Repetitive DNA Dynamics and Polyploidization in the Genus *Nicotiana* (Solanaceae)

7

Steven Dodsworth, Ales Kovarik,
Marie-Angèle Grandbastien,
Ilia J. Leitch, and Andrew R. Leitch

Abstract

Large variations in genome size are observed in angiosperms as a result of whole-genome duplications and the balance between amplification and deletion of repetitive DNA, together explaining the observed variation in plant genome size. In the genus *Nicotiana*, there are 42 cytogenetically diploid species that have been classified into eight sections. There are also six allopolyploid *Nicotiana* sections that have evolved from species in different diploid sections. The phylogenetic relationships among these *Nicotiana* species, along with recurrent polyploidization events, permits the divergence of repetitive content in both diploid and allopolyploid genomes to be compared through evolutionary time. In this chapter, we review genome size variation in *Nicotiana* that reveals both genome upsizing and genome downsizing in different polyploid species. We discuss the divergence of specific repetitive elements, including tandemly repeated satellite DNAs, retroelements, and intergenic spacers as well as the sub-repeats contained in 35S rDNA. The lag-phase hypothesis, which describes post-polyploid radiations, is posed as a potentially important mechanism of evolution in *Nicotiana* section *Suaveolentes*, the largest polyploid section that consists of over half the current species diversity.

S. Dodsworth
School of Life Sciences, University of Bedfordshire, Luton LU1 3JU, UK

A. Kovarik
Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno 61265, Czech Republic

M.-A. Grandbastien
Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, University Paris-Saclay, 78000 Versailles, France

I. J. Leitch
Department of Comparative Plant and Fungal Biology, Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3DS, UK

A. R. Leitch (✉)
School of Biological and Chemical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
e-mail: a.r.leitch@qmul.ac.uk

7.1 Introduction to Plant Repetitive DNA

There is, approximately, a 2,400-fold range in genome size (GS; measured as the amount of DNA in giga base pairs [Gbp] in the haploid unreplicated 1C nucleus) across angiosperms, ranging from ~0.061 Gbp/1C in *Genlisea aurea* to ~148 Gbp/1C in *Paris japonica* (Pellicer et al. 2010), a larger range than that observed for any comparable group of eukaryotes (Leitch and Leitch 2012; Pellicer et al. 2018). This variation in GS arises from two sources: polyploidy

(whole-genome duplication) and the balance between processes that amplify repetitive DNA (i.e., (retro)transposition and recombination-based amplification) and those that delete repeats and other types of non-essential DNA (i.e., recombination-based deletion) (Kejnovsky et al. 2012).

Despite the large range in GS found across angiosperms, most species have a small GS (modal GS 0.587 Gbp/1C, mean GS 5.020 Gbp/1C, n = 10,768 species, Pellicer et al. 2018). This is interesting given the context of multiple ancestral polyploid events in most, if not all, angiosperm lineages. For example, within seed plants, the lineage leading to Solanaceae is predicted to have been through multiple rounds of whole-genome duplication, such that all species are at least paleo-36-ploid (Wendel 2015). Such high ploidal levels are not generally reflected in large genomes, most likely as a consequence of genome downsizing that typically follows polyploidy and involves the deletion of both genes and repeats during the process of post-polyploidization diploidization (Leitch and Bennett 2004; Dodsworth et al. 2016). Another general feature reported for repeats in plants with a small to medium GS (<10 Gbp/1C) is that many are highly dynamic, with amplification and deletion processes occurring so rapidly such that the half-life of any particular repeat is reported to be only 3–4 million years in Poaceae, Brassicaceae, Fabaceae, and Vitaceae (Ma et al. 2004; El Baidouri and Panaud 2013).

The proportion of a plant genome that is repetitive is highly variable between taxa. This partly reflects differences in the balance between amplification and deletion processes, which in turn influence the occurrence and longevity of (micro)satellite repeats, (retro)transposable elements, and their truncated derivatives. For example, conifers typically have larger GS than most lineages of angiosperms, probably reflecting contrasting genome dynamics between gymnosperms and angiosperms (Leitch and Leitch 2012). There are also differences in repetitive content of the genome depending on GS, such that species with the smallest GS have the smallest genome proportion occupied by repeats. For example, the estimated genome proportion in *Utricularia gibba* with a GS of 0.088 Gbp/1C is just 3% (Ibarra-Laclette et al. 2013) versus a reported 91.6% genome proportion in *Aegilops tauschii*, with a GS of 4.3 Gbp/1C (Li et al. 2004).

7.2 Introduction to *Nicotiana*

There are 42 species of *Nicotiana* that are cytogenetically diploid, falling into eight sections: *Alatae* with eight species (haploid chromosome count, n = 9–10), *Noctiflorae* with six species (n = 12), *Petunioides* with eight species (n = 12), *Paniculatae* with seven species (n = 12), *Sylvestres* with a single species (n = 12), *Tomentosae* with five species (n = 12), *Trigonophyllae* with two species (n = 12), and *Undulatae* with five species (n = 12). Homoploid hybrids have been reported to form between some of the sections, particularly in species belonging to *Petunioides* and *Noctiflorae* (Knapp et al. 2004; Kelly et al. 2010).

