universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Genome evolution of European *Polygonatum* Mill. (Asparagaceae) species"

verfasst von / submitted by

Daniela Hanusch BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree program code as it appears on the student record sheet:	UA 066 832
Studienrichtung lt. Studienblatt / degree program as it appears on the student record sheet:	Masterstudium Botanik
Betreut von / Supervisor:	Univ-. Prof. Mag. Dr. Hanna Schneeweiss, Privatdoz.

Acknowledgments

A master thesis is a lot of work and clearly, I would have never been able to accomplish this project without help. Therefore, I would especially like to thank Hanna Schneeweiss for sharing her knowledge with me and supporting me whenever needed. Her patience was a relief in times of sorrows and low energy.

I would also like to thank Vildana Suljevic and Eva Temsch for introducing me to lab work and different methods. Working with both of you kept me motivated and our talks about various topics made time fly even faster. The openness about personal experiences made me feel welcomed and comfortable.

Thank you also to the other members of our group, Nadine Jany, Martin Piringer, Nahuel Palombo and Carolina Carrizo Garcia; working is much more fun with nice people around.

Thank you to Gerald Schneeweiss for helping me with the first identifications of my *Polygonatum* plants. The information about ecology and small morphological differences were very valuable and helpful for the field work.

And finally, I would like to thank my family and friends for collecting plants with me, supporting me, and motivating me through this intense – and sometimes difficult – time. The support from each one of you kept me going, helped me relax, and gain new energy.

Table of Contents

Introduction

More than 3000 species of angiosperms grow in Austria due to its high levels of geological and climatic diversity. One of the families represented in the Austrian flora is the family Asparagaceae, named after the culinary valuable asparagus. The family comprises more than 100 genera (Christenhusz and Byng, 2016), one of those being *Polygonatum*, formerly placed in the families Liliaceae (Engler, 1888), Ruscaceae (Dahlgren et al., 1985) or Convallariaceae (e.g. Rudall et al., 2000). Species of the genus *Polygonatum* are often used as ornamental and medicinal plants.

The genus *Polygonatum*

The genus *Polygonatum* of the tribe Polygonateae comprises about 60 species (Zhao et al., 2014a; Wang et al., 2016). The genus is restricted to the Northern hemisphere with its diversification center in south-western China (Meng et. al, 2014; Zhao et al., 2014a; Wang et al., 2016) and its highest diversity in the eastern Himalaya and the Indo-Burma biodiversity hotspots (Wang et al., 2016; Floden and Schilling, 2018). The geographic distribution and species diversity allowed formulation of the hypothesis of an origin of the genus within or near this area. The genus has most likely diversified and evolved in southern East Asia, followed by a split in northern and southern lineages as a result of a biogeographical barrier formed in the Miocene (Meng et al., 2014; Wang et al., 2016). The dispersal to North America, Central Asia and Europe has likely occurred more recently and regionally from East Asia (Wang et al., 2016).

The monophyly of the tribe Polygonateae (Rudall et al., 2000) as well as the genus (Tamura et al., 1997; Meng et al., 2014) was inferred in several phylogenetic studies. Analyses of the relationships of *Polygonatum* and related genera inferred *Polygonatum* to be a sister group to a clade including the genera *Heteropolygonatum* and *Disporopsis* (Meng et al., 2014). The genus *Polygonatum* comprises two sections, *Polygonatum* and *Verticillata* (Tamura et al., 1997; Meng et al., 2014). The section *Polygonatum* was found to be monophyletic, whereas the section *Verticillata* was inferred to be polyphyletic (Meng et al., 2014).

Four *Polygonatum* species are recognized In Europe*: P. latifolium, P. multiflorum*, *P. odoratum* and *P. verticillatum* (Janchen 1956-1960). All of these four species occur also in Austria and are distinct morphologically, ecologically, and karyologically. Three of the species belong to sect. *Polygonatum,* whereas *P. verticillatum* was placed in sect.

Verticillata. The relationships between the four European species have not yet been resolved, as most phylogenetic studies focused mostly on Asian *Polygonatum* species with only some of the European species included in the analyses. Morphologically, *P. verticillatum* was most distinct, which has also been shown in phylogenetic studies including Asian accessions of *P. multiflorum*, *P. odoratum* and *P. verticillatum* (Meng et al., 2014). Karyologically, *P. odoratum* is relatively easily distinguishable from *P. latifolium* and *P. multiflorum* (Suomalainen, 1947; Therman, 1953), indicating a closer relationship between the latter two species.

P. latifolium, P. odoratum and *P. multiflorum* were placed in the sect. *Polygonatum*. The species in this section are characterized by the presence of alternate leaves, thick staminal filaments, and a perforated pollen exine (Meng et al., 2014). They have been reported to have base chromosome numbers of x = 9, 10 or 11 (Meng et al., 2014).The only species of the sect. *Verticillata* growing in Austria and Europe is *P. verticillatum*. Species within this section are characterized by the presence of opposite or verticillate leaves, slender staminal filaments, predominantly reticulate pollen exines and a karyotype with a base chromosome number of x = 14 or 15 (Meng et al., 2014).

The morphology of the European *Polygonatum* species is typical for their respective sections (see morphology). Phylogenetic studies and ancestral character state reconstructions allowed the inference of the ancestral states of the leaf arrangements on the stem (phyllotaxy), one of the most important diagnostic characters defining the sections. Alternate phyllotaxy was inferred to represent ancestral state whereas opposite and whorled/verticillate arrangements were inferred as derived states (Meng et al., 2014).

European species of the genus *Polygonatum* are less known than their Asian relatives. One of the reasons may be that various *Polygonatum* species have been used as medicinal plants especially in the Traditional Chinese Medicine (TCM). More than 37 species and one variety of *Polygonatum* have been or still are used for medicinal purposes, especially their rhizomes (Zhao et al., 2014b). *P. odoratum* is used in TCM to treat osteoporosis, feebleness, fatigue, as well as diabetes and lung disorders (Zhao et al., 2014b). *P. verticillatum* is used in Ayurveda as an aphrodisiac and against weakness (Zhao et al., 2014b). *P. odoratum* and *P. multiflorum* are also used in traditional medicine in Europe (Fischer et al., 2008), mainly for wound healing, but also as emetics. Though the rhizome of some species is used for medicinal purposes, all parts of the four European *Polygonatum* species, especially the berries, are (slightly) toxic (Fischer et al., 2008).

European/Austrian *Polygonatum* species

Morphology

The rhizomes of species of the genus *Polygonatum* are very characteristic seal-resembling structures, thus earning them their German name of “Salomon’s Siegel” (*Figure 1a*).They are found only a few centimeters under the soil and can be up to several dozen centimeters long. The aboveground body of the plant, depending on the species, can be up to 80 (100) cm high (*P. latifolium, P. multiflorum, P. verticillatum*; Fischer et al., 2008).

The identification of *Polygonatum* species based on morphological characters alone can be difficult when the plants do not flower and if no information about the ecology is provided. Clear species identification is possible with adult, flowering plants, and/or in the habitat context. The morphological characters in this section are described using own field observations and Fischer et al. (2008).

The identification of *P. latifolium* was relatively easy due to its unique hairy and shiny abaxial leave surface (*Figure 1b*). The adaxial leaf surface of this species was also shinier than that of other species (*Figure 1b*). The second species that was easily identified was *P. verticillatum*, as it is the only species of the sect. *Verticillata* with whorled/verticillate phyllotaxis (*Figure 1c*). In combination with its ecological preferences, this species could be clearly distinguished from the other three species.

In contrast, the determination of *P. multiflorum* and *P. odoratum* was more difficult. Both species possessed matt and hairless abaxial and adaxial leaf surfaces (*Figure 1b*). However, *P. odoratum* plants mostly carried one, at most two, flowers per node (*Figure 1c*), the flowers were not contracted over the ovules and the filaments were hairless. In addition, plants of this species had an angular stem. In contrast, *P. multiflorum* carried more than one flower per node, the flowers were contracted over the ovule, the filaments hairy, and the stem was roundish (*Figure 1c*).

Figure 1: Morphological characters of the four European *Polygonatum* species. *P. latifolium (P. l), P. multiflorum (P. m), P. odoratum (P. o)* and *P. verticillatum (P. v)*; (a) rhizome with seal-like structure; (b) adaxial leaf surface (above) and abaxial leaf surface (below) with hairs in *P. latifolium (P. l)* and a matt surface in the remaining three species; (c) typical habitus of all four species, with one flower per node in *P. multiflorum (P. m)* and verticillate leaves in *P. verticillatum (P. v).*

Ecology

Plants of the genus *Polygonatum* typically grow in fertile soil in forests or bushes (Zhao et al., 2014b). European species, however, are also found in drier habitats and on poor soil. Most of the European *Polygonatum* species differ in ecological preferences and thus in habitat.

P. latifolium grows in hardwood forests, old parks, and hard floodplain forests (Fischer et. al, 2008), for example at the Donauauen *(Figure 2a)* or the Augarten. *P. multiflorum* is found in hardwood forests and floodplain forests (Fischer et. al, 2008), for example at the Harzberg *(Figure 2b)*. These two species can also easily be distinguished morphologically. *P. odoratum* is calcipole and grows mainly in dry grasslands and pinewoods (Fischer et. al, 2008), for example at the Semmering *(Figure 2c)*. *P. multiflorum* and *P. odoratum* are both found on low mountains, but the former grows on the wetter sites, while the latter grows on the drier sites. *P. verticillatum* grows in upper montane and subalpine areas, e.g., at the Gesäuse *(Figure 2d),* and in spruce forests; this species is morphologically and ecologically easy to identify.

Figure 2: Typical habitats for (a) *P. latifolium* (Donauauen), (b) *P. multiflorum* (Harzberg), (c) *P. odoratum* (Semmering) and (d) *P. verticillatum* (Gesäuse).

Karyotype

In addition to morphology and ecology, the karyotypes of *Polygonatum* species can be used to aid their determination. The chromosome numbers, karyotype morphology, and number of secondary constrictions/satellites are important diagnostic features. The secondary constriction/satellite is the region carrying active 35S rDNA loci. The rDNA within an active 35S rDNA locus is decondensed, as it is transcribed during most of the cell cycle. The secondary constrictions/satellites are also referred to as the nucleolar organizing regions (NORs) (Hirai, 2020).

Base chromosome numbers of x = 9, 10 and 14 were reported for European *Polygonatum* species (Therman, 1953), whereas for American species, the only reported base chromosome number was that of x = 10 (Therman, 1950; Zhao et al., 2014a). Hybrids were

also reported in European species (Therman, 1950; Therman, 1953), and polyploidization events were documented in several American, European, and Asian species (Therman, 1950; Zhao et al., 2014a). The polyploids were often found in separate populations (Zhao et al., 2014a). The diploid and triploid hybrids were not fertile (Therman, 1953). Asian species harbor the highest karyotypic variation with multiple diploid chromosome numbers reported from $2n$ = 14 to $2n$ = 91 (Zhao et al., 2014a). Although the origin of the polyploids often remains unknown, allopolyploidization events were inferred for some Asian species (Dale and Kiger, 2002; Zhao et al., 2014a). In the Eurasian species, four karyotype types have been observed, based on chromosome/karyotype morphology and the asymmetry indices: 2B, 3B, 2C and 3C (Deng et al., 2009). The karyotypes 2C and 3C were more asymmetric than 2B and 3B, with 3C being the most asymmetrical (Deng et al., 2009). The hypothesis of an evolutionary trend of increasing karyotype asymmetry accompanied by changes of base chromosome numbers has been proposed (Deng et al., 2009). The number and localization of rDNA loci (NORs) also varied between species, sometimes even between populations in the same species (Wang et al. 1987; Deng et al. 2009; Zhao et al., 2014a). In contrast, the American *Polygonatum* species showed very similar karyotypes between species and populations (Therman, 1950).

Three base chromosome numbers were reported for European *Polygonatum* species: *P. latifolium* and *P. odoratum* with x = 10 ($2n$ = 20) and two secondary constrictions, *P. multiflorum* with x = 9 ($2n$ = 18) and two secondary constrictions, and *P. verticillatum* with x = 14 ($2n$ = 28) and one secondary constriction (Therman, 1953). The karyotypes of *P. latifolium*, *P. multiflorum* and *P. odoratum* consisted mainly of metacentric and submetacentric chromosomes (sect. *Polygonatum*), while the karyotype of *P. verticillatum* consisted mainly of long subtelomeric and short metacentric chromosomes (sect. *Verticillata*) (Tamura et al., 1997). In *P. odoratum*, chromosomal polymorphisms were also observed (Weiss-Schneeweiss and Jang, 2003; Deng et al., 2009), indicating ongoing chromosomal rearrangements.

Genome evolution

Fluorescence *in situ* hybridization (FISH)

FISH is a 30-year-old technique that allows mapping of known RNA or DNA sequences to homologous chromosomal regions (Jiang, 2019). Thus, FISH allows to determine the number, size, and localization of loci of interest in chromosomes. Karyological polymorphisms can also be detected (Schwarzacher and Heslop-Harrison, 2000). Probes

used most often for FISH are coding regions of rRNA genes. They are very conserved and no *a priori* information of genome sequence is required. Hence, size and number of rDNA loci can be established and changes in chromosome structure can be inferred in various plant species, not only in model organisms. In *Hypochaeris*, for example, a genus within the Asteraceae, FISH was used to localize the rDNA regions in diploids and polyploids, and therefore show the karyotype evolution in this system (Weiss-Schneeweiss et al., 2008).

In addition to conserved rDNA sequences, satellite DNAs (satDNAs) are also often used as probes for FISH. These tandemly repeated sequences are fast evolving and often species/genus specific (Plohl et al., 2008). Hence, they are used for evolutionary karyotype analyses in closely related taxa. Their dynamic evolution makes their isolation difficult, but recent advances in next generation sequencing (NGS) and the development of pipelines dedicated to the analyses of repeatomes allow for quick identification of all satDNAs in any genome of interest.

Next generation sequencing (NGS)

Recent years have witnessed the development of NGS methods (van Dijk et al., 2014). These are high throughput methods that enable massive parallel sequencing of any genome of interest without *a priori* knowledge of its composition. Combined high throughput and a lower per base-cost (van Dijk et al., 2014), make it a method of choice for phylogenetic (e.g. Zimmer and Wen, 2015), populational phylogenetic and/or demographic analyses (e.g. Hou and Li, 2020) as well as for genome assemblies (Davey and Blaxter, 2011; van Dijk et al., 2014; Zimmer and Wen, 2015) or transcriptome analyses (Strickler et al., 2012). NGS data can also be used to analyze repeatomes (profiles of total repetitive DNA fraction of the genome) (e.g. Macas et al., 2015; Simon, 2020), especially in combination with RepeatExplorer, a dedicated pipeline for repeat analyses (Weiss-Schneeweiss et al., 2015; Novák et al., 2020). Such analyses allow profiles of repetitive DNA families to be established including tandem repeats, such as rDNA, as well as various types of dispersed repeats, such as transposable elements (Weiss-Schneeweiss et al., 2015; Novák et al., 2020). Additionally, comparative evolutionary analyses of the repeat profiles in groups of closely related species can be performed providing novel information pertaining to the evolution of groups of taxa (e.g. Dodsworth et al., 2016).

Repeatome

Formerly treated as junk DNA (Palazzo and Gregory, 2014), non-coding repetitive sequences are now a subject of intensive research, as they form a substantial proportion of plant genomes and are important drivers of genome evolution (Weiss-Schneeweiss et al., 2015). In addition, the repeatome is responsible for the C-value paradox (Weiss-Schneeweiss et al., 2015), which refers to a lack of correlation between genome size, number of genes and biological complexity of the taxa (Choi et al., 2020b). In the absence of polyploidy, dynamic changes in repetitive DNAs amount and content affect genome size of the taxa, often without affecting the chromosome numbers. Hence, there is no correlation between genome size and chromosome number. Two major types of repeats are distinguished: tandem repeats and dispersed repeats (Heslop-Harrison and Schmidt, 1998; Weiss-Schneeweiss and Schneeweiss, 2013; Weiss-Schneeweiss et al., 2015).

Tandem repeats

Tandem repeats are DNA monomers arranged tandemly in one to many loci. The monomers can be of different length, ranging from a few base pairs up to a few thousand base pairs (Weiss-Schneeweiss et al., 2015). The tandem repeats include ribosomal RNA genes (5S and 35S rDNA) and satDNAs (Weiss-Schneeweiss et al., 2015).

Ribosomal DNA monomers are up to a few thousand base pairs long and repeated up to a few hundred times in a locus (Weis et al., 2015). Two different rDNA types are found in the eukaryotic genomes: 35S rDNA, encoding two components of the big subunit (5.8S rDNA and 25S rDNA) and one for the small subunit (18S rDNA) of the ribosome, and the 5S rDNA region for the large subunit of the ribosome (Weis et al., 2015). The genic regions of rDNA are highly conserved because ribosomes are essential for the function of all organisms as they translate mRNA into proteins. However, between the conserved genic regions (18S rRNA, 5.8S rRNA and 25S rRNA of 35SrDNA monomer and 5S rRNA gene of 5S rDNA monomer) variable spacers are found (non-transcribed spacers [NTS] of 5S rDNA; intergenic spacers [IGS], internal transcribed spacers [ITS] and external transcribed spacers [ETS] of 35SrDNA). These regions are often used for phylogenetic analyses especially between species or genera (White et al., 1990).

