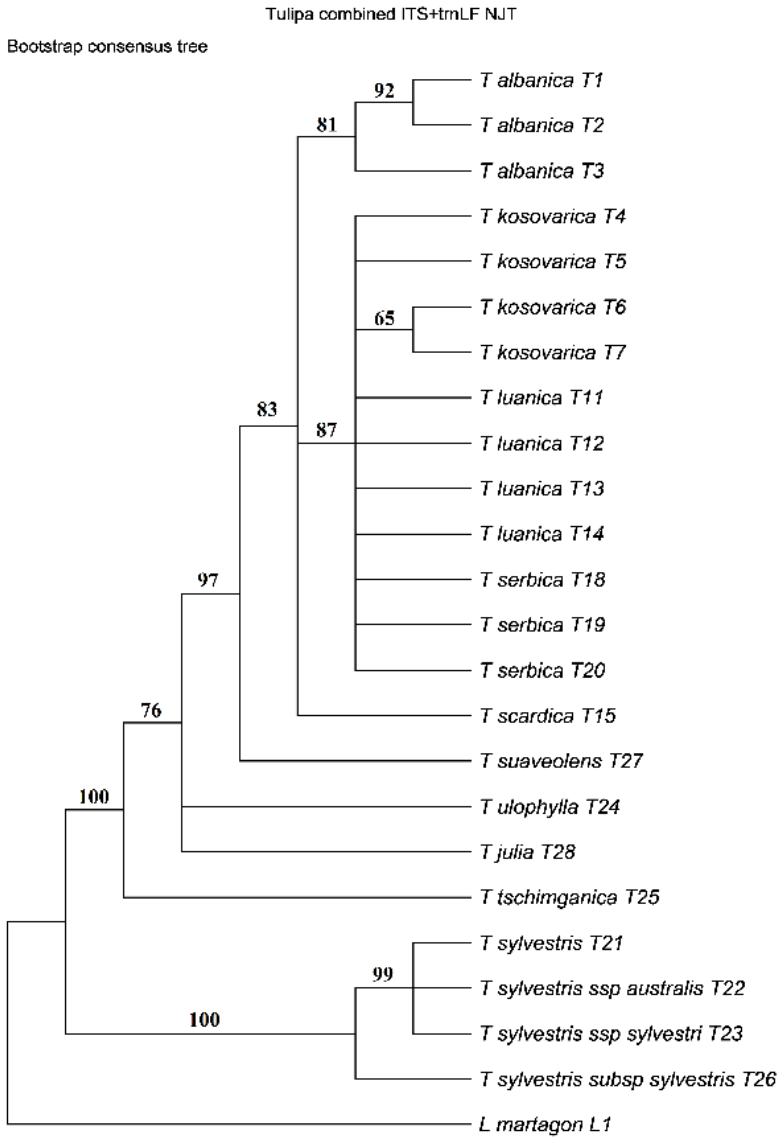


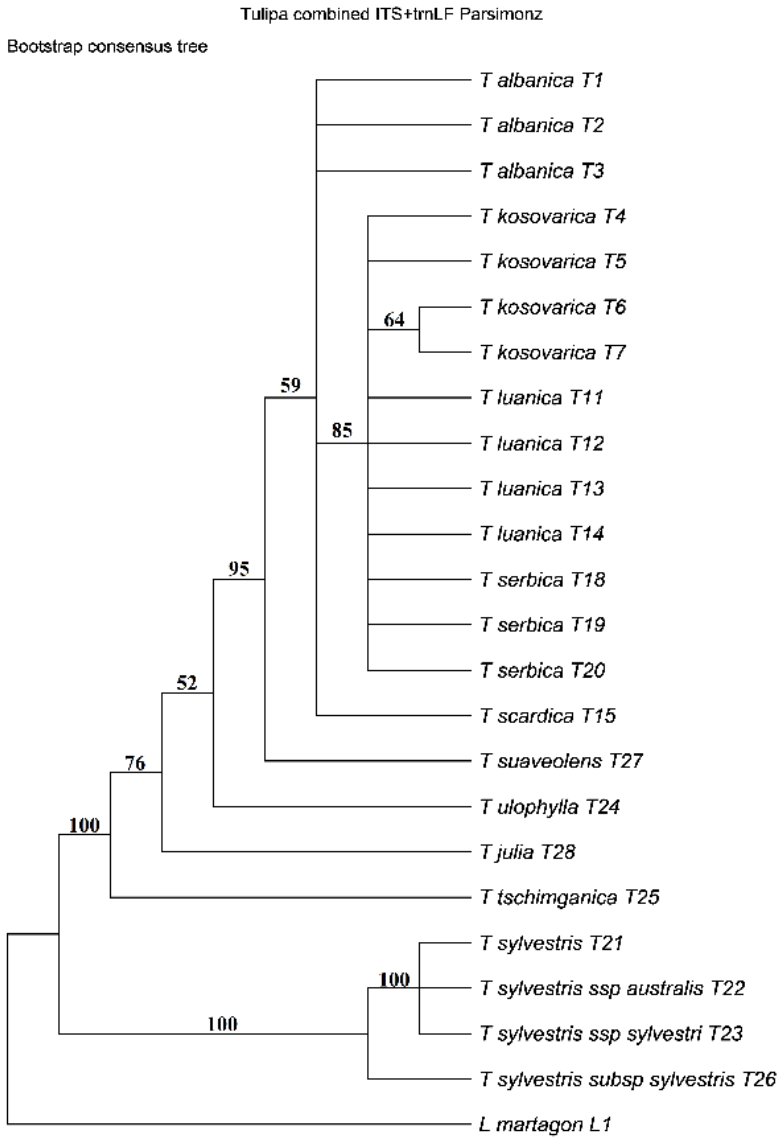
A. B.