In addition, there are six allopolyploid *Nicotiana* sections, thought to involve species from different diploid sections, and providing an opportunity to study and compare repeat divergence in diploid lineages and the genomic melting pot generated by allopolyploidy itself. All the allopolyploids retain a signature of polyploidy in chromosome counts, and evidence of interspecific hybridization is seen in phylogenetic patterns from multiple nuclear markers, including the internally transcribed spacer sequences of ribosomal DNA (rDNA) (Clarkson et al. 2010, 2017), paralogs of the nuclear genes *ADH*, *GS*, *FLO/LFY*, *WAXY*, and *MADS1/FUL* (Kelly et al. 2010), and comparisons between these and plastid sequences from the maternal genome donor (Clarkson et al. 2004, 2010). Figure 7.1 summarizes the phylogenetic relationships in genus *Nicotiana*.

(1) The oldest allopolyploid section, *Nicotiana* section *Suaveolentes*, contains at least 40 species (Knapp et al. 2004). This section has radiated substantially in Australia, and there are likely to be many more taxa than are currently
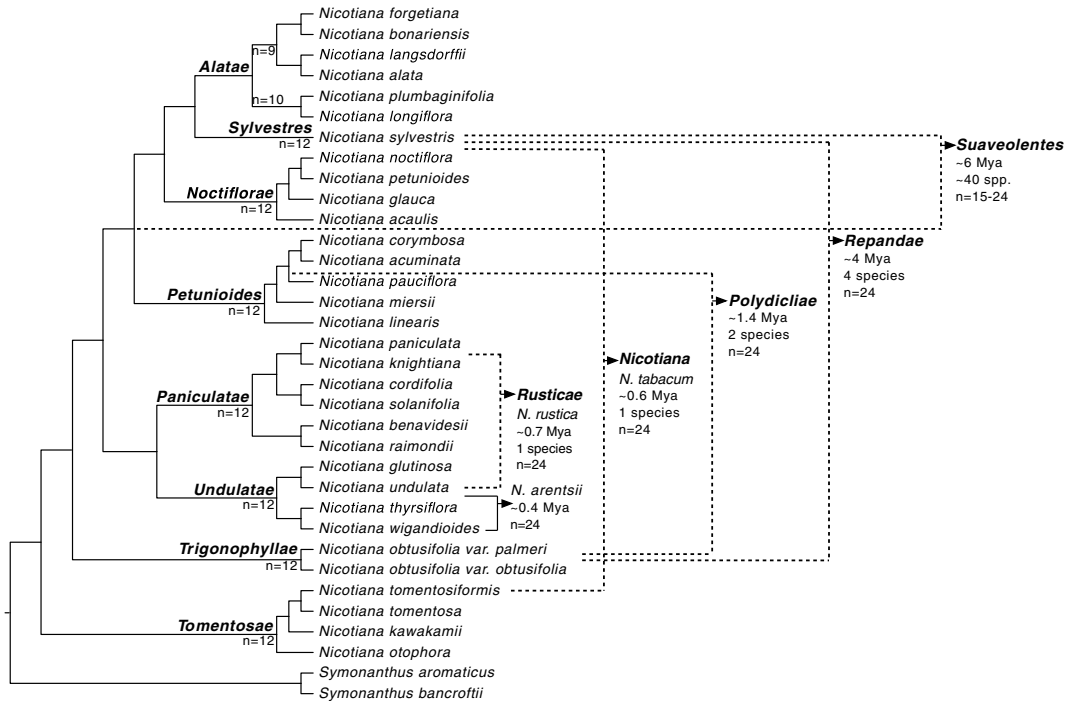
Fig. 7.1 Summary of phylogenetic relationships in the genus *Nicotiana*, including the dates of formation (Clarkson et al. 2017) and number of species in each of the allopolyploid sections (Knapp et al. 2004). The diploid phylogenetic tree was reconstructed from combined plastid and nrITS data and is summarized from Clarkson et al. (2004) and Kelly et al. (2010). Haploid chromosome numbers (*n*) are given for each section. Section names in bold have been given. There are three polyploid sections containing multiple species (*Nicotiana* sections *Suaveolentes*, *Repandae* and *Polydicliae* and two monotypic polyploid sections, *Nicotiana* section *Nicotiana* (*N. tabacum*) and *Rusticae* (*N. rustica*)). The polyploid species *N. arentsii* is in *Nicotiana* section *Undulatae*, which also includes diploid species

recognized (four new species were described recently by Chase et al. 2018). The section includes *N. benthamiana*, a species widely used in studies of plant–pathogen interactions, and whose genome has been fully sequenced (Bombarely et al. 2012). Time-calibrated phylogenetic reconstructions suggest that the section probably formed about 6 million years ago (Mya). The diploid progenitors of section *Suaveolentes* are difficult to determine, likely because of their formation from one or more diploid hybrid taxa. Phylogenetic evidence suggests the involvement of several diploid sections (*Noctiflorae*, *Alatae*, and *Petunioides*) as the maternal genome donor, with *N. sylvestris* (section *Sylvestres*) as the most closely related extant relative of the paternal genome donor (Clarkson et al. 2010, 2017; Schiavinato et al. 2020). Chromosome numbers

vary from the expected n = 24 (the sum of the parental chromosome counts, for an allotetraploid) down to n = 15, depending on the species. The species with lower chromosome numbers are presumed to have arisen from postpolyploidization dysploidy involving chromosome rearrangements, as part of a broader diploidization process.