SatDNAs, including satellites, mini- and microsatellites, are non-coding repeats which are often found at the telomeres or centromeres, but also in interstitial chromosomal positions (Garrido-Ramos, 2017). They form tandemly repeated clusters of monomers as short as one

base pair to more than a thousand base pairs (Garrido-Ramos, 2017). Due to their repetitive nature, the frequency of the four bases in satDNAs often differs from the rest of the genome and therefore, these DNA sequences tend to be lighter or heavier in a density gradient (Kit., 1961). Fast evolving satDNAs are not under strong selective pressure and mutate frequently and are thus often used for phylogenetic studies in closely related groups of taxa (Plohl et al., 2008). Not only monomer sequence but also the abundance of satDNAs can be very variable between the genomes of different lineages (Ellegren, 2004).

Dispersed repeats

The dispersed repeats are represented by mobile genetic elements, so-called transposable elements (TEs; Kumar and Bennetzen, 1999; Kejnovský et al., 2012; Weiss-Schneeweiss et al., 2015) (*Figure 4*). Within the TEs, retrotransposons (Class I; copy-and-paste mechanism), DNA transposons (Class II; cut-and-paste mechanism) and Helitrons (now also Class II; rolling cycle replication; Kapitonov and Jurka, 2001) are distinguished (Weiss-Schneeweiss et al., 2015). Whereas the copy number of DNA transposons increases only slowly due to their mobility mechanism, the copy numbers of active retrotransposons increase much faster. Within the retrotransposons, LTR (long terminal repeats) and non-LTR retrotransposons are found. LTR retrotransposons include Ty1*/copia* and Ty3*/gypsy* lineages, while non-LTR retrotransposons are represented by LINEs and SINEs (Weiss-Schneeweiss et al., 2015). In plant genomes, Ty1*/copia* and Ty3*/gypsy* elements are generally most abundant (Kejnovský et al. 2012). The number of non-autonomous TE copies i.e., TEs which lost the ability to transpose by themselves, is much higher than the number of autonomous elements in most genomes (Schulmann and Wicker, 2013). These non-autonomous TEs represent ancestral and often mutated elements, for example, SINEs, MITEs, and TRIMs (Schulmann and Wicker, 2013).

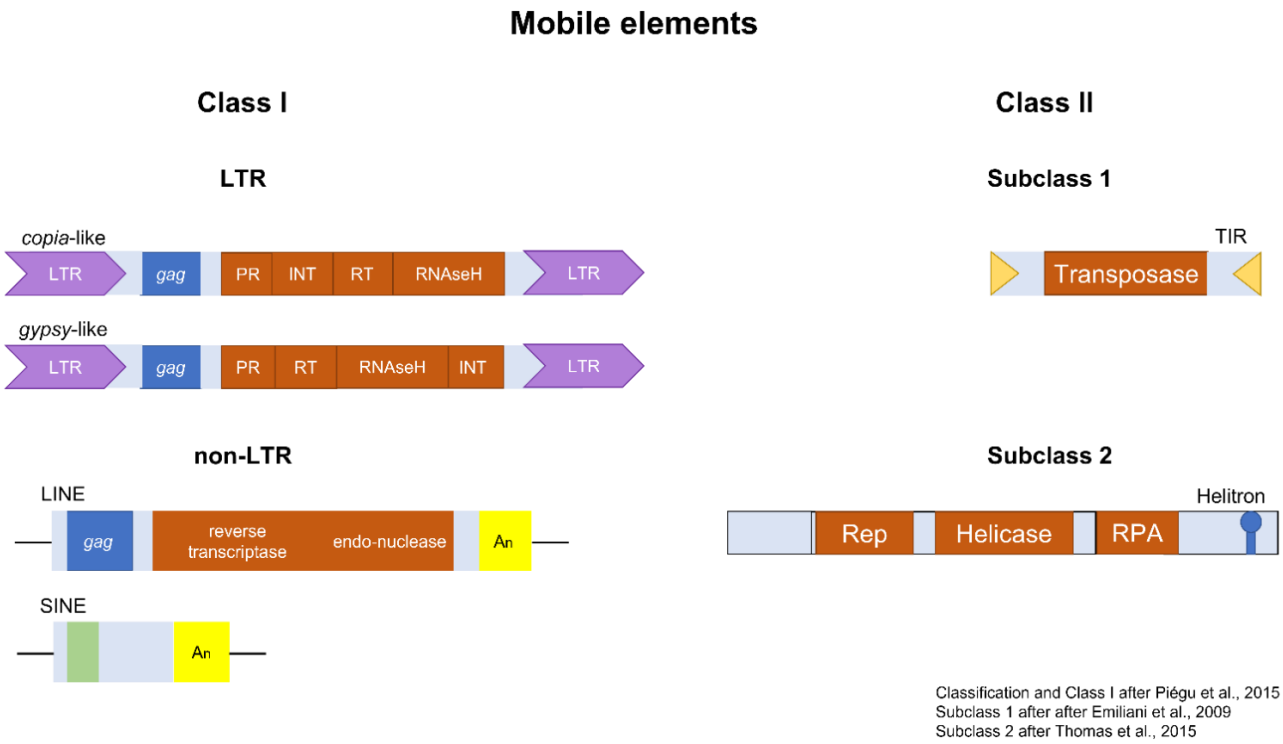


Figure 3: Overview of mobile genetic elements (TEs). Class I mobile genetic elements, including LTR retrotransposons (*copia*-like and *gypsy*-like) and non-LTR retroposons (LINE [long-interspersed element], SINE [small-interspersed element]); Class II mobile genetic elements, including DNA transposons (Subclass 1) and Helitrons (Subclass 2); LTR: long-terminal repeats, PR: protease, INT: integrase, RT: reverse transcriptase, RNAseH: RNAse H, A_n: Poly-A, Rep: replication initiator, RPA: replication protein A, blue structure: stem loop.

The aims of the study

Plant genomes are highly variable concerning their genome sizes, chromosome numbers, karyotype structure and repeat profiles. Changes in the repetitive DNA fraction are responsible for changes of the genome structure and size in the absence of polyploidy. Repetitive DNAs are therefore important drivers of genome evolution in plants (Weiss-Schneeweiss et al., 2015). Analyses of these different characters enable evolutionary analyses of related plant genomes. Little is known about the genomes of the genus *Polygonatum* species, except for their chromosome numbers and numbers of secondary constrictions. The aim of this study is therefore to shed light on the genome evolution of the genus *Polygonatum* by analyzing the four European *Polygonatum* species, *P. latifolium*, *P. multiflorum*, *P. odoratum* and *P. verticillatum*. To accomplish this goal, (1) the karyotypes were analyzed, (2) the genome sizes was measured, (3) the phylogenetic relationships were inferred (4) the repeat profiles of individual species were established, and (5) comparative analyses of the repeatomes of all four species were performed.

Chromosome numbers, karyotype structure, number and localization of rDNA loci were analyzed using classical and molecular cytogenetic techniques (Feulgen staining and FISH, respectively). For the calibration of chromosome numbers and genome sizes, the DNA content of the same individuals was measured using flow cytometry. In addition, the ITS1 and ITS2 regions were sequenced and analyzed phylogenetically to aid species identification and recover species relationships. Lastly, the genomes of the same plants were sequenced using NGS Illumina technology and analyzed using RepeatExplorer to establish a repeat profile for each species and then to compare the profiles among the species. All these approaches allowed for the comparative and comprehensive evolutionary analyses of the four European *Polygonatum* species and the formulation of a new hypotheses on their relationships.

Material and Methods

Plant material and pretreatment for cytological analyses

58 individuals representing 17 populations collected from 11 different locations were cultivated in the Botanical Garden of the University of Vienna (*Table 1*). The old roots were removed, and the rhizomes were placed in soil. New actively growing roots were harvested after two to three weeks and pretreated with colchicine following the protocol of Choi et al. (2020a). Briefly, the roots were washed with water to remove any soil particles and incubated in 0.05% colchicine solution for 4.5 hours in the dark. The material was then washed with water and fixed in 3 : 1 ethanol : acidic acid mixture. Fixed roots were stored at -20°C until further use.

Table 1: Individuals used for the study on European *Polygonatum* species. The ID consists of the population number and individual number; Geier (Geier Christian), Hanusch (Hanusch Daniela), Primer (Primer Clara).

Species	ID (Pop.Ind)	Locality/date/collector	Ecology
P. latifolium	4.1	Roter Berg/16.04.21/Hanusch	forest edge
P. latifolium	4.2	Roter Berg/16.04.21/Hanusch	forest edge
P. latifolium	4.3	Roter Berg/16.04.21/Hanusch	forest edge
P. latifolium	4.4	Roter Berg/16.04.21/Hanusch	forest edge
P. latifolium	6.1	Park bei Lainzerstrasse/Hanusch	roadside
P. latifolium	6.2	Park bei Lainzerstrasse/Hanusch	roadside
P. latifolium	7.1	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.2	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.3	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.4	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.5	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.6	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.7	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.8	Prater/24.04.21/Hanusch	roadside
P. latifolium	7.9	Prater/24.04.21/Hanusch	roadside
P. latifolium	8.1	Augarten/Geier	forest edge
P. latifolium	8.2	Augarten/Geier	forest edge
P. latifolium	9.2	Augarten/Geier	forest edge
P. latifolium	9.3	Augarten/Geier	forest edge
P. latifolium	10.1	Augarten/Geier	forest edge
P. latifolium	10.2	Augarten/Geier	forest edge
P. latifolium	10.3	Augarten/Geier	forest edge
P. latifolium	12.4	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest
P. latifolium	12.5	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest
P. latifolium	12.6	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest
P. multiflorum	1.1	Kallusweg/10.04.21/Hanusch	forest edge
P. multiflorum	1.2	Kallusweg/10.04.21/Hanusch	forest edge
P. multiflorum	1.3	Kallusweg/10.04.21/Hanusch	forest edge
P. multiflorum	5.2	Maurerwald/16.04.21/Hanusch	forest
P. multiflorum	6.5	Park bei Lainzerstrasse/Hanusch	young forest
P. multiflorum	12.1	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest
P. multiflorum	12.2	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest
P. multiflorum	12.3	Hundsheimer Berg (N)/01.05.21/Hanusch	young forest

P. multiflorum	14.1	Harzberg/13.05.21/Hanusch	young forest
P. multiflorum	14.2	Harzberg/13.05.21/Hanusch	young forest
P. multiflorum	14.3	Harzberg/13.05.21/Hanusch	young forest
P. multiflorum	14.4	Harzberg/13.05.21/Hanusch	young forest
P. multiflorum	15.1	Huehnerberg/13.05.21/Hanusch	young forest
P. multiflorum	15.2	Huehnerberg/13.05.21/Hanusch	young forest
P. multiflorum	15.3	Huehnerberg/13.05.21/Hanusch	young forest
P. multiflorum	15.4	Huehnerberg/13.05.21/Hanusch	young forest
P. odoratum	11.1	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	11.2	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	11.3	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	11.4	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	11.6	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	11.7	Hundsheimer Berg/01.05.21/Hanusch	shrubland
P. odoratum	20.1	Semmering/11.08.21/Hanusch	montane
P. odoratum	20.2	Semmering/11.08.21/Hanusch	montane
P. odoratum	20.3	Semmering/11.08.21/Hanusch	montane
P. odoratum	20.4	Semmering/11.08.21/Hanusch	montane
P. verticillatum	13.1	Ausschlag-Zoebern/13.05.21/Hanusch	spruce forest
P. verticillatum	13.2	Ausschlag-Zoebern/13.05.21/Hanusch	spruce forest
P. verticillatum	13.3	Ausschlag-Zoebern/13.05.21/Hanusch	spruce forest
P. verticillatum	13.4	Ausschlag-Zoebern/13.05.21/Hanusch	spruce forest
P. verticillatum	16.1	Naturfreunde Haus Huette Knofele-ben/05.06.21/Primer	young forest
P. verticillatum	16.2	Naturfreunde Haus Huette Knofele-ben/05.06.21/Primer	young forest
P. verticillatum	16.4	Naturfreunde Haus Huette Knofele-ben/05.06.21/Primer	young forest

Feulgen staining

Feulgen staining was performed to establish chromosome number and karyotype structure of the analyzed *Polygonatum* samples following a modified protocol of Weiss-Schneeweiss et al. (2009). Briefly, the pretreated and fixed root meristems were incubated in a glass tube with 5N HCl for 20 minutes and washed with tap water. Schiff's reagent (Merck, Vienna, Austria) was added to the material to stain the DNA for up to one hour in the darkness. Afterwards, the Schiff's reagent was removed, the meristems were placed on a glass slide in a drop of 60% acetic acid, fragmented with entomological needles and covered by a cover slip. The preparations were then gently tapped with a needle to spread the material and squashed with the thumb. The chromosome number was established for at least two individuals per population. Preparations with at least three well spread metaphase plates were analyzed. Complete and well-dispersed mitotic metaphases were photographed using a CCD camera mounted on a light microscope Imager.A1 and the AxioVision software (Carl Zeiss, Vienna, Austria). Chromosomes were cut out and arranged into karyotypes using Corel Photo-Paint 12.0 (Corel Corporation, Ottawa, Canada).

Flow cytometry

Genome size was measured using flow cytometry. *Pisum sativum* (Fabaceae; "Kleine Rheinländerin"; 1C = 4.42 pg; Greilhuber and Ebert, 1994) served as internal standard for *P. latifolium, P. multiflorum* and *P. odoratum*, while *Solanum pseudocapsicum* (Solanaceae; 1C = 1.29 pg; Temsch et al., 2010) was used as standard for *P. verticillatum*. Approximately 0.5 cm^2 of fresh leaf material of the standard and 0.5 cm^2 - 1 cm^2 of fresh leaf material or – if no leaf material was available – about 1 cm peeled and washed rhizome of the probe were co-chopped in 550 µL of citrate buffer (Otto et al., 1981) using a razor blade (Galbraith et al., 1983). Another 550 µL of citrate buffer were added to completely suspend the chopped material, followed by a filtering step using a 30 µm nylon mesh (Saatilene Hitex, Sericol, Germany) to separate the plant tissue debris and the isolated nuclei. Subsequently 50 µL of RNaseA (Sigma Aldrich, Vienna, Austria) were added to the filtrate to remove the RNAs, followed by incubation in a water bath for 30 min at 37°C. The suspensions were then aliquoted into two tubes, and one aliquot was stained with 2 mL propidium iodide (1 mg/mL in Otto's buffer II; Temsch et al. 2010; Sigma-Aldrich, Vienna, Austria) for at least 30 minutes in the fridge at 4°C. Samples were analyzed using CyFlow Space flow cytometer (Sysmex Austria GmbH, Vienna, Austria) equipped with a green laser (5432 nm, 100 mW, Cobolt Samba, Cobolt AB, Stockholm, Sweden). Each run included one sample of the standard alone to calibrate the machine. Three runs (measurements) were performed for each individual, resulting in total of 9999 nuclei (particles) per sample/plant. These particles were then visualized graphically as a Gaussian bell-curve using FloMax Software (Partec, Münster, Germany).

The calculation of the genome size (1C-value) of each sample was performed in Microsoft Excel using the following formula:

$$Genome\ size\ sample\ (pg) = \frac{(mean\ G1\ peak\ sample)}{(mean\ G1\ peak\ standard)} \times genome\ size\ standard\ (pg)$$

The average genome size per species as well as the standard deviation were calculated in Microsoft Excel.

FISH

FISH was performed using a modified protocol of Jang and Weiss-Schneeweiss (2015). Briefly, fixed root meristems were digested with an enzyme mix containing 0.4% pectolyase (Sigma-Aldrich, Vienna, Austria), 0.4% cytohelicase (Sigma-Aldrich, Vienna, Austria) and 1% cellulase (Serva, Vienna, Austria) in a citrate buffer (pH = 4.8) at 37°C for about 1 to 1.5 hours, depending on the quality of the material. The enzyme mix was removed, the root tips washed for 10 - 20 minutes in citric buffer and the meristems were transferred onto a slide in a drop of 60% acetic acid. The meristem was fragmented, a cover slip applied, and the preparation was squashed. In contrast to Feulgen staining, the material was very soft after enzymatic digestion and therefore the squashing was performed more carefully. The preparations were analyzed for quality using phase contrast of an Imager.A1 microscope (Carl Zeiss, Vienna, Austria) and only preparations with at least five well-spread complete mitotic metaphase plates were used for FISH. Slides were frozen at -80°C to facilitate the removal of coverslips, air-dried at room temperature and stored at 4°C until use, maximally for a few days.

The preparations were re-fixed in 3 : 1 ethanol : acidic acid mixture for 20 minutes, followed by dehydration in 96% ethanol for 20 minutes and air-dried for at least 40 minutes at room temperature. RNase treatment was omitted. The slides were washed in 2xSSC and incubated in the pepsin solution (Sigma-Aldrich, Vienna, Austria) (pH = 2, 10 µg/mL) for 20 minutes at 37°C to remove proteins. After the pretreatment, the slides were washed three times with 2xSSC and incubated in 4% paraformaldehyde to re-fix the histones. The slides were washed in three changes of 2xSSC and dehydrated for 3 minutes in cold 70% ethanol (4°C) and for 3 minutes in cold 96% ethanol (4°C). The preparations were then air-dried for 45 minutes.

The hybridization mix contained 0.02xSSC, 10% dextran sulphate (Promega, Madison, USA), 100 ng/µL blocking DNA (single stranded; Sigma-Aldrich, Vienna, Austria), sterile water and 10-100 ng of each labelled probe. The rDNA probes were cloned (18S/25S rDNA isolated from *Arabidopsis thaliana* in pSK+ plasmids; Jang and Weiss-Schneeweiss, 2015) or commercially synthesized and labelled (5S rDNA; IBA, Germany). 18S/25S rDNA probes were labelled with digoxygenin by nick-translation using dig nick translation kit (Roche, Vienna, Austria). The hybridization mix was denatured on a thermo-block at 96°C for 3 minutes followed by incubation for 5 minutes on ice. After air-drying, 8 µl of the hybridization mix were added to each slide, covered with a coverslip, and sealed with

Fixogum. Slides with the hybridization mix were denatured on a PCR *in situ* block (PeqlabGMBH, Erlangen, Germany) using the following cycling: 4 minutes at 72°C, 1 minute at 65°C, 1 minute at 55°C, and 1 minute at 45°C. The slides were then transferred to the incubator at 37°C for overnight hybridization.