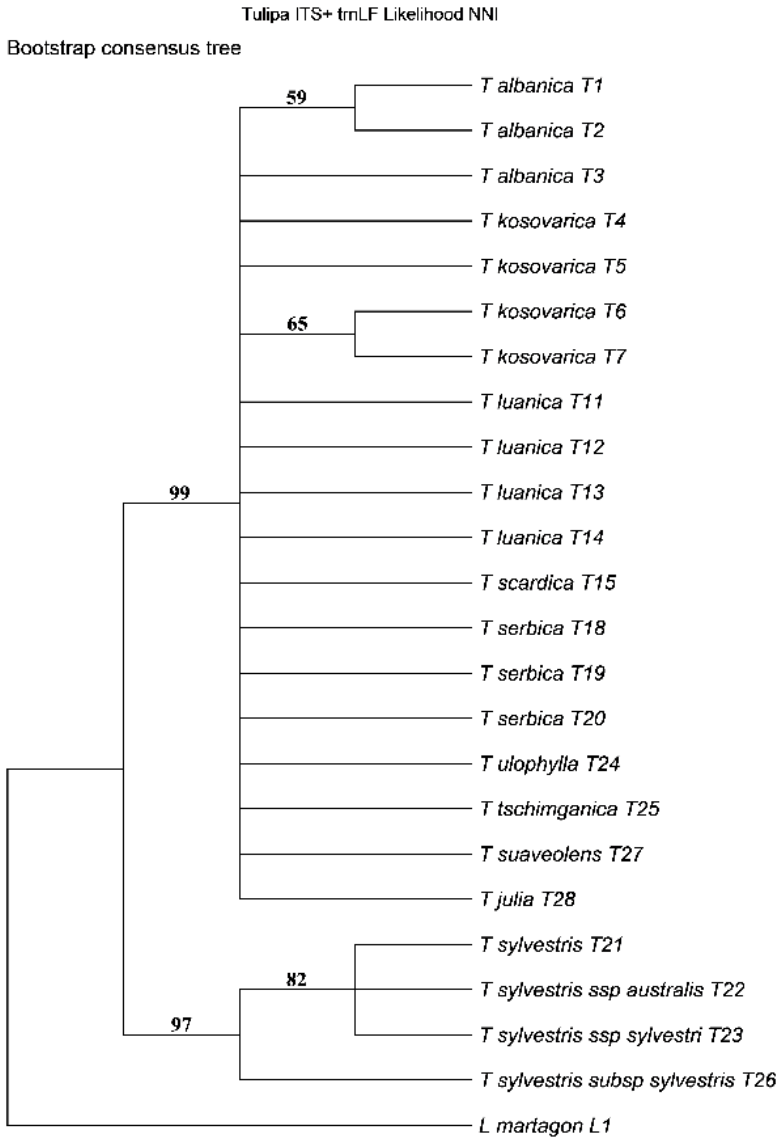
C.

Figure 2. Different phylogenetic trees based on*trnL-trnF* sequences including Bootstrap values based on 1000 replicates. A. Neighbor-joining tree. B. Maximum parsimony tree. C. Maximum likelihood tree.





A. B.



C.

Figure 3 . Different phylogenetic trees based on a combined ITS+*trnL-trnF* sequence set including Bootstrap values based on 1000 replicates. A. Neighbor-joining tree. B. Maximum parsimony tree. C. Maximum likelihood tree.