After overnight hybridization, the Fixogum was removed from the slides and the slides were placed in a Coplin jar in 2xSSC at room temperature until the cover slips fell off. The following stringent washes were performed in a water bath: 2xSSC at 39°C, three changes of 2xSSC at 39°C and 2xSSC+0.2%Tween20 at 39°C. Labelled probes were detected using Streptavidin-Cy3 (for biotin detection; Sigma Aldrich, Vienna, Austria) and anti-digoxygenin-FITC (for detection of digoxygenin; Sigma Aldrich, Vienna, Austria). Briefly, slides were incubated with blocking solution (2% BSA [Sigma-Aldrich, Vienna, Austria] in 2xSSC+0.2%Tween20) for 20 minutes at 37°C and then incubated in a detection solution (blocking solution containing 2µg/mL streptavidin-Cy3 and 5µg/mL anti-digoxygenin-FITC) at 37°C for 60 minutes. Subsequently slides were washed twice with 2xSSC at 42°C and once with 2xSSC+0.2%Tween20 at 42°C. Finally, the antifade buffer Vectashield (Vector Laboratories, Burlingame, USA) containing DAPI (2µg/mL; Sigma-Aldrich, Vienna, Austria) was added to each slide and new coverslips were applied. The preparations were stored in the dark at 4°C for at least overnight to stabilize the fluorescence. Slides were analyzed using a Zeiss Imager.M2 microscope (Carl Zeiss, Vienna, Austria). Images were taken by a CCD camera and visualized by the AxioVision 4.8 software (Zeiss, Vienna, Austria). The pictures were contrasted using Adobe (Adobe Inc., San José, USA) and/or Corel Photo Paint (Corel Corporation, Ottawa, Canada) using only the functions that applied equally to the whole image. The chromosomes were cut out using Corel Photo Paint (Corel Corporation, Ottawa, Canada).

DNA extraction and Sanger sequencing

Genomic DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). Briefly, 5 - 25 mg of silica dried leaves were ground with glass beads in a TissueLyser II (QIAGEN, Hilden, Germany) after pre-cooling the tubes in liquid nitrogen. This powder was then mixed with buffer AP1 and RNaseA to digest RNA in the samples, modifying the protocol by increasing the incubation time with buffer AP1 and RNase to 22 minutes and the following cooling to 20 minutes. After incubation, the solution was centrifuged, the supernatant was collected and buffer P3 was added. The lysate was transferred in a QIAshredder spin column to remove cell walls and other debris. After centrifuging, the lysate

was transferred into a new tube and buffer AW1 was added. The lysate was pipetted in a DNeasy Mini spin column and centrifuged. Afterwards, the membrane with bound DNA was washed with ethanol and dried. Buffer AE was added after drying to elute the DNA. DNA extracts were checked on 1.5% agarose gel at 86V for 30 minutes.

PCR was performed using a thermal cycler GeneAmp PCR machine (PCR System 9700; ThermoFisher Scientific, Vienna, Austria) to amplify the ITS1 and ITS2 regions using primers ITS4 (5'- TCCTCCGCTTATTGATATGC-3') and ITS5 (5'- GGAAGTAAAAGTCGTAACAAGG-3') (White et al., 1990). The master mix for PCR included the following chemicals: 1x DreamTag Green PCR master mix (ThermoFisher Scientific, Vienna, Austria), forward and reverse primers (final concentration 0.5 µM), BSA (final concentration 0.2 µg/µL; Sigma-Aldrich, Vienna, Austria) and trehalose (final concentration 0.16 M; Sigma-Aldrich, Vienna, Austria). The PCR reaction was performed using the following cycling profile: 1x 3 minutes at 95°C; 40x 30 seconds at 95°C, 30 seconds at 50°C, 1 minute at 72°C; 1x 5 minutes at 72°C; hold at 10°C.

PCR products were checked on a 1.5% agarose gel and afterwards purified by incubation with Exonuclease I (1 U/µL; ThermoFisher Scientific, Vienna, Austria) and FastA (1 U/µL; ThermoFisher Scientific, Vienna, Austria) for 30 minutes, at 37°C. The temperature was increased to 85°C for 15 minutes to inactivate the enzymes.

After purification, cycle sequencing was performed in a GeneAmp PCR thermal cycler (PCR System 9700; ThermoFisher Scientific, Vienna, Austria) using standard protocol and BigDye Terminator v3.1 chemistry (ThermoFisher Scientific, Vienna, Austria) (Tremetsberger et al., 2005) Sequencing was performed using the ABI 3730 DNA Analyzer (“Long Read Sequencing” module; ThermoFisher Scientific, Vienna, Austria).

The sequences were assembled using SeqMan II Version 5.05 (DNASTAR, Madison, USA) with manual editing of ambiguities. Ambiguities with unclear peaks were retained to avoid over-interpretation. Sequences were then aligned in BioEdit (Hall, 1999), GenBank samples were added, and the sequences were then trimmed to 573 bp. The phylogenetic trees with and without GenBank samples were built using the software Mega11 (Kimura 2-parameter model, 1000 bootstraps [BS]) (Tamura et al., 2021).

NGS

The quality of selected genomic DNA extracts was checked on a 1.5% agarose gel. The concentration of extracts was measured using a Qubit assay (ThermoFisher Scientific, Vienna, Austria) following the manufacturer's protocol. Genomic DNA was fragmented into 500 - 800 bp long fragments and the library was constructed at CSF-NGS (Vienna Biocenter, Vienna, Austria). The samples (one sample of each of four *Polygonatum* species) were sequenced using 150 bp paired-end NextSeq550 Medium technology (Illumina, San Diego, USA).

The NGS data were processed and analyzed utilizing the RepeatExplorer2 pipeline (Novák et al., 2020, http://repeatexplorer.org). The reads were checked for quality and the number of overlapping and non-overlapping pairs (*Appendix 3*). The read pairs were then interlaced and clustered using default settings in RepeatExplorer2. Clustering was based on all-to-all comparison of all reads resulting in clusters and superclusters representing all major lineages of repeats present in the analyzed genomes (Novák et al., 2020). Automatic annotation of clusters was then verified manually using, among others, BLAST. Clustering allowed for the identification of satDNAs via TAREAN. These were subsequently manually verified via Dotter (Sonnhammer and Durbin, 1995).

All four samples/species were analyzed individually, which allowed to use more reads for clustering, resulting in more in-depth repeatome characterization. Subsequently, the four datasets were concatenated, and comparative analyses of all four species were performed.

Results

Chromosome numbers and karyotypes

Chromosome numbers were established, and karyotypes arranged using preparations stained with Schiff's reagent (Feulgen staining; *Figure 4 a-d*). Since NORs can but not always are visible in Feulgen-stained preparations, their number could not always be established unequivocally.

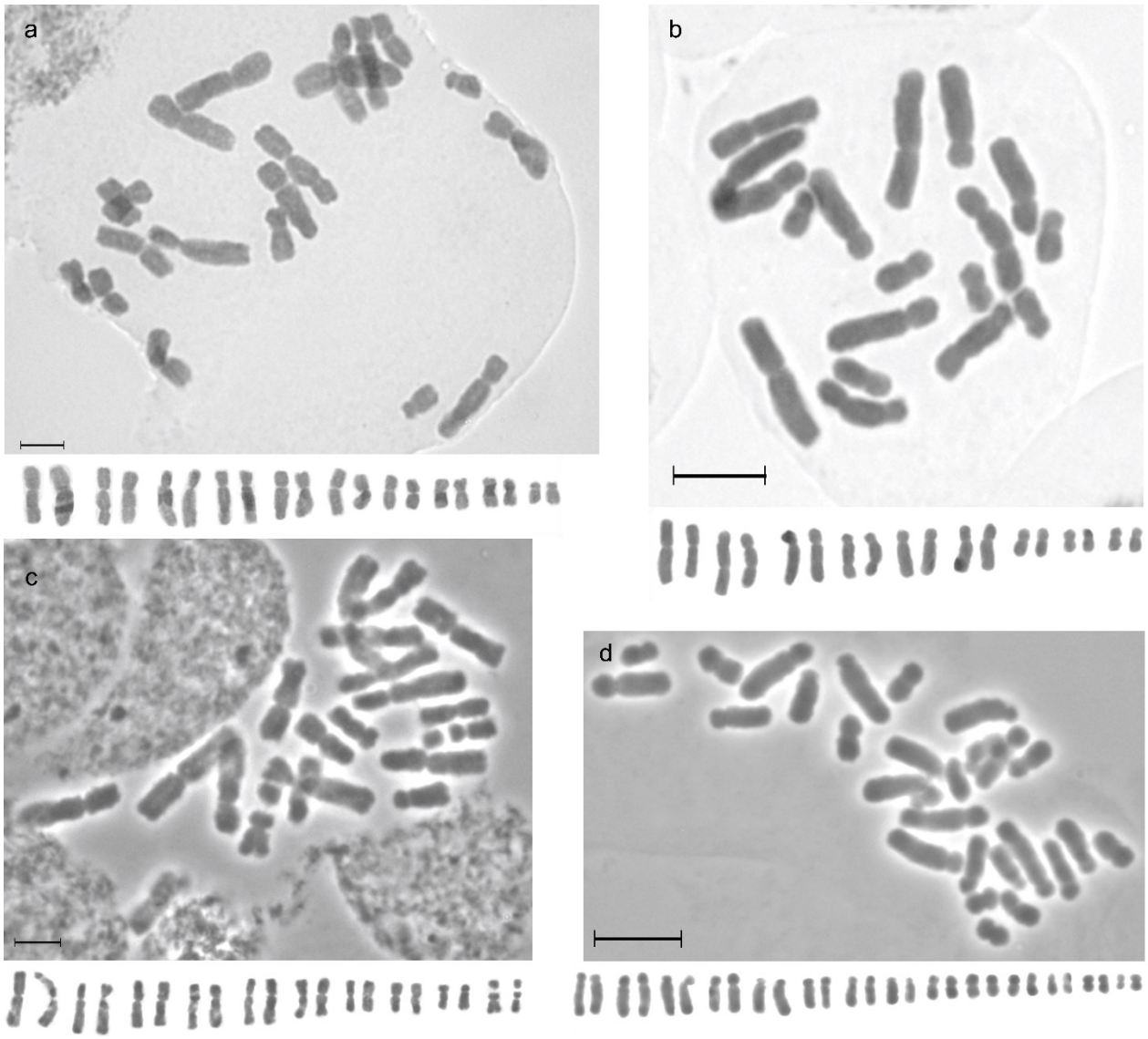


Figure 4: Mitotic metaphase chromosomes and karyotypes of (a) *P. latifolium* (2*n* = 20; ID 9.3), (b) *P. multiflorum* (2*n* = 18; ID 14.2), (c) *P. odoratum* (2*n* = 20; ID 20.1) and (d) *P. verticillatum* (2*n* = 28; ID 13.3); scale bar = 5 µm.

Table 2: Chromosome numbers and genome size measurements of the four European *Polygonatum* species, *P. latifolium, P. multiflorum, P. odoratum* and *P. verticillatum.*

Species	ID	2*n*	Genome size (pg)
latifolium	4.1	20	10.1152
latifolium	4.2	20	10.0207
latifolium	4.3	20	9.7398
latifolium	4.4	20	10.6060
latifolium	6.1	20	10.1087
latifolium	6.2	20	10.0415
latifolium	7.1	20	10.4348
latifolium	7.2	20	10.0067
latifolium	7.3	20	-
latifolium	7.4	20	-
latifolium	7.5	20	10.0290
latifolium	7.6	20	10.0191
latifolium	7.7	20	-
latifolium	7.8	20	10.1152
latifolium	7.9	20	9.8760
latifolium	8.1	20	10.0943
latifolium	8.2	20	10.0359
latifolium	9.2	20	10.0768
latifolium	9.3	20	10.0190
latifolium	10.1	20	9.4397
latifolium	10.2	20	10.1120
latifolium	10.3	20	9.8493
latifolium	12.4	20	9.6090
latifolium	12.5	20	9.9247
latifolium	12.6	-	10.4418
multiflorum	1.1	18	10.3791
multiflorum	1.2	18	10.4144
multiflorum	1.3	18	10.3433
multiflorum	5.2	18	11.0402
multiflorum	6.5	18	10.3237
multiflorum	12.1	18	10.4381
multiflorum	12.2	18	10.3306
multiflorum	12.3	-	10.3231
multiflorum	14.1	18	10.3498
multiflorum	14.2	18	10.7915
multiflorum	14.3	-	10.4689
multiflorum	14.4	-	10.3880
multiflorum	15.1	-	10.3193
multiflorum	15.2	-	10.4275
multiflorum	15.3	18	10.4925
multiflorum	15.4	18	10.2521
odoratum	11.1	20	11.5255
odoratum	11.2	-	11.7822
odoratum	11.3	20	11.7918
odoratum	11.4	20	12.1936
odoratum	11.6	-	11.1867
odoratum	11.7	-	11.1949
odoratum	20.1	20	10.0943
odoratum	20.2	-	11.2835
odoratum	20.3	-	-
odoratum	20.4	-	11.4645
verticillatum	13.1	-	8.9756

verticillatum	13.2	-	8.9445
verticillatum	13.3	-	8.9084
verticillatum	13.4	-	8.8761
verticillatum	16.1	28	8.7401
verticillatum	16.2	28	9.1807
verticillatum	16.4	28	8.8855

P. latifolium

24 individuals of seven populations of this species were analyzed. They all showed consistent chromosome number of 2*n* = 20 with two pairs of chromosomes carrying a secondary constriction on their long arms (*Figure 4a*). The chromosomes were submetacentric or metacentric (longest pair and two of the shortest pairs). No apparent polymorphisms were detected in homologous chromosome pairs. No B chromosomes were found.

P. multiflorum

All 11 individuals, representing six different populations of this species, showed a consistent chromosome number of 2*n* = 18, and no B chromosomes were found (*Figure 4b*). The chromosomes were submetacentric or metacentric (longest pair and two of the shortest pairs). Two pairs of chromosomes carried a secondary constriction on their long arms (*Figure 4b*). In population 14, a potential polymorphism of the second chromosomal pair was found (*Figure 4b*). Neither other individuals of this population nor any individuals of other populations carried polymorphisms.

P. odoratum

All four individuals of the two analyzed populations of this species showed a consistent chromosome number of 2*n* = 20 and two pairs of chromosomes with secondary constrictions, one within the long arm of a submetacentric pair and the other within the arm of the smallest metacentric pair (*Figure 4c*). The chromosomes were submetacentric and metacentric (the longest pair and three of the medium-sized pairs). No polymorphisms were detected.

P. verticillatum

The chromosomes of *P. verticillatum* were small and therefore more difficult to analyze. A consistent chromosome number of 2*n* = 28 was found in three individuals analyzed representing one population (*Figure 4d*). Due to smaller size of chromosomes and thus their higher condensation degree, chromosomal polymorphisms might have been overlooked. *P. verticillatum* had only one pair of chromosomes carrying secondary constriction on their

long arms. The chromosomes of this species were submetacentric or metacentric (the five shortest pairs).

Genome size variation

Flow cytometry was used to measure the genome size of the four European *Polygonatum* species. Altogether, 53 samples were measured, representing 17 populations: seven of *P. latifolium*, six of *P. multiflorum*, two of *P. odoratum* and two of *P. verticillatum*. The FloMax Software analyses resulted in a Gaussian peak for each sample and standard (*Figure 5*). *Pisum sativum* was used as a standard for *P. latifolium*, *P. ,multiflorum* and *P. odoratum*. As the G1 peak (genome size before replication) of *P. verticillatum* lied within the G2 peak (after replication) of the standard *Pisum sativum*, *Solanum pseudocapsicum* had to been used as standard for this taxon (*Figure 5d*).

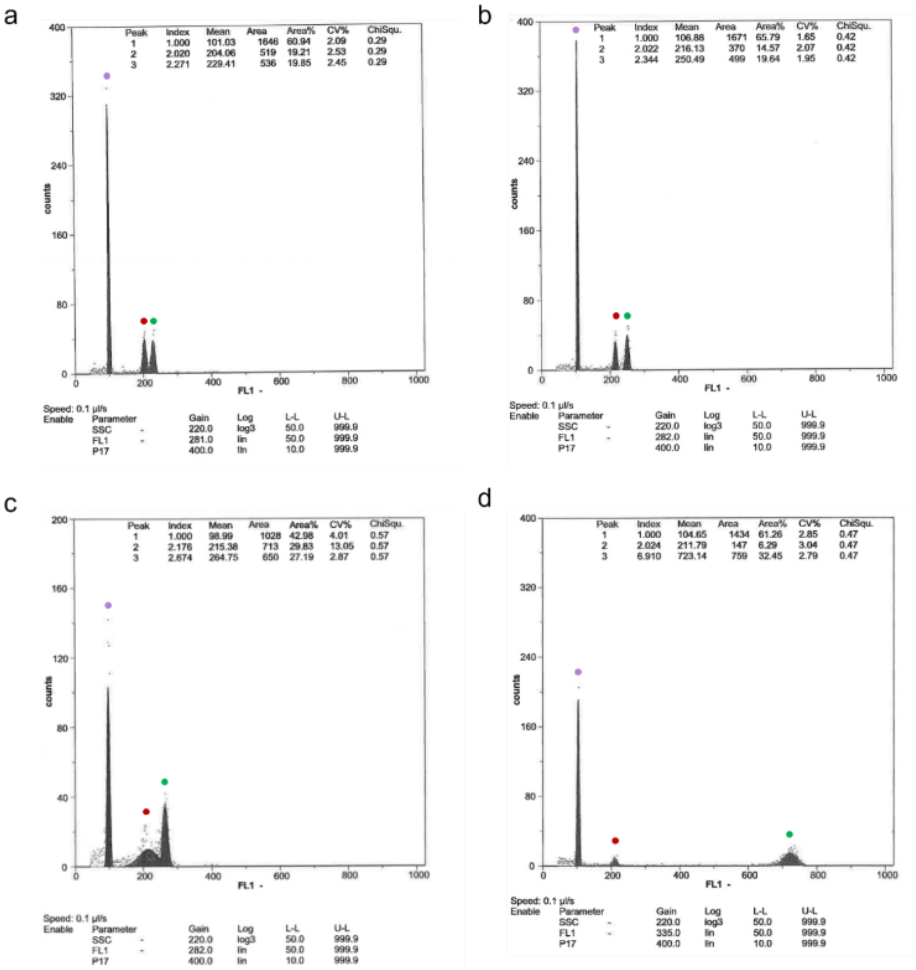


Figure 5: Histograms of the genome size measurements of the four *Polygonatum* species analyzed; (a) *P. latifolium*, (b) *P. multiflorum*, (c) *P. odoratum*, and (d) *P. verticillatum* showing the G1 peak of the standard (purple circle), the G2 of the standard (red circle) and the G1 peak of the sample (green circle).

Genome size of *P. verticillatum* was the smallest (1C = 8.93 pg ±, 0.123 pg) (*Table 2*). Genome sizes of *P. latifolium* (1C = 10.03 pg ± 0.250 pg), and *P. multiflorum* (1C = 10.41 pg, ± 0.121 pg) were more similar to one another (*Table 2*). The largest genome size of 1C = 11.39 pg ± 0.553 pg was observed in *P. odoratum* (*Table 2*). The genome size was distinct for each species. No polyploid individuals were found after calibrating the genome size using chromosome number of the same (selected) individuals.

Phylogenetic relationships of European *Polygonatum* species

The size of the sequenced ITS1 and ITS2 fragments was 634 bp and the length of the alignment after trimming was 573 bp. The sequence of *P. multiflorum* was used as reference, as the individuals showed no variation, and 33 variable characters were observed in total (*Figure 6*). The maximum likelihood (ML) analyses of the aligned ITS1 and ITS2 sequences representing 43 individuals (20 of *P. latifolium*, 15 of *P. multiflorum*, five of *P. odoratum*, three of *P. verticillatum*) revealed clear differences between species and proved to be sufficient for species delimitation (*Figure 7, Appendix 1* [with GenBank sequences]). All individuals of each of the four species were recovered in clearly defined species-specific clades. Hence, the trees were used to aid/confirm species identification. The clade of *P. verticillatum* was very well supported (BS 99), as was the clade of *P. odoratum* (BS 99). The split of *P. latifolium* and *P. multiflorum* was rather poorly supported (BS 41), but the clade of *P. multiflorum* itself had good support (BS 82).

Figure 6: Variable characters of ITS1 and ITS2 regions in the four *Polygonatum* species. *P. multiflorum* (multi), *P. latifolium* (lati), *P. odoratum* (odor) and *P. verticillatum* (vert); numbers indicate plant IDs (see *Table 1*).

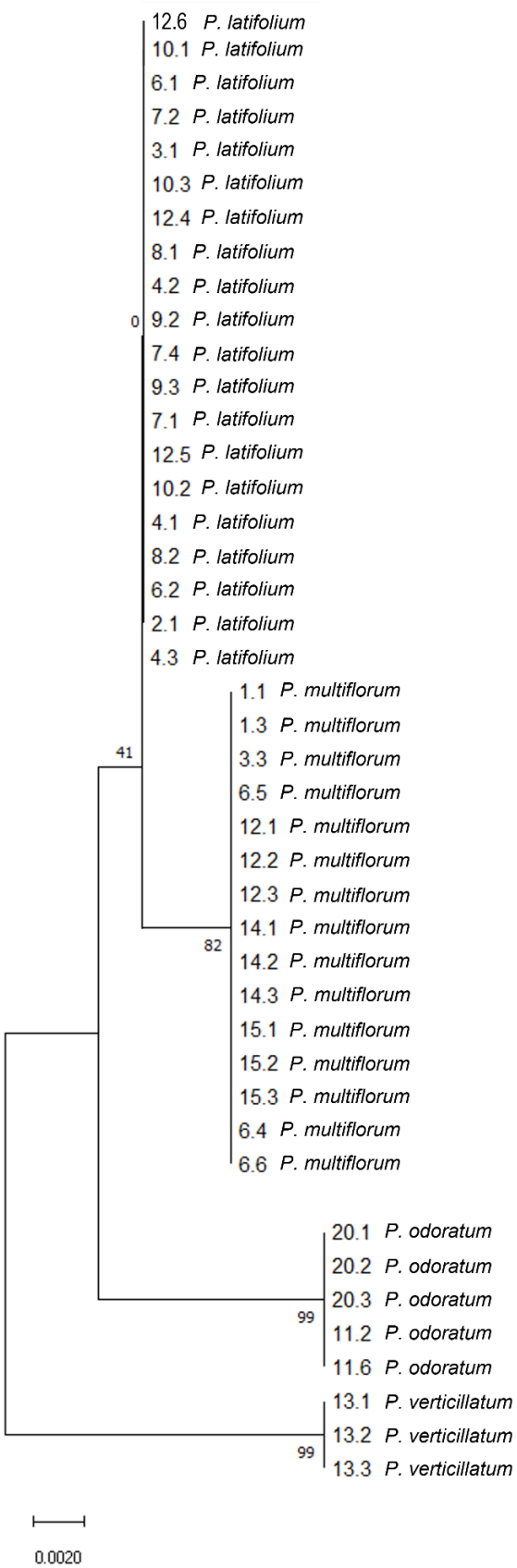


Figure 7: Phylogenetic tree based on Maximum Likelihood (ML) analyses of ITS1 and ITS2 regions of 35S rDNA of the four *Polygonatum* species . Numbers indicate the individual IDs (see *Table 1*).

Mapping of 5S and 35S rDNA loci

Well-spread mitotic metaphase plates were used to map coding regions of 35S and 5S rDNAs in chromosomes of all four analyzed species. Clearly distinguishable chromosomes carrying rDNA loci were cut out (*Figure 8 a-d, Appendix 2*).

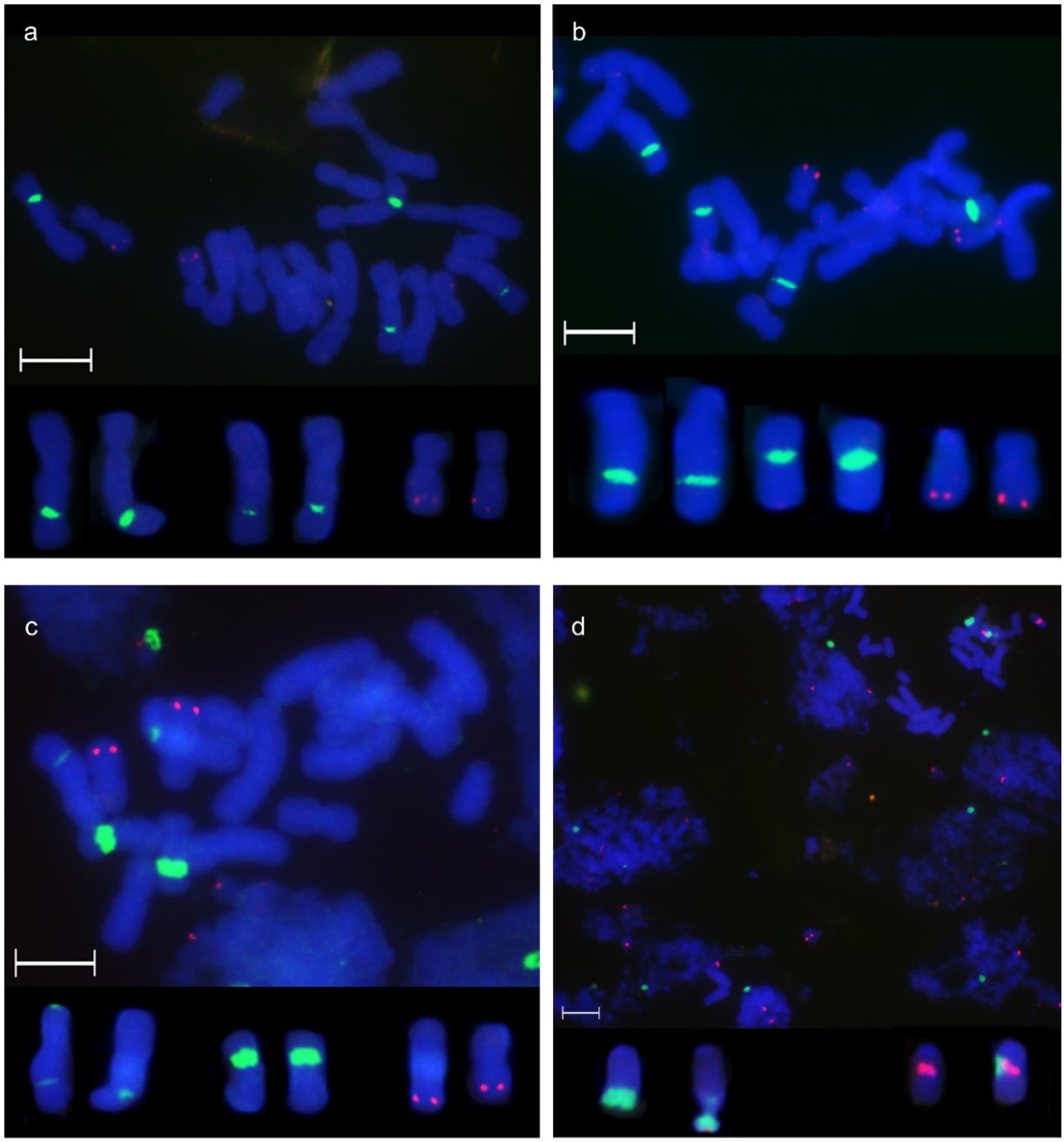


Figure 8: Mapping of the 5S and 35S rDNA loci in four *Polygonatum* species. (a) *P. latifolium* (ID 8.2), (b) *P. multiflorum* (ID 12.1), (c) *P. odoratum* (ID 11.7) and (d) *P. verticillatum* (ID 13.1), 35S rDNA loci are in green and the 5S rDNA loci in red. Clearly visible is the difference in signal strength between the 35S rDNA-carrying chromosome pairs in (c) *P. odoratum*; scale bar = 5 µm.

P. latifolium

Two 35S rDNA loci, one in the long arm of a submetacentric chromosome pair and one in the long arm of a metacentric pair were detected, as well as one 5S rDNA locus on the long arm of a submetacentric pair (*Figure 8a*). No clear polymorphism were detected. The signal size and strength were similar within all rDNA-carrying chromosome pairs.

P. multiflorum

Two 35S rDNA loci in the long arms of two submetacentric chromosome pairs were detected (*Figure 8b, Appendix 2a*). The 5S rDNA locus was localized within the long arm of a small submetacentric chromosome pair. The signals of the rDNA loci within individual chromosome pairs were of similar strength.

P. odoratum

Two pairs of chromosomes with 35S rDNA loci, one in the long arm of a submetacentric pair and one in an arm of the shortest metacentric chromosome pair were found consistently in all populations. One short submetacentric chromosome pair carried the 5S rDNA locus within the long arm (*Figure 8c, Appendix 2b*).

In population 11, collected at the Hundsheimer Berge, an individual with a difference in the signal strength between the 35S rDNA-carrying chromosome pairs was found: one pair of chromosomes carried a larger green signal, whereas the second pair of chromosomes only carried a weak green signal (*Figure 8c*). This difference in size of the 35S rDNA locus seen in population 11 was not observed in the other examined individual of population 20 (*Appendix 2b*); however, not all individuals of the two populations were used for FISH, so it is unclear whether this difference is seen in all plants of population 11.

P. verticillatum

Only one chromosome pair carried a 35S rDNA locus within its long arm (*Figure 8d*). Similarly, one pair of chromosomes carried the 5S rDNA locus within the short arm (*Figure 8d).* No polymorphism were found.

Repeat profiles

NGS sequence data of all four *Polygonatum* species were analyzed using the RepeatExplorer2 pipeline. For each species, forward and reverse reads were interlaced, and such datasets were used for the individual species analyses. The number of input reads and reads assigned to clusters are provided in *Table 3*.

Table 3: Number of input reads and reads assigned to clusters of the single species analyses of the four *Polygonatum* species.

Species	Sample ID	Number of input reads	Reads in clusters
P. latifolium	7.5	21242400	5128317
P. multiflorum	15.4	21917510	5077254
P. odoratum	20.2	20292938	4884113
P. verticillatum	13.1	18458286	5104966

Repeat composition in individual species

The analyses of individual/species-specific NGS datasets allowed for characterization of repeatomes of all four analyzed *Polygonatum* species (*Figure 9*). Subsequently, comparative analyses of all four species were performed, providing additional information about the proportions of all repeat types in these genomes (*Figure 10*).

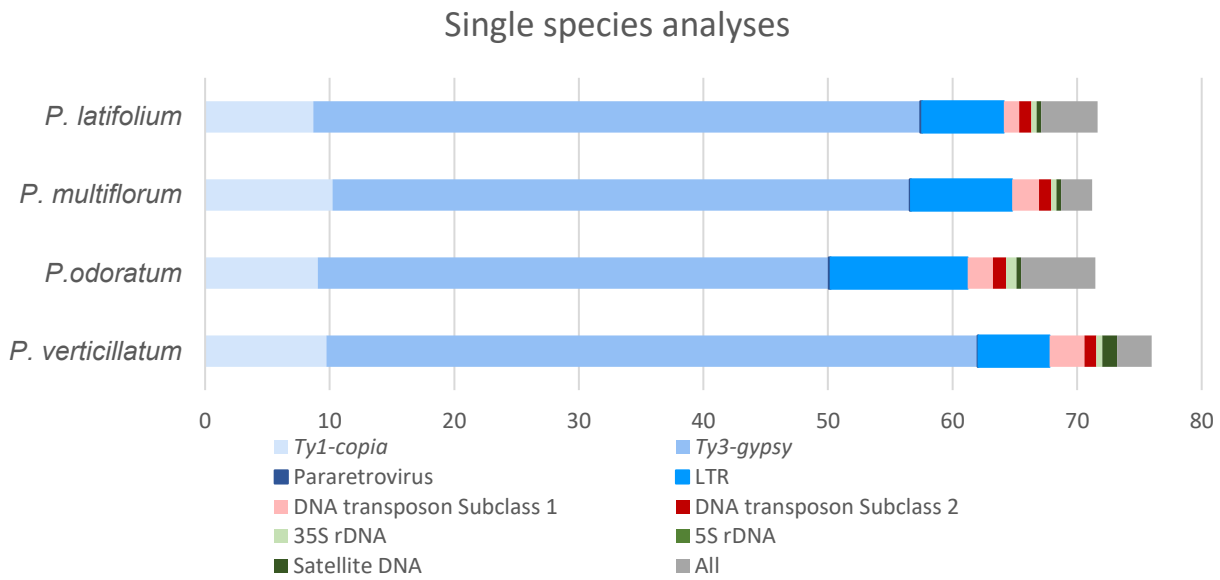


Figure 9: Repeatome composition of the four European *Polygonatum* species inferred from the single species analyses. Class I mobile elements are indicated in blue: LTR retrotransposons in shades of light blue, non-LTR elements in dark blue, non-annotated LTR elements in cyan; Class II mobile elements are indicated in red: Subclass 1 in light pink, Subclass 2 in red; Tandem repeats are indicated in green and black: rDNAs in green: 35S rDNA in light green, 5S rDNA in dark green, satDNAs in black; non-annotated reads in grey.

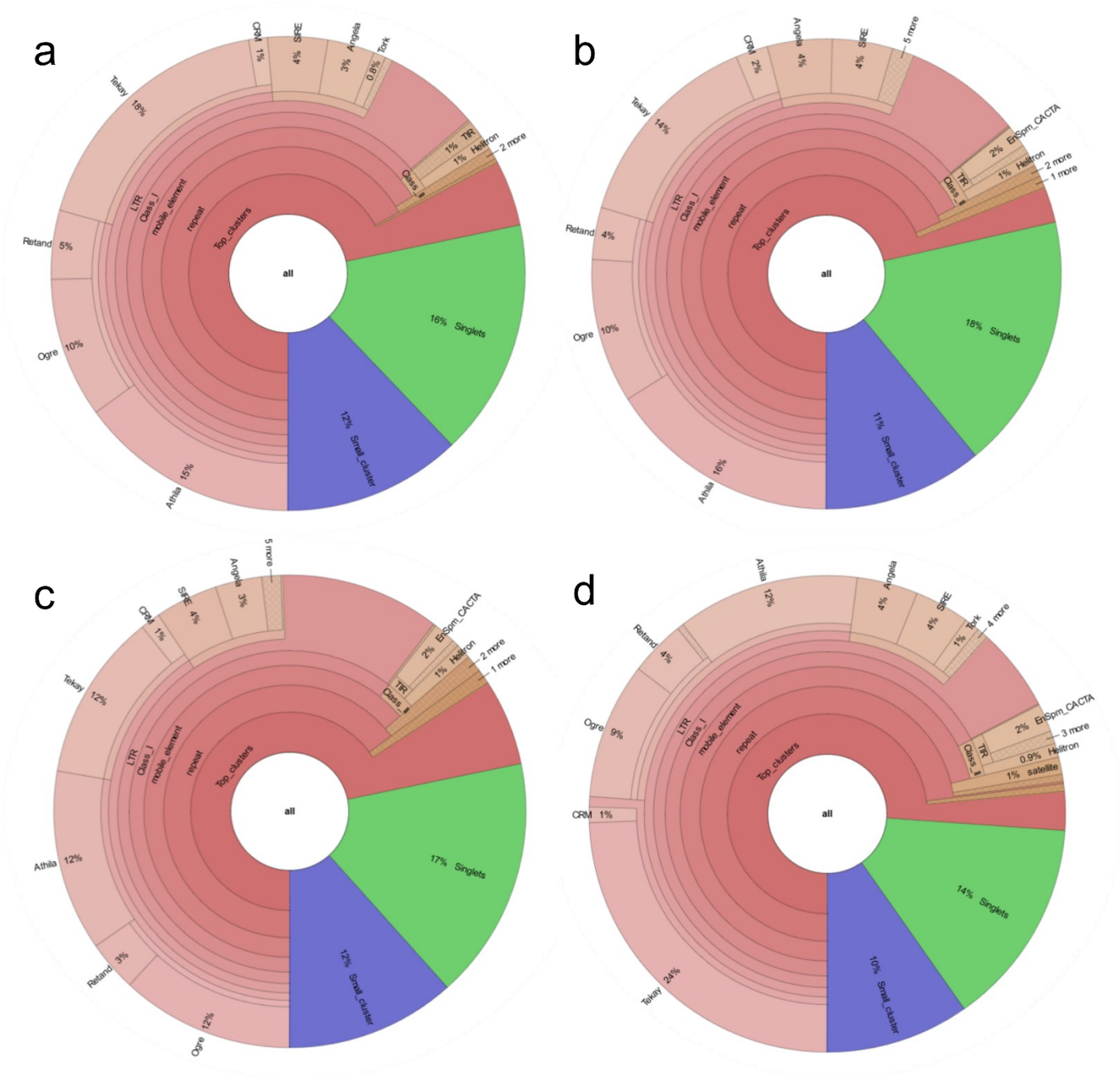


Figure 10: Single species repeat profiles of the four European *Polygonatum* species. (a) *P. latifolium*, (b) *P. multiflorum*, (c) *P. odoratum* and (d) P. *verticillatum*; top clusters in red, singlets in green and small clusters in blue; genome proportions given in percentages [%].

P. latifolium

The analyses of NGS data of *P. latifolium* revealed that 71.6% of the genome were represented by repeats (*Figure 9*; *Table 4*). The largest proportion of repetitive sequences was represented by retroelements (Class I) making up 64.0% of the genome. Ty3/*gypsy* elements (48.7%) were dominating, with Tekay lineage (18.1%) followed by Athila (15.1%), Ogre (9.6%; *Figure 11a*) and Retand lineages (4.7%). CMR Ty3/*gypsy* elements constituted only 1.3% of the genome. The Ty1/*copia* lineage was less abundant (8.7%) and mostly represented by SIRE (4.1%), and Angela (3.2%) lineages. DNA transposons were only present in low proportions (2.2%) with CACTA being most abundant (1.2%), followed by hAT (0.3%), Mutator (0.2%) and MITEs (0.1%). Helitrons (Class II, Subclass 2) represented 1.0% of the genome. 6.5% of the genome were identified as LTR retrotransposons but could not be annotated to the specific lineage. Non-annotated repetitive DNAs amounted to 4.5%. In addition, 0.8% of the genome was classified as tandem repeats, representing 35S rDNA (0.4%), 5S rDNA (0.01%) and satDNAs (0.4%).

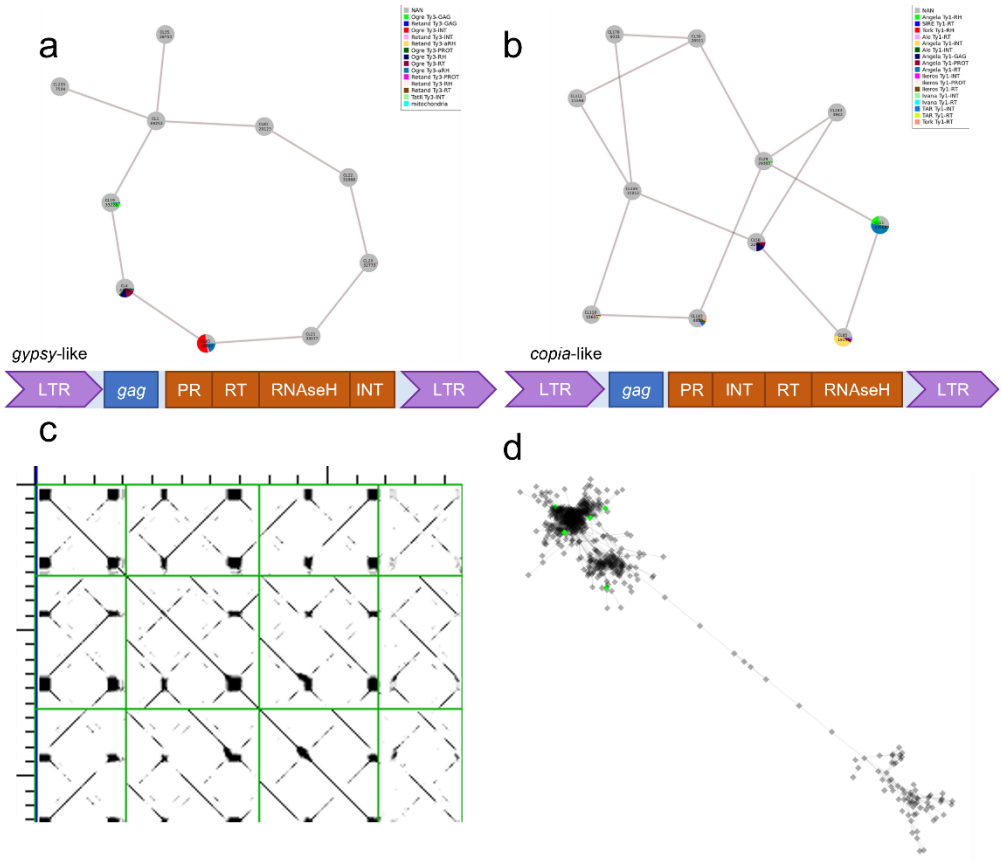


Figure 11: Examples of clusters and superclusters representing various types of repeats recovered in single species analyses and comparative analyses of the repeatomes of the four European *Polygonatum* species. (a and b) Superclusters representing (a) an abundant Ty3/*gypsy* Ogre element in *P. latifolium* genome and (b) an Ty1/*copia* Angela element in *P. odoratum* genome; (c-d) (c) graph and (d) dotter of cluster CL324 representing a MITE element recovered in the *P. verticillatum* genome.

Table 4: Repeat composition of *P. latifolium* genome.

Repeat Type	*Superfamily*	*Family*	*Genome Proportion [%]*	*No. of reads*
Retrotransposons			**64.0**	3911603
	Ty1-copia		**8.7**	529611
		SIRE	**4.1**	250351
		Angela	**3.2**	195706
		Tork	**0.8**	50906
		Ale	**0.3**	16772
		TAR	**0.2**	11326
		Bianca	**0.05**	3660
		Ivana	**0.01**	890
	Ty3-gypsy		**48.7**	2973204
		Tekay	**18.1**	1102173
		Athila	**15.1**	920263
		Ogre	**9.6**	584869
		Retand	**4.7**	287617
		CRM	**1.3**	78282
	Other/non-LTR		**6.7**	
		Other-LTR	**6.5**	399054
		Pararetrovirus	**0.2**	9734
DNA-transposons			**2.2**	135458
	Subclass 1		**1.2**	75335
		CACTA	**0.5**	33400
		hAT	**0.3**	20219
		Mutator	**0.2**	14582
		MITE	**0.1**	7134
	Subclass 2		**1.0**	60123
		Helitron	**1.0**	60123
Tandem repeats			**0.8**	47866
	rDNA		**0.4**	25854
		35S rDNA	**0.4**	25214
		5S rDNA	**0.01**	640
	Satellite DNAs		**0.4**	22012
Unclassified			**4.5**	275458
Total repeats			**71.6**	4370385

P. multiflorum

Repeats made 71.2% of the genome of *P. multiflorum* (*Figure 9*; *Table 5*). The most abundant repetitive sequences were represented by retroelements (Class I) with 64.8%. In this class, Ty3*/gypsy* elements dominated (46.3%) with highest proportion of Athila (16.2%), followed by Tekay (14.3%), Ogre (10.0%), Retand (3.6%) and CRM (2.2%). The next largest group was represented by the superfamily Ty1*/copia* (10.2%), mostly by Angela (4.5%) and SIRE (4.2%) lineages. DNA transposons made 3.1% of the genome, with Subclass 1 (2.1%) represented by CACTA (1.6%), Mutator (0.4%) and hAT elements (0.2%), and Subclass 2 by Helitrons (1.0%). 8.1% of the repeats were identified as LTR-retroelements, but could not

be further classified, and 2.5% of the repeatome remained unclassified. Tandem repeats made up 0.8% of the genome with 35S rDNAs(0.4%), the 5S rDNAs(0.01%) and satDNAs (0.4%).

Table 5: Repeat composition of *P. multiflorum* genome.

Repeat Type	*Superfamily*	*Family*	*Genome Proportion [%]*	*No. of reads*
Retrotransposons			**64.8**	3959546
	Ty1-copia		**10.2**	624825
		Angela	**4.5**	274714
		SIRE	**4.2**	256722
		Tork	**0.8**	46883
		Ale	**0.4**	22056
		TAR	**0.3**	18037
		Bianca	**0.08**	4848
		Ivana	**0.03**	1565
	Ty3-gypsy		**46.3**	2830467
		Athila	**16.2**	989469
		Tekay	**14.3**	872595
		Ogre	**10.0**	611370
		Retand	**3.6**	220476
		CRM	**2.2**	136557
	Other/non-LTR		**8.2**	503546
		Other-LTR	**8.1**	496324
		Pararetrovirus	**0.2**	9734
DNA-transposons			**3.1**	190880
	Subclass 1		**2.1**	131164
		CACTA	**1.6**	94821
		Mutator	**0.4**	21492
		hAT	**0.2**	14851
	Subclass 2		**1.0**	59716
		Helitron	**1.0**	59716
Tandem repeats			**0.8**	49363
	rDNA		**0.4**	27052
		35S rDNA	**0.4**	26281
		5S rDNA	**0.01**	771
	Satellite DNAs		**0.4**	22311
Unclassified			**2.5**	151532
Total repeats			**71.2**	4351321

P. odoratum

Repeats were inferred to represent 71.5% of the genome of *P. odoratum* (*Figure 9*; *Table 6*). Retroelements were found in the highest proportion (61.2%) and among these Ty3/*gypsy* elements were most abundant (41.0%). The biggest superfamily of the Ty3/*gypsy* elements was Athila (12.5%), followed by Ogre (12.0%) and Tekay (11.6%). Less abundant were elements of Retand (3.5%) and CRM (1.4%) lineages. 9.0% of the repeatome consisted of

Ty1/*copia* elements, with SIRE (4.2%) and Angela (3.2%; *Figure 11b*) being most abundant. DNA transposons represented 3.1% of the repetitive sequences in the genome, with 2.0% in Subclass 1 (1.5% of CACTA, 0.2% of MITEs, 0.2% of hAT, 0.08% of Mutator) and 1.1% in Subclass 2 (Helitrons). 11.1% of repeats could be identified as LTR-retrotransposon, but could not be further classified, and 6.0% of repeats could not be classified at all. 1.2% of the genome were found to represent tandem repeats, with 0.8% of 35S rDNA, 0.01% of 5S rDNA, and 0.4% of satDNAs.

Table 6: Repeat composition of *P. odoratum* genome.

Repeat Type	***Superfamily***	***Family***	***Genome Proportion [%]***	***No. of reads***
Retrotransposons			**61.2**	3550607
	Ty1-copia		**9.0**	523941
		SIRE	**4.3**	251780
		Angela	**3.2**	186490
		Tork	**0.8**	43895
		Ale	**0.3**	16686
		TAR	**0.2**	10847
		Bianca	**0.08**	4412
		Ivana	**0.02**	1008
	Ty3-gypsy		**41.0**	2378849
		Athila	**12.5**	725670
		Ogre	**12.0**	695244
		Tekay	**11.6**	674440
		Retand	**3.5**	204049
		CRM	**1.4**	79446
	Other/non-LTR		**11.2**	647817
		Other-LTR	**11.1**	646365
		Pararetrovirus	**0.2**	10275
DNA-transposons			**3.1**	178673
	Subclass 1		**2.0**	117298
		CACTA	**1.5**	89681
		MITE	**0.2**	12517
		hAT	**0.2**	10634
		Mutator	**0.08**	4466
	Subclass 2		**1.1**	61375
		Helitron	**1.1**	61375
Tandem repeats			**1.2**	68885
	rDNA		**0.8**	48195
		35S rDNA	**0.8**	47544
		5S rDNA	**0.01**	651
	Satellite DNAs		**0.4**	20690
Unclassified			**6.0**	346744
Total repeats			**71.5**	4144909

P. verticillatum

Repetitive sequences constituted 76% of the genome of *P. verticillatum* (*Figure 9; Table 7*). The highest portion of these sequences was represented by retroelements (67.8%) and within those, by Ty3*/gypsy* elements (52.3%). Ty3/*gypsy* elements of the Tekay lineage (24.5%) were most abundant, followed by Athila (12.4%), Ogre (9.3%), Retand (4.2%) and CMR (1.1%). The second biggest repeat class was represented by Ty1/*copia* elements (9.7%) with Angela being most abundant (4.2%), followed by SIRE (3.7%) and Tork (1.0%). DNA transposons amounted to 3.8%, of those 2.8% in Subclass 1 (1.4% of CACTA 0.4% of Mutator, 0.3% of hAT and 0.1% of MITEs [*Figure 11c-d]*) and 1.0% in the Helitron lineage (Subclass 2). Within the LTR-retrotransposons, 5.7% could not further be annotated and 2.6% of the repetitive sequences remained unclassified. Tandem repeats made up 1.7% of the genome of *P. verticillatum* and were represented by 35S rDNA (0.5%), 5S rDNA (0.03%) and satDNAs (1.2%).

Table 7: Repeat composition of *P. verticillatum*.

Repeat Type	*Superfamily*	*Family*	*Genome Proportion [%]*	*No. of reads*
Retrotransposons			**67.8**	4009251
	Ty1-copia		**9.7**	575253
		Angela	**4.2**	246281
		SIRE	**3.7**	221213
		Tork	**1.0**	61083
		Ale	**0.4**	22978
		TAR	**0.3**	18912
		Bianca	**0.05**	2765
		Ivana	**0.03**	2021
	Ty3-gypsy		**52.3**	3092363
		Tekay	**24.5**	1450751
		Athila	**12.4**	735759
		Ogre	**9.3**	550147
		Retand	**4.2**	250730
		CRM	**1.1**	62996
	Other/non-LTR		**5.8**	341635
		Other-LTR	**5.7**	336783
		Pararetrovirus	**0.08**	4852
DNA-transposons			**3.8**	221851
	Subclass 1		**2.8**	165609
		CACTA	**1.9**	113820
		Mutator	**0.4**	24252
		hAT	**0.3**	20396
		MITE	**0.1**	7141
	Subclass 2		**1.0**	56242
		Helitron	**1.0**	56242
Tandem repeats			**1.7**	98238
	rDNA		**0.5**	28959
		35S rDNA	**0.5**	27326
		5S rDNA	**0.03**	1633
	Satellite DNAs		**1.2**	69279
Unclassified			**2.6**	156700
Total repeats			**76.0**	4494603

Comparative Analysis

The comparative analyses allowed further insight in the repeat composition and evolution of the four *Polygonatum* species (*Table 8*). The highest portion of repetitive sequences was found in the genome of *P. multiflorum* (76.6%) followed by *P. odoratum* (74.4%), *P. latifolium* (72.9%) and *P. verticillatum* (72.3%). Retroelements were the most abundant type and ranged from 62.0% in *P. verticillatum* to 67.7% in *P. multiflorum*. Ty3/*gypsy* elements were highly abundant with 47.1% in *P. multiflorum* to 45.3% in *P. latifolium*. Elements of the Tekay lineage were more abundant in *P. verticillatum* (24.8%) than in the other three species (18.1% in *P. latifolium*, 15.8% in *P. multiflorum*, 17.1% in *P. odoratum*). In

P. multiflorum, elements of the Athila superfamily were more abundant than those of the Tekay lineage (17.7%). Hence, Athila elements were the second largest superfamily of Ty3/*gypsy* elements in *P. latifolium* (15.8%), *P. odoratum* (14.1%) and *P. verticillatum* (12.3%). The third largest superfamily of Ty3/*gypsy* elements was the Ogre lineage ranging from 6.5% in *P. verticillatum* to 12.5% in *P. odoratum*. More than 10% of repetitive elements were represented by Ty1/*copia* elements (10.6% in *P. latifolium*, 13.8% in *P. multiflorum*, 11.1% in *P. odoratum*, 11.4% in *P. verticillatum*), mainly SIRE and Angela lineages. Approximately 3% of the genome consisted of DNA transposons (3.2% in *P. latifolium*, 3.4% in *P. multiflorum*, 3.4% in *P. odoratum*, 3.7% in *P. verticillatum*) with CACTA being most abundant in all four species. Between 3.8% in *P. verticillatum* and 7.8% in *P. latifolium* of LTR-retrotransposons could not further be identified. The number of unclassified sequences was similar in all four species around 5%.The highest percentage of tandem repeats was found in *P. verticillatum* (1.9%), followed by *P. odoratum* (1.3%) and *P. latifolium* and *P. multiflorum*, both with 0.9%.

The three species of sect. *Polygonatum*, *P. latifolium, P. multiflorum* and *P. odoratum*, were more similar to one another than any of them was to *P. verticillatum* (*Figure 12*). The first few unannotated clusters shown in *Figure 12* had high representation in the genome of *P. verticillatum* but were nearly absent in the other three species. In contrast, some other clusters, e.g., unannotated cluster CL320, were present in *P. latifolium* (0.04%), *P. multiflorum* (0.02%) and *P. odoratum* (0.02%) but absent in *P. verticillatum*. Some clusters were also specific for other species. CL120 (Tekay, 0.7%) and CL156 (unannotated, 0.6%) were well represented in *P. odoratum*, but were less abundant in the remaining three species (less or equal 0.1%). Similarly, CL245 (unannotated) was mostly composed of *P. latifolium* reads (0.2%). Repeat profiles of *P. latifolium* and *P. multiflorum* in general were more similar to each other than to *P. odoratum.* However, one cluster (CL160) was shared by *P. latifolium* (0.05%), *P. odoratum* (0.02%) and by *P. verticillatum* (0.8%), the latter with much higher read number, but did not include reads of *P. multiflorum.* This cluster represented a satDNA, PolSat2, and was not found in the comparative analysis of *P. multiflorum* (*Figure 12*; *Table 9, Table 10*). The other clusters, which are absent in at least one of the four species, were mostly unannotated (*Figure 12*).

Table 8: Proportions of repetitive DNA elements per species inferred from the comparative analyses of all four European *Polygonatum* species; *P. latifolium* (*P. lat.*), *P. multiflorum* (*P. mul.*), *P. odoratum* (*P. odo.*), *P. verticillatum* (*P. ver.*).

Repeat Type	*Superfamily*	*Family*	*P. lat.*	*P. mul.*	*P. odo.*	*P. ver.*
Retrotransposons			**63.8**	**67.7**	**64.7**	**62.0**
	Ty1-copia		**10.6**	**13.8**	**11.1**	**11.4**
		SIRE	**4.2**	**4.6**	**4.8**	**4.1**
		Angela	**3.6**	**5.1**	**3.6**	**4.4**
		Tork	**0.8**	**0.8**	**0.7**	**1.0**
		Ale	**0.3**	**0.4**	**0.3**	**0.4**
		TAR	**0.2**	**0.3**	**0.2**	**0.3**
		Bianca	**0.07**	**0.08**	**0.08**	**0.04**
		Ivana	**0.01**	**0.03**	**0.02**	**0.03**
	Ty3-gypsy		**45.3**	**47.1**	**46.6**	**46.8**
		Tekay	**18.1**	**15.8**	**17.1**	**24.8**
		Athila	**15.8**	**17.7**	**14.1**	**12.3**
		Ogre	**7.8**	**10.4**	**12.5**	**6.5**
		Retand	**3.6**	**3.1**	**2.9**	**3.2**
		CRM	**1.4**	**2.5**	**1.4**	**1.1**
	Other/non-LTR		**8.0**	**6.8**	**7.1**	**3.9**
		Other-LTR	**7.8**	**6.7**	**6.9**	**3.8**
		Pararetrovirus	**0.2**	**0.1**	**0.2**	**0.1**
DNA-transposons			**3.2**	**3.4**	**3.4**	**3.7**
	Subclass 1		**2.3**	**2.4**	**2.3**	**2.7**
		CACTA	**1.7**	**1.7**	**1.7**	**2.0**
		Mutator	**0.3**	**0.3**	**0.3**	**0.3**
		hAT	**0.2**	**0.2**	**0.2**	**0.3**
		MITE	**0.1**	**0.1**	**0.1**	**0.1**
	Subclass 2		**0.9**	**1.0**	**1.1**	**1.0**
		Helitron	**0.9**	**1.0**	**1.1**	**1.0**
Tandem repeats			**0.9**	**0.9**	**1.3**	**1.9**
	rDNA		**0.5**	**0.5**	**0.9**	**0.6**
		35S rDNA	**0.5**	**0.5**	**0.9**	**0.5**
		5S rDNA	**0.01**	**0.01**	**0.02**	**0.03**
	Satellite DNAs		**0.4**	**0.4**	**0.4**	**1.3**
Unclassified			**5.0**	**4.5**	**5.0**	**4.7**
Total repeats			**72.9**	**76.6**	**74.4**	**72.3**

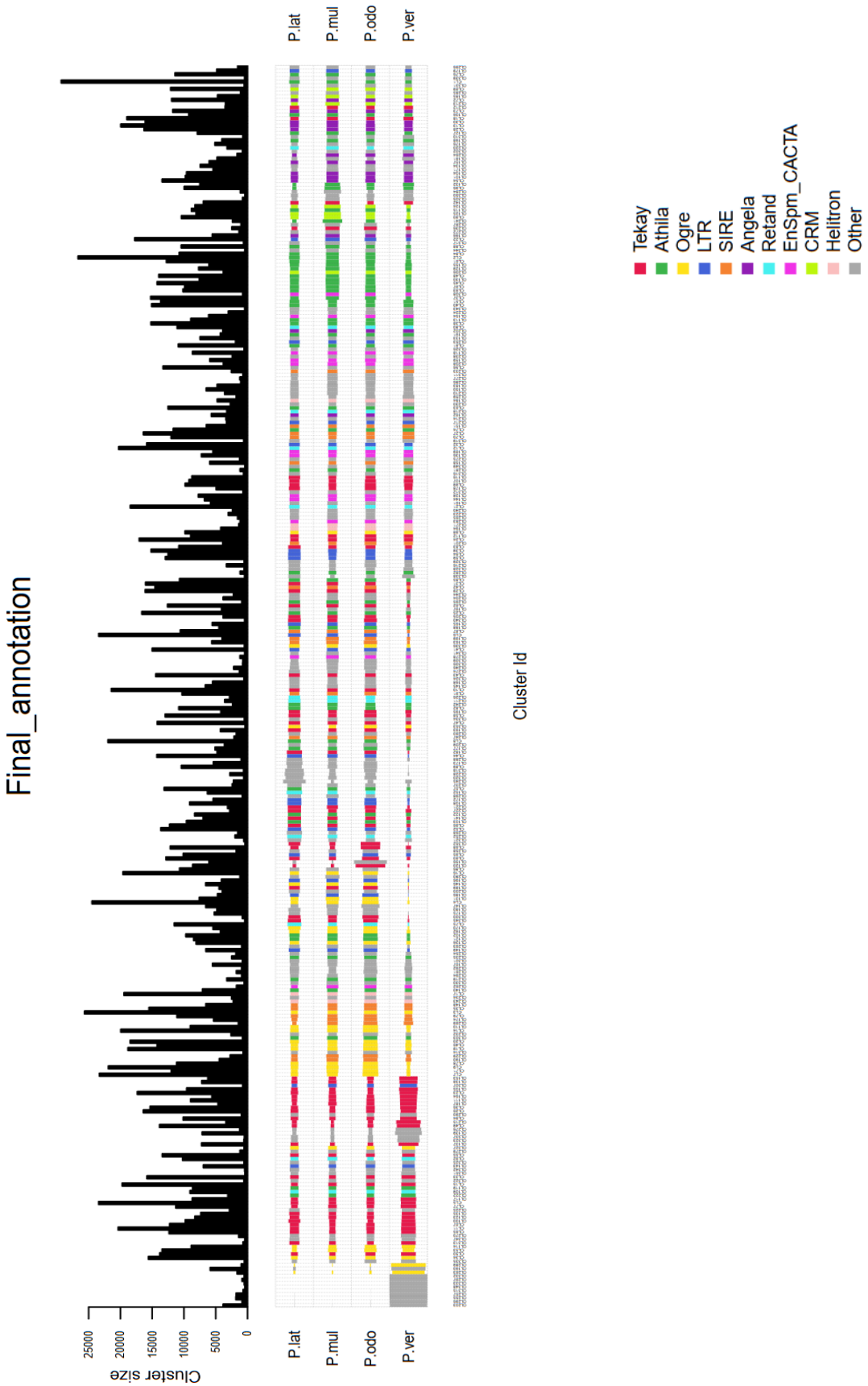


Figure 12: Comparative analysis of all four *Polygonatum* species, showing the differences in repeat proportions. *P. latifolium* (P. lat), *P. multiflorum* (P. mul), *P. odoratum* (P. odo) and *P. verticillatum* (P. ver); "Other" indicates unannotated clusters.

Satellite DNAs

Two different satDNAs with a monomer size of 69 bp and 176 bp, respectively, were identified in *Polygonatum*. PolSat1 was recovered both in the comparative analyses and in all single species analyses (*Table 9*). This satDNA of a 69 bp long monomer was reconstructed as a very dense spherical cluster in the comparative analysis (*Table 10*). The second satDNA, PolSat2, with monomer size of 176 bp was recovered in the comparative analysis and encompassed reads of *P. latifolium*, *P. odoratum* and *P. verticillatum*. It was also recovered in the individual analyses of two species, *P. latifolium* and *P. verticillatum*. This satDNA was most abundant in *P. verticillatum* and was completely absent in *P. multiflorum* (*Table 10; Figure 12*). Reads of *P. latifolium* and *P. odoratum* were also recovered in this satellite cluster, albeit with lower read numbers. All sequences of satDNAs found in the single species analyses were plotted against the sequences of the satDNAs identified in comparative analyses. SatDNAs recovered in single species analyses were identical to the satDNAs of the same monomer length identified in comparative analyses (*Figure 13*). Hence, European species of genus *Polygonatum* harbor two newly identified satDNAs: PolSat1 and PolSat2 were found in *P. latifolium*, *P. odoratum* and *P. verticillatum*, and PolSat1 in *P. multiflorum* (*Table 9*).

Table 9: Presence-absence table indicating the occurrence (black bar) of satDNAs in the four *Polygonatum* species.

	P. latifolium	*P. multiflorum*	*P. odoratum*	*P. verticillatum*
PolSat1				
PolSat2				

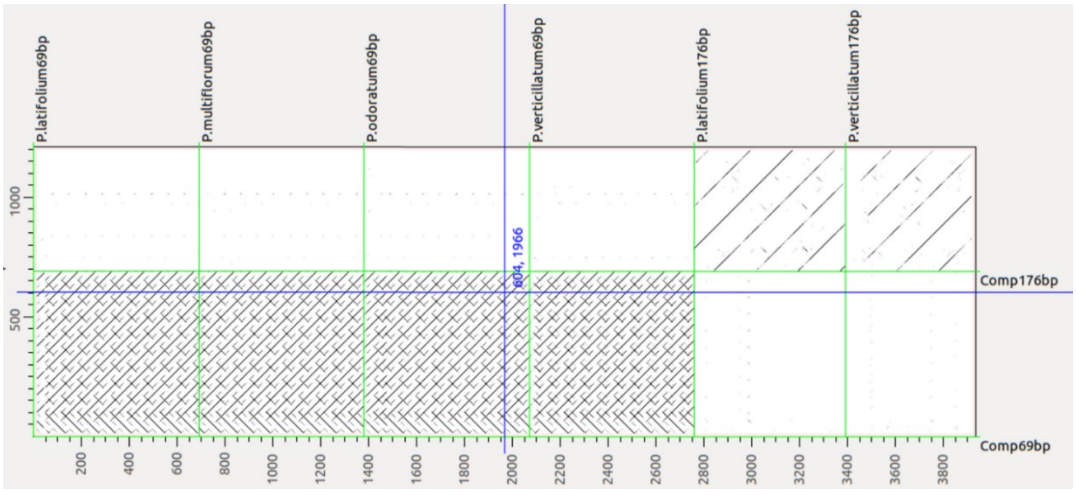
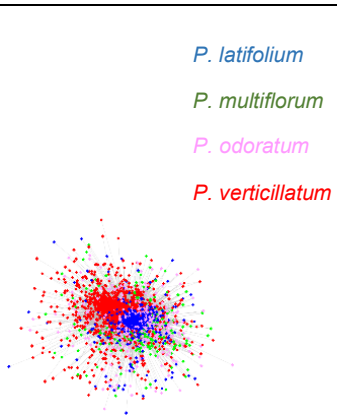
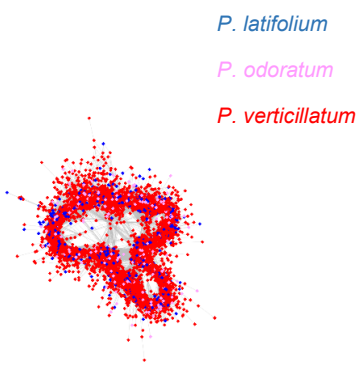


Figure 13: Dotter of the reconstructed satDNA sequences derived from the single species analyses (y-axis) against the satDNA sequences identified in the comparative analyses of the four *Polygonatum* species (x-axis). Comp69bp: PolSat1 identified in comparative analysis, Comp176bp: PolSat2 identified in comparative analysis.

Table 10: Characterization of two novel satDNAs identified in the comparative analysis of the four *Polygonatum* species.

satDNA	Monomer length	Reconstructed monomer sequence	Cluster graph with abundances in analyzed species
PolSat1	69bp	TCGATTAGTGATTAGGATCGGGTCTATTGAGGGAACCGATATTGATTAGTAT CGGGTCCTTCAATGAAC	
PolSat2	176bp	ATGGTTTGGAGTTAGAATTTATTAGGAAAAAATACGCCAAAATTCAAAAACC AGTCTTTATCCATCCGAATGCCTCGAGAATTGGTCGTTAAGCTCCATTTAGT GAGCGCATACATGTGTCCAGGTCTCGTTTTCTTGCAATGGATGGTTTTAGA ATTACGTTTATAATTTGTATTA	

Discussion

Little is known about genomes of European *Polygonatum* species. As the diversity hotspot of the genus is found in Asia, most phylogenetic and cytological studies have dealt with Asian species and accessions. In addition, the information about the repeatomes of this genus, whether in European, American, or Asian species, is completely missing. Based on the methods applied in this study, the four European species could clearly be distinguished from each other, the repeat profiles could be established, and their phylogenetic relationship could be inferred.

Chromosome numbers and genome sizes

The chromosome numbers of several different *Polygonatum* species have been reported previously (e.g. Therman, 1953; Weiss-Schneeweiss and Jang, 2003). Chromosome numbers obtained in this study for the four analyzed species were largely consistent with previous studies (Therman, 1953; Weiss-Schneeweiss and Jang, 2003). The karyotypes presented here were also similar to previously reported karyotypes of European species (Therman, 1953) and showed a clear difference between the two sections of the genus as well as differences between the two karyotype types identified earlier (*latifolium*-type and *odoratum-type*; Suomalainen, 1947; Therman, 1953; Tamura et al., 1997). Some of the previous reports of chromosome numbers for Chinese *P. odoratum* and *P. verticillatum* plants (Shao et al., 1993; Deng et al., 2009) indicated other chromosome numbers. The first study described two *P. odoratum* populations from different locations with either $2n = 16$ or $2n = 18$ and heterozygosity regarding the presence of satellites/NORs (Shao et al., 1993). *P. verticillatum* was reported to possess two cytotypes with chromosome numbers of either $2n = 18$ or $2n = 24$ (Shao et al., 1993). Another study of Chinese material reported a chromosome number of $2n = 30$ for *P. verticillatum* (Deng et al., 2009). Some studies reported that *P. multiflorum* and *P. odoratum* could form triploid hybrids (Vestergren, 1925; Therman, 1953) with chromosome numbers of $2n = 28$, 29 or 30 (Therman, 1953). As the diversity hotspot of the genus is hypothesized to be in China, the karyotype diversity, including heterozygosity and the presence of different chromosome numbers, including polyploidy, and potentially also the number of hybrids or hybrid populations might be higher in Asian than in European populations. To clearly understand the differences between individuals of the same species of European and Chinese material, further studies are needed.

Comparison of chromosome numbers and genome sizes in analyzed *Polygonatum* species has revealed that there is no correlation between the two characters. Former studies indicated that the genome sizes can be negatively correlated with the chromosome numbers in certain families (Melanthiaceae; Pellicer et al., 2014) or genera (*Carex*; Nishikawa et al., 1984). The latter correlation was hypothesized to be connected to the fragmentation of large holocentric chromosomes into smaller ones, as *Carex* carried holocentric chromosomes (Nishikawa et al., 1984). In contrast, no or only a weak positive or negative correlation between genome size and chromosome numbers at higher phylogenetic levels was reported (Chung et al., 2012). In the genus *Polygonatum*, considering only four European species, the species with the highest chromosome number, *P. verticillatum*, had the smallest chromosomes and the smallest genome size compared to the other three species (*Figure 14, Figure 15*). However, in the three remaining species, *P. latifolium, P. odoratum* and *P. multiflorum*, no negative trend could be inferred. *P. multiflorum,* with the lowest chromosome number of 2*n* = 18 had the second largest genome size. Although the European species did not show any trend regarding genome size and chromosome number (*Figure 14*), no clear correlation can be drawn for the whole genus, as the data from the diversity hotspot in China are missing. In addition, the genome sizes of the four analyzed species were distinct and thus can provide an additional diagnostic character.

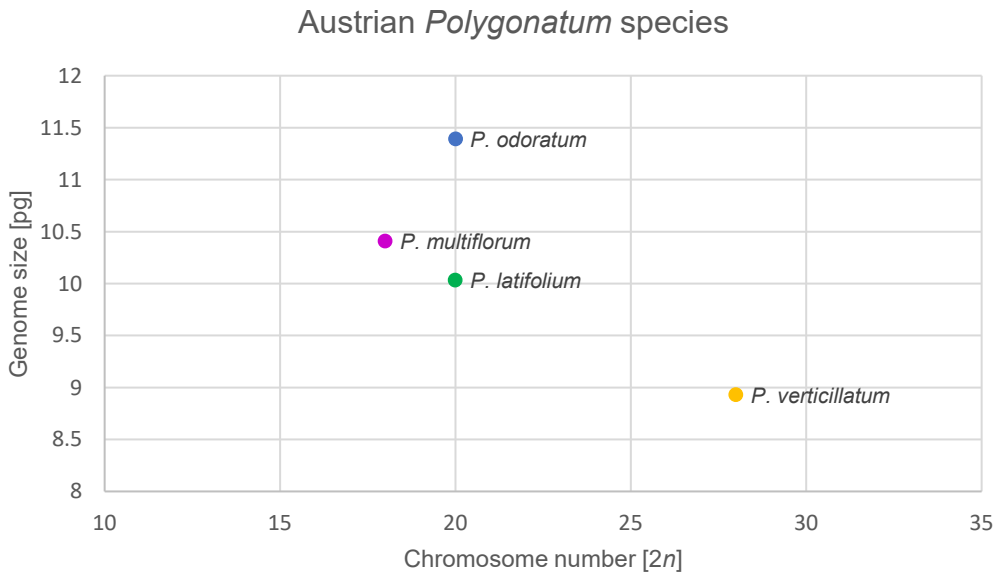


Figure 14: Correlation between chromosome number and genome size in the four European *Polygonatum* species. Colors indicate species: *P. latifolium* (green), *P. multiflorum* (purple), *P. odoratum* (blue) and *P. verticillatum* (yellow).

Evolution of rDNA loci

The position of the 35S rDNA loci of *P. latifolium* and *P. multiflorum* was expected to be similar, as both species have a *latifolium*-type karyotype (Suomalainen, 1947; Therman, 1953). This was indeed the case for the analyzed individuals of *P. latifolium* and *P. multiflorum* (*Figure 15*). *P. odoratum*, on the other hand, has an *officinale*-type karyotype (Suomalainen, 1947) and was expected to carry the 35S rDNA loci on the long arms of a submetacentric chromosome pair and in pericentromeric position in a metacentric pair (Therman, 1953; *Appendix 2*). This was not observed in this study, as the position of the 35S rDNA locus was not in the pericentromeric region in the analyzed individuals (*Figure 15*; *Appendix 2b*). The position of the 5S rDNA locus was reported here for the first time for European *Polygonatum* species: it was clearly located on one of the small metacentric chromosome pairs in all four species. As FISH was not used in previous studies, the position of the 5S rDNA locus/loci in Asian or American species is still unknown.

Previous studies reported chromosomal polymorphisms in many species of the (former) family Liliaceae (Greilhuber, 1995) or in *Polygonatum* itself (Weiss-Schneeweiss and Jang, 2003; Deng et al., 2009). Weiss-Schneeweiss and Jang (2003) reported an asymmetry in the length of the long arm of a chromosome pair as well as the loss of a distal segment next to the secondary constriction in *P. odoratum*. In the same species, Deng et al. (2009) described heterozygosity of the satellites attached to the long arm of the first and second chromosome pair, respectively. Such polymorphisms were not observed in the currently analyzed populations of *P. odoratum* or any other of the three species. It is likely that analyses of more populations and individuals will reveal polymorphisms. However, a difference in the signal strength between the 35S rDNA-carrying chromosome pairs was found in an individual of one population collected at the Hundsheimer Berge. As only one individual per population was examined, it is unclear whether this difference is an exception or the rule.

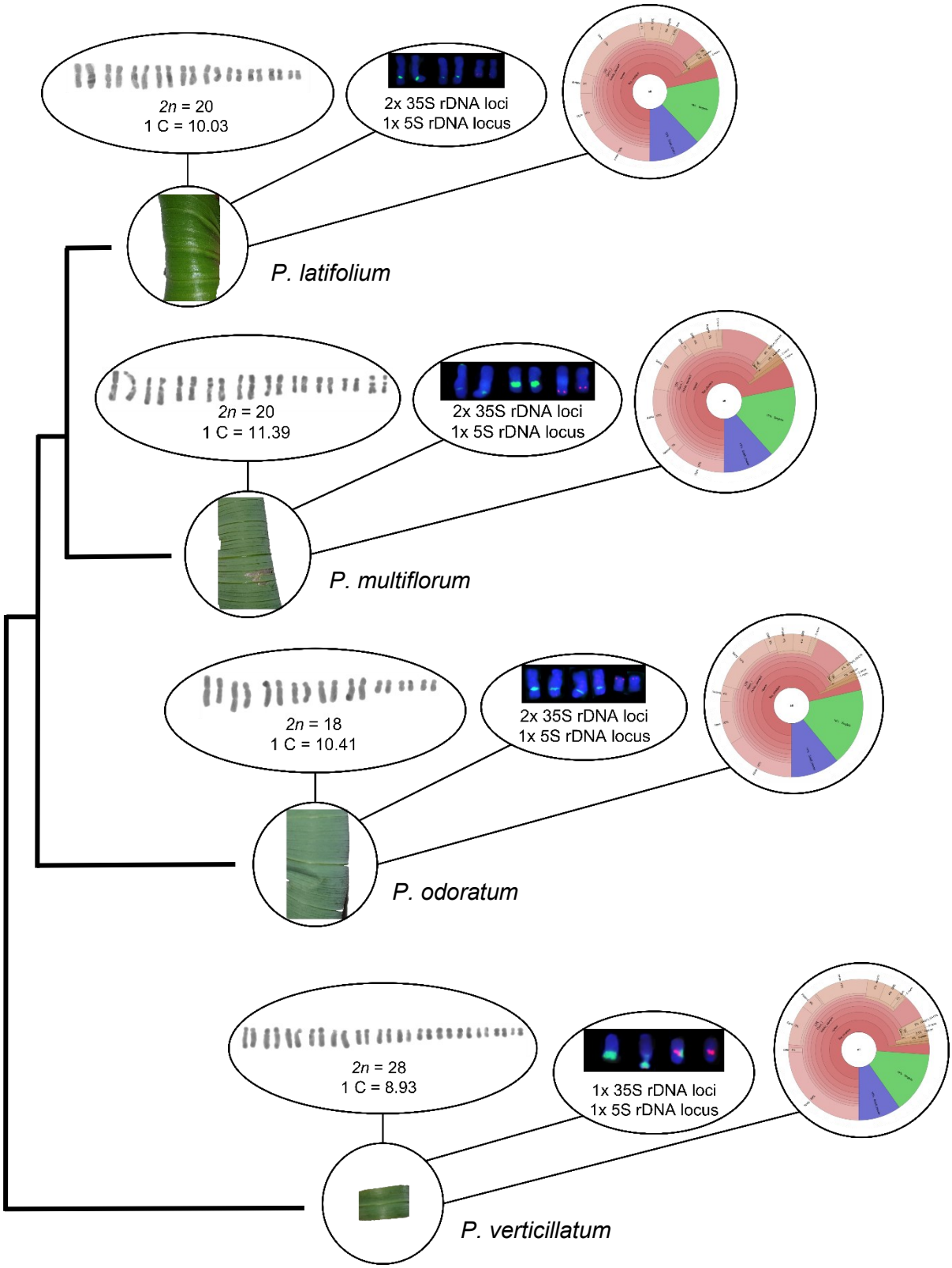


Figure 15: Summary of the data of the four European *Polygonatum* species, including the karyotype, chromosome number, genome size, chromosome pairs with rDNA loci and repeat landscapes. (a) *P. latifolium*, (b) *P. multiflorum*, (c) *P. odoratum* and (d) *P. verticillatum*.

Repeatome evolution

The analyses of NGS data of *P. latifolium, P. multiflorum*, *P. odoratum* and *P. verticillatum* provided first insight in the repeatome evolution of these four species. Repeatome profiles can be compared with the other genomic characters analyzed in the current study, genome sizes, chromosome numbers and karyotypes. Previous studies of other plant groups clearly indicated that there is a correlation between genome size and the amount of repetitive DNAs in the genomes of plants, particularly number of copies of transposable elements and in some cases also satDNAs, especially in the absence of polyploidy. Small genomes typically have fewer copies of large transposable elements, mostly LTR-retrotransposons, than bigger genomes (Civán et al., 2011, Novák et al., 2014). LTR retrotransposons in the genome use “copy-and-paste” mechanisms and each activation results in genome size increase (Kim 2017). Thus, Ty1*/copia* and Ty3*/gypsy* elements were shown to be most abundant in the majority of plant genomes (Kejnovský et al., 2012; Vitte et al., 2013; Macas et al., 2015). The same patterns were also observed in the current study, as elements of the type Ty3*/gypsy* made up the highest proportion of the repeatome of all four species of *Polygonatum* and the genome proportion of all LTR-retrotransposons in all four species was at or above 60%. Whether Ty3/*gypsy* or Ty1/*copia* elements dominate in the analyzed genomes, is strongly group-specific with no particular pattern observed across the families of angiosperms (Weiss-Schneeweiss et al., 2015). In contrast to the four *Polygonatum* species, the *Musa* genome, a monocotyledonous herb found in the tropics and subtropics, showed a dominance of Ty1*/copia* elements, mostly Maximus/SIRE and Angela lineages (Novák et al., 2014). However, the overall proportion of repetitive sequences is largely similar (and high) across the monocotyledonous families (Novák et al. 2014). This is also the case in *Polygonatum* species.

Previous study hypothesized that LTR-retrotransposons of the Ogre lineage may only be present in eudicots, and absent in non-eudicots (Macas et al., 2015). This very long LTR-retrotransposon was first described in the family of Fabaceae, where it can make up to 40% of the genome (Macas and Neumann, 2007). However, the current study identified Ogre elements in all four *Polygonatum* species, with high abundance (*Figure 15*). Those elements have not been reported in other monocot species to date, but analyses of more genomes might provide evidence to the contrary.

The comparative analyses of the repeatomes of all four species of *Polygonatum* revealed significant differences in their repeat profiles. *P. latifolium*, *P. multiflorum* and *P. odoratum*

had very similar profiles, while *P. verticillatum* had different abundances of various repetitive elements. This further supports the hypothesis that the first three species placed in sect. *Polygonatum* are closer related to one another than to *P. verticillatum* in sect. *Verticillata* (Meng et al., 2014). Phylogenetic analyses of ITS1 and ITS2 DNA sequences performed in this study also confirmed the clear distinction of the two groups (*Figure 15*). Repeat profiles have been shown to differ more in phylogenetically more distant plant groups. Three of the four species of sect. *Repandae* in genus *Nicotiana* (Solanaceae), which are closer related to each other, had similar repeat profiles, whereas the fourth and phylogenetically more distant species, *N. nudicaulis*, had a quite different repeat profile (Dodsworth et al., 2016). As an example of a monocotyledonous genus, thus closer related to *Polygonatum* than *Nicotiana*, serves *Musa* (Musaceae). Among the studied species, clear differences in the repeat profiles could be observed and therefore, phylogenetic relationships drawn in earlier studies could be supported (Novák et al., 2014). The species *Ensete gilletii*, an ornamental banana distantly related to other *Musa* species analyzed, had most distinct and different repeats (Novák et al., 2014). Within the *Musa* species, the phylogenetic signal was also clear in repeat profiles with closely related *M. accuminata* and *M. ornata* having very similar profiles and early diverging *M. balbisiana* having a more different prolife of repeats (Novák et al., 2014).

Similar patterns can be inferred from the presence and copy numbers of satDNAs in *P. verticillatum* and the three other species. *P. verticillatum* had a higher proportion of satDNAs than the other three species. The two satDNAs found in this study, PolSat1 and PolSat2, had very similar monomer sequences in all species, suggesting their common origin in the ancestor of the two sections. Mapping of these novel satDNAs in chromosomes of all four species will provide more information about their chromosomal distribution and will allow more detailed analyses of the karyotype evolution among the four species.

Combining all gathered data, the genomic distinction between the four European *Polygonatum* species was clearly supported. Thus, chromosome numbers, genome size and repeat profiles as well as the phylogenetic tree reveal clear differences between *P. verticillatum* of sect. *Verticillata* and the other three species of sect. *Polygonatum*, as well as the clear distinction between *P. odoratum* and the two closer related species *P. latifolium* and *P. multiflorum*. Hence, it supports the current taxonomic classification of these species.

Literature

Choi B, Weiss-Schneeweiss H, Temsch EM, So S, Myeong HH, Jang TS. **2020a**. Genome size and chromosome number evolution in Korean *Iris* L. species (Iridaceae Juss.). *Plants* (Basel) **9**: 1284.

Choi IY, Kwon EC, Kim NS. **2020b**. The C- and G-value paradox with polyploidy, repeatomes, introns, phenomes and cell economy. *Genes and Genomics* **42**: 699–714.

Christenhusz MJM, Byng JW. **2016**. The number of known plant species in the world and its annual increase. *Phytotaxa* **261**: 201–217.

Civán P, Svec M, Hauptvogel P. **2011**. On the coevolution of transposable elements and plant genomes. *Journal of Botany* **2011**: 1–9.

Dahlgren RMT, Clifford HAT, Yeo PF. **1985**. *The families of the monocotyledons*. Springer, Vienna, Austria.

Dale W, Kiger E. **2002**. *Polygonatum*. *Flora of North America* **26**: 210–212. Oxford University Press.

Davey JW, Blaxter ML. **2011**. RADseq: next-generation population genetics. *Briefings in Functional Genomics* **9**: 416–423.

Deng XY, Wang Q, He XJ. **2009**. Karyotype of 16 populations of eight species in the genus *Polygonatum* (Asparagaceae) from China. *Botanical Journal of the Linnean Society* **169**: 245–254.

Dodsworth S, Jang TS, Struebig M, Chase MW, Weiss-Schneeweiss H, Leitch AR. **2016**. Genome-wide repeat dynamics reflect phylogenetic distance in closely related allotetraploid *Nicotiana* (Solanaceae). *Plant Systematics and Evolution* **303**: 1013–1020.

Ellegren H. **2004**. Microsatellites: simple sequence with complex evolution. *Nature Reviews Genetics* **5**: 435–445.

Emiliani G, Paffetti D, Giannini R. **2009**. Identification and molecular characterization of LTR and LINE retrotransposable elements in *Fagus sylvatica* L.. *iForest – Biogeosciences and Forestry* **2**: 119–126.

Engler A. **1888**. Liliaceae. *Die natürlichen Pflanzenfamilien* **2**: 10–91.

Fischer MA, Oswald K, Adler W. **2008**. *Exkursionsflora für Österreich, Liechtenstein und Südtirol.* Land Oberösterreich, oberösterreichische Landesmuseen, Linz, Austria, pp. 1066–1067.

Floden A, Schilling EE. **2018**. Using phylogenomics to reconstruct phylogenetic relationships within tribe Polygonateae (Asparagaceae), with a special focus on *Polygonatum*. *Molecular Phylogenetics and Evolution* **129**: 202–213.

Galbraith DW, Harkins KR, Maddox JR, Ayres NM, Sharma D, Firoozabady E. **1983**. Rapid flow cytometric analysis of the cell cycle in intact plant tissue. *Science* **220**: 1049–1051.

Garrido-Ramos MA. **2017**. Satellite DNA: an evolving topic. *Genes* **8**: 230.

Greilhuber J, Ebert I. **1994**. Genome size variation in *Pisum sativum*. *Genome* **37**: 4.

Heslop-Harrison JS, Schmidt T. **1998**. Genomes, genes and junk: The large-scale organization of plant genomes. *Trends in Plant Science* **3**: 195–199.

Hirai H. **2020**. Chromosome dynamics regulating genomic dispersion and alteration of nucleolus organizer regions (NORs). *Cells* **9**: 971.

Hou Z, Li A. **2020**. Population genomics reveals demographic history and genomic differentiation of *Populus davidiana* and *Populus tremula*. *Frontiers in Plant Science* **11**: 1103.

Jang TS, Weiss-Schneeweiss H. **2015**. Formamide-free genomic *in situ* hybridization allows unambiguous discrimination of highly similar parental genomes in diploid hybrids and allopolyploids. *Cytogenetic Genome Research* **146**: 325–331.

Janchen E. **1956–1960**. *Catalogus florae austriae*. Springer, Vienna, Austria.

Jiang J. **2019**. Fluorescence *in situ* hybridization in plants: recent developments and future applications. *Chromosome Research* **27**: 153–165.

Kapitonov VV, Jurka J. **2001**. Rolling-circle transposons in eukaryotes. *PNAS* **98**: 8714–8719.

Kejnovský E, Hawkins J, Feschotte C. **2012**. Plant transposable elements: biology and evolution. In: Wendel J, Greilhuber J, Dolezel J, Leitch IJ (eds.). *Plant genome diversity, Vol. 1, Plant genomes, their residents, and their evolutionary dynamics*. Springer, Vienna, Austria, pp.17–34.

Kim NS. **2017**. The genomes and transposable elements in plants: are they friends or foes? *Genes and Genomics* **39**: 359–370.

Kit S. **1961**. Equilibrium sedimentation in density gradients of DNA preparations from animal tissues. *Journal of Molecular Biology* **3**: 711–716.

Kumar A, Bennetzen JL. **1999**. Plant retrotransposons. *Annual Review Genetics* **33**: 479–532.

Macas J, Neumann P. **2007**. Ogre elements – a distinct group of plant *Ty3/gypsy*-like retrotransposons. *Gene* **390**: 108–116.

Macas J, Novák P, Pellicer J, Čížková J, Koblížková A, Neumann P, Fuková I, Doležel J, Kelly LJ, Leitch IJ. **2015**. In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe *Fabeae*. *PLoS ONE* **10**: e0143424.

Meng Y, Nie ZL, Deng T, Wen J, Yang YP. **2014**. Phylogenetics and evolution of phyllotaxy in the Solomon's seal genus *Polygonatum* (Asparagaceae: Polygonateae). *Botanical Journal of the Linnean Society* **176**: 435–451.

Nishikawa K, Furuta Y, Ishitobi K. **1984**. Chromosome evolution in the genus *Carex* as viewed from nuclear DNA content, with special reference to its aneuploidy. *Japanese Journal of Genetics* **59**: 465–472.

Novák P, Hřibova E, Neumann P, Koblížková A, Doležel J, Macas J. **2014**. Genome-wide analysis of repeat diversity across the family Musaceae. *PLoS ONE* **9**: e98918.

Novák P, Neumann P, Macas J. **2020**. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nature Protocols* **15**: 3745–3776.

Otto FJ, Oldiges H, Göhde W, Jain VK. **1981**. Flowcytometric measurement of nuclear DNA content variations as a potential *in vivo* mutagenicity test. *Cytometry* **2**: 189–191.

Palazzo AF, Gregory TR. **2014**. The case for junk DNA. *PLoS Genetics* **10**: e1004351.

Piégu B, Bire S, Arensburger P, Bigot Y. **2015**. A survey of transposable element classification systems – a call for a fundamental update to meet the challenge of their diversity and complexity. *Molecular Phylogenetics and Evolution* **86**: 90–109.

Pellicer J, Kelly LJ, Leitch IJ, Zomlefer WB, Fay MF. **2014**. A universe of dwarf and giants: genome size and chromosome evolution in the monocot family Melanthiaceae. *New Phytologist* **201**: 1484–1497.

Plohl M, Luchetti A, Mestrović N, Mantovani B. **2008**. Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. *Gene* **409**: 72–82.

Rudall PJ, Conran JG, Chase MW. **2000**. Systematics of Ruscaceae/Convallariaceae: a combined molecular and morphological investigation. Bot*anical Journal of the Linnean Society* **134**: 73–92.

Schulmann AH, Wicker T. **2013**. In: Federoff N. V. (eds). *Plant transposons and genome dynamics in evolution*. John Wiley and Sons, Hoboken, USA, pp.15–40.

Schwarzacher T, Heslop-Harrison P. **2000**. *Practical* in situ *hybridization*. BIOS Scientific Publishers Ltd., Oxford, United Kingdom.

Shao JZ, Zhang DC, Yang JG, Yao CH, Qian F. **1993**. Karyotype analysis of 5 species of *Polygonatum* Mill. *Acta Phytotaxonomica Sinica* **31**: 353–361.

Simon AU. **2020**. *Evolution of selected satellite DNAs in cultivated* Capsicum *species.* Master's thesis, University of Vienna, Vienna Austria.

Sonnhammer ELL, Durbin R. **1995**. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**: GC1–GC10.

Strickler SR, Bombarely A, Mueller LA. **2012**. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *American Journal of Botany* **99**: 257–266.

Suomalainen E. **1947**. On the cytology of the genus *Polygonatum*. Group Alternifolia. *Annales Academiæ Scientiarum Fennicæ IV* **13**: 1–63.

Tamura MN, Schwarzbach AE, Kruse S, Reski R. **1997**. Biosystematic study on the genus *Polygonatum* (Convallariaceae) IV. Molecular phylogenetic analysis based on restriction site mapping of the chloroplast gene *trn*K. *Feddes Repertorium* **108**: 159–168.

Tamura K, Stecher G, Kumar S. **2021**. MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**: 3022–3027.

Temsch EM, Greilhuber J, Krisai R. **2010**. Genome size in liverworts. *Preslia* **82**: 63–80.

Therman E. **1950**. Chromosome numbers in American *Polygonatum* species. *American Journal of Botany* **37**: 407–413.

Therman E. **1953**. Chromosomal evolution in the genus *Polygonatum*. *Hereditas* **39**: 277–288.

Thomas J, Pritham E. **2015**. Helitrons, the eukaryotic rolling-circle transposable elements. *Microbiology Spectrum* **3(4)**.

Tremetsberger K, Weiss-Schneeweiss H, Stuessy T, Samuel R, Kadlec G, Ortiz MA, Talavera S. **2005**. Nuclear ribosomal DNA and karyotypes indicate a NW African origin of South American *Hypochaeris* (Astereraceae, Cichorieae). *Molecular Phylogenetics and Evolution* **35**: 102–116.

Vestergren T. **1925**. *Polygonatum multiflorum* (L.) ALL. x *officinale* ALL. i Sverige. *Svensk Botanisk Tidskrift* **19**: 495–519.

Vitte C, Estep MC, Leebens-Mack J, Bennetzen JL. **2013**. Young, intact and nested retrotransposons are abundant in the onion and asparagus genomes. *Annals of Botany* **112**: 881–889.

Wang JW, Li MX, Li LX. **1987**. Studies on the cytotaxonomy of *Polygonatum* I. Karyotypes and evolution of eight species of *Polygonatum* in China. *Wuhan Botanical Research* **5**: 1–10.

Wang JJ, Yang YP, Sun H, Wen J, Deng T, Nie ZL, Meng Y. **2016**. The biogeographic south-north divide of *Polygonatum* (Asparagaceae Tribe Polygonateae) within eastern Asia and its recent dispersals in the Northern Hemisphere. *PLoS ONE* **1**: e0166134.

Weis BL, Kovacevic J, Missbach S, Schleiff E. **2015**. Plant-specific features of ribosome biogenesis. *Trends in Plant Science* **20**: 729–740.

Weiss-Schneeweiss H, Jang CG. **2003**. Cytotaxonomic analysis of Austrian *Polygonatum* species (Ruscaceae). *Neilreichia* **2–3**: 245–250.

Weiss-Schneeweiss H, Tremetsberger K, Schneeweiss GM, Parker JS, Stuessy TF. **2008**. Karyotype diversification and evolution in diploid and polyploid South American *Hypochaeris* (Asteraceae) inferred from rDNA localization and genetic fingerprint data. *Annals of Botany* **101**: 909–918.

Weiss-Schneeweiss H, Villaseñor JL, Stuessy TF. **2009**. Chromosome numbers, karyotypes, and evolution in *Melampodium* (Asteraceae). *International Journal of Plant Science* **170**: 1168–1182.

Weiss-Schneeweiss H, Schneeweiss GM. **2013**. Karyotype diversity and evolutionary trends in angiosperms. In: Leitch IJ, Greilhuber J, Dolezel J, Wendel JF (eds.). *Plant genome diversity. Vol. 2. Physical structure, behavior and evolution of plant genomes*. Springer, Vienna, Austria, pp. 209–230.

Weiss-Schneeweiss H, Leitch AR, McCann J, Jang TS, Macas J. **2015**. Employing next-generation sequencing to explore the repeat landscape of the plant genome. In: Hörandl E, Appelhans MS (eds.). *Next-generation sequencing in plant systematics*. International Association for Plant Taxonomy, Bratislava, Slovakia, pp. 155–179.

White TJ, Bruns T, Lee S, Taylor J. **1990**. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: EDITORS? *PCR Protocols*. Academic Press, Cambridge, United Kingdom: , pp. 315–322.

Zhao LH, Zhou SD, He XJ, Wang ZX, Peng L. **2014a**. A cytotaxonomic analysis of Chinese *Polygonatum* (Asparagaceae) species. *Nordic Journal of Botany* **32**: 441–451.

Zhao P, Zhao C, Li X, Gao Q, Huang L, Xiao P, Gao W, **2014b**. The genus *Polygonatum*: A review of ethnopharmacology, phytochemistry and pharmacology. *Journal of Ethnopharmacology* **214**: 274–291.

Zimmer EA, Wen J. **2015**. Using nuclear gene data for plant phylogenetics: progress and prospects II. Next-gen approaches. *Journal of Systematics and Evolution* **53**: 371–379.

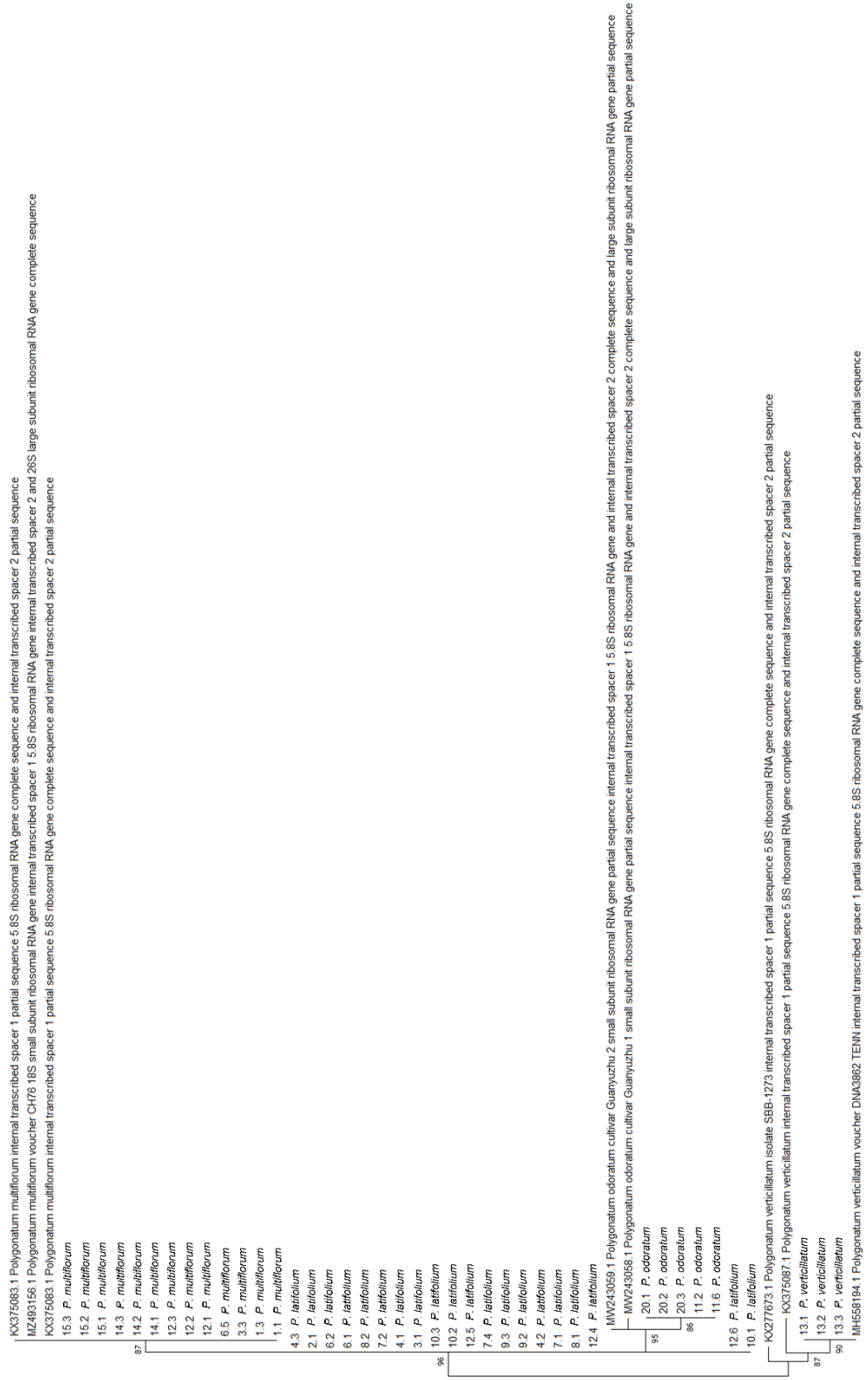
Abstract

Polygonatum is a genus within the family of Asparagaceae with its diversity hotspot in South-Western China and the Indo-Burma region. Four species of this genus are found in Europe, *P. latifolium, P. multiflorum, P. odoratum* and *P. verticillatum*. In this study, all four species were sampled from multiple populations in Austria and analyzed chromosomally and genomically using Feulgen staining, fluorescence *in situ* hybridization, flow cytometry, phylogenetic analyses of the ITS1 and ITS2 sequences and analyses of repeat profiles using next generation sequencing data and the dedicated RepeatExplorer pipeline. The study provides novel data on basic karyotype structure and genome size as well as the localization of rDNA loci, and the repeat profiles of the analyzed taxa. The analyses of NGS data provide first insight into the composition of repetitive DNA fraction in the genus and reveal that Athila and Tekay lineages of Ty3*/gypsy* retrotransposons are the most abundant and active repeats in all four species. The comparative repeat analyses confirmed the distinctiveness of section *Polygonatum* represented here by three species and section *Verticillata* represented by *P. verticillatum,* inferred also from morphological and cytogenetic data. Molecular synapomorphies were observed and described. Two novel satellite DNAs have been identified and will provide excellent tools for more in-depth genome and karyotype analyses in the genus.

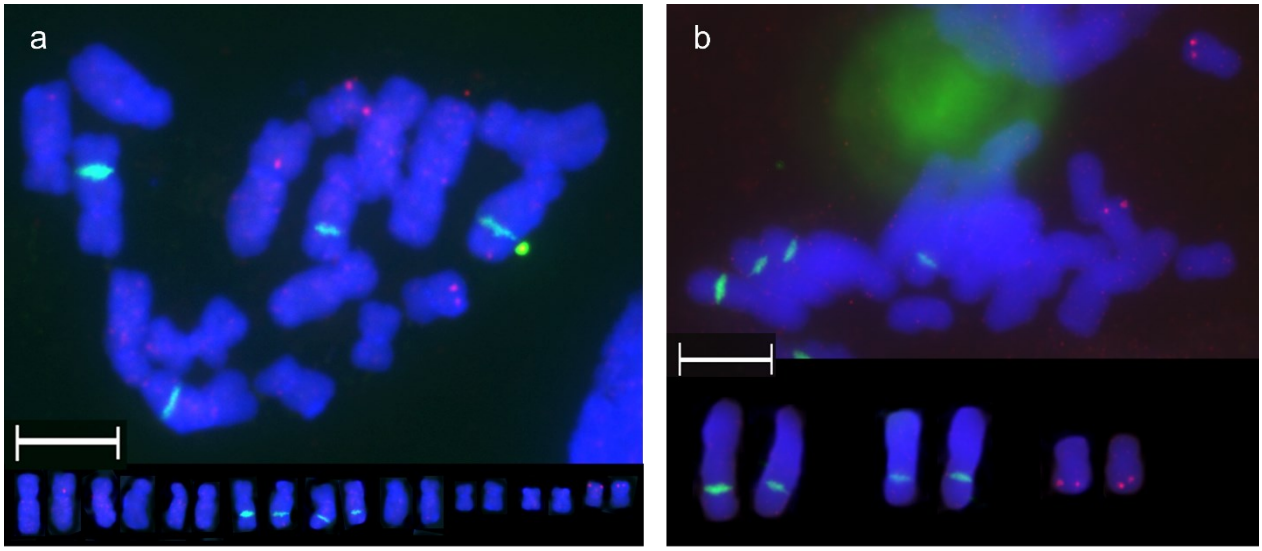
Zusammenfassung

Polygonatum ist eine Gattung in der Familie Asparagaceae innerhalb der Monokotyledonen. Der Biodiversitätshotspot dieser Gattung liegt in Süd-West China und der Indo-Burma-Region; in Europa wie auch in Österreich sind vier Arten zu finden, *P. latifolium*, *P. multiflorum*, *P. odoratum* und *P. verticillatum*. Diese Masterarbeit behandelt die Morphologie der Chromosomen, die Lage der rDNA, Polymorphismen des Karyotypes, sowie die repetitiven DNA-Sequenzen. Hierfür wurden verschiedenste Methoden verwendet, wie Feulgen staining, fluorescence *in situ* hybridization, flow cytometry, ITS sequencing und next generation sequencing. Die Anzahl der Chromosomen konnte für jede Art eindeutig festgestellt werden und ist kongruent mit vorangegangen Studien, ebenfalls die Anzahl der rDNA Regionen. Die NGS Daten zeigten, dass die Elemente Ty3*/gypsy* Athila und Ty3*/gypsy* Tekay am häufigsten in allen vier Arten vorkommen. Die vergleichenden Analysen des Repeatomes aller vier Arten konnten außerdem molekulare Synapomorphien aufdecken, die die Arten der zwei Sektionen (sect. *Polygonatum* und sect. *Verticillata*) klar unterscheiden. Zusätzlich wurden größere Ähnlichkeiten im Karyotyp und dem Repeat-Profil der beiden näher verwandten Arten, *P. latifolium* und *P. multiflorum*, festgestellt. Zu guter Letzt konnten zwei bislang unbekannte satDNAs in ihrer Länge und Sequenz beschrieben werden.

Appendix



Appendix 1: Phylogenetic tree of the European/Austrian *Polygonatum* species using the ITS1 and ITS2 region, including GenBank sequences. Numbers indicate the species ID (see *Table 1*).



Appendix 2: FISH plates and cut chromosomes for (a) *P. multiflorum* (ID 12.1), (b) *P. odoratum* (ID 20.1).

Appendix 3: Number of overlapping and non-overlapping reads for all four *Polygonatum* species. The majority of reads should not overlap.

Species	**Overlapping**	**Non-overlapping**
P. latifolium	4.493.918	16.748.482
P. multiflorum	3.824.284	18.093.226
P. odoratum	4.673.194	15.619.744
P. verticillatum	3.448.248	15.010.038