(2) *Nicotiana* section *Repandae* has four allotetraploid species, each with n = 24 chromosomes, with two species that are endemic to the Revillagigedo Islands. The section is thought to have formed about 4 Mya (Clarkson et al. 2017) from a maternal parent most closely related to extant *N. sylvestris* and a paternal species related to *Nicotiana* section *Trigonophyllae* (e.g., *N. obtusifolia* var. *obtusifolia*/*N. obtusifolia* var. *palmeri*).

(3) *Nicotiana* section *Polydicliae* has two species, each with n = 24 chromosomes, and one of which, *N. quadrivalvis*, was cultivated by native North Americans, presumably for recreational use. The section formed about 1.4 Mya (Clarkson et al. 2017) from progenitors of *N. obtusifolia* (section *Trigonophyllae*; maternal genome donor) and a member of section *Petunioides* as the paternal genome donor.

(4) *Nicotiana rustica* (section *Rusticae*, n = 24) formed about 0.7 Mya (Clarkson et al. 2017) from progenitors of *N. undulata* (section *Undulatae*; paternal) and section *Paniculatae* (either *N. paniculata* or *N. knightiana*) as the maternal progenitor.

(5) *Nicotiana tabacum* (section *Nicotiana,* n = 24), the best-known *Nicotiana* allopolyploid, formed about 0.6 Mya (Clarkson et al. 2017) from progenitors of *N. sylvestris* (section *Sylvestris*; maternal genome donor) and *N. tomentosiformis* (section *Tomentosae*; paternal).

(6) *Nicotiana arentsii* (section *Undulatae*, n = 24) formed about 0.4 Mya (Clarkson et al. 2017) from two other species in section *Undulatae*, most closely related to extant *N. undulata* (section *Undulatae*; maternal parent) and *N. wigandioides* (section *Undulatae*; paternal parent).

7.3 Genome Size Variation in *Nicotiana*

An analysis of angiosperm GSs, grouped into presumed ploidal levels based on chromosome counts (from diploid to octoploid), indicated that higher ploidal levels were not associated with larger GS, as might be predicted (Leitch and Bennett 2004). These findings suggested that following polyploidy, there is a tendency toward the loss of genomic DNA, termed genome downsizing, perhaps caused by selection against large GS in angiosperms (Leitch and Leitch 2012; Guignard et al. 2016). Genome downsizing is predicted as a likely occurrence given enough time post-polyploidization. Serial genome rearrangements associated with genome downsizing lead to the so-called "wondrous cycles of polyploidy" reported for many angiosperm lineages (Wendel 2015).

To predict the direction of GS change following ancient polyploidy, it is necessary to reconstruct the size of ancestral diploid genomes at the point of polyploid formation. Such ancestral GSs can be estimated by summing the GSs of the most closely related extant diploids (Leitch et al. 2008). Alternatively, they can be derived by reconstructing the ancestral diploid GSs of diploid lineages at the point of allopolyploidy. This is achieved using ancestral reconstruction approaches, which requires a confident phylogenetic assignment of parental lineages. With estimates of the ages of polyploid events, rates of genome size change can be estimated. The predicted GS of ancestral polyploids can then be compared with the actual GS of the extant polyploid taxa. When such an analysis is conducted on *Nicotiana*, some polyploid species are predicted to have experienced genome upsizing, while most experienced genome downsizing.

Given the GS of the progenitor diploid taxa, the allotetraploids *N. tabacum* and *N. arentsii* are thought to have undergone genome downsizing of ~3.7% and ~3.9%, respectively, over ~0.4–0.6 million years. Similarly, *N. rustica* is likely to have downsized by ~1.9–5.4% over a similar time frame. In contrast, in section *Polydicliae*, divergence over a longer time frame of ~1.4 million years is thought to have given rise to approximately 2.5% (*N. clevelandii*) and 7.5% (*N. quadrivalvis*) increases in GS. Both genome upsizing and downsizing have been reported in species in section *Repandae*, over ~4 million years, with genome upsizing in *N. repanda* (~26.6%), *N. nesophila* (~19.1%), and *N. stocktonii* (~19.1%), and downsizing in *N. nudicaulis* (~14.3%) (Leitch et al. 2008). In the oldest section, *Suaveolentes*, which originated ~6 Mya, almost all species have undergone substantial genome downsizing associated with their divergence (Fig. 7.2). This has generated genomes that are similar in size to the mean GS of extant *Nicotiana* diploids (i.e., ~3.2 Gbp/1C, 34 species) and not much larger than *N. sylvestris* (2.6 Gbp/1C, paternal parent) and the average GS of species in sections *Noctiflorae*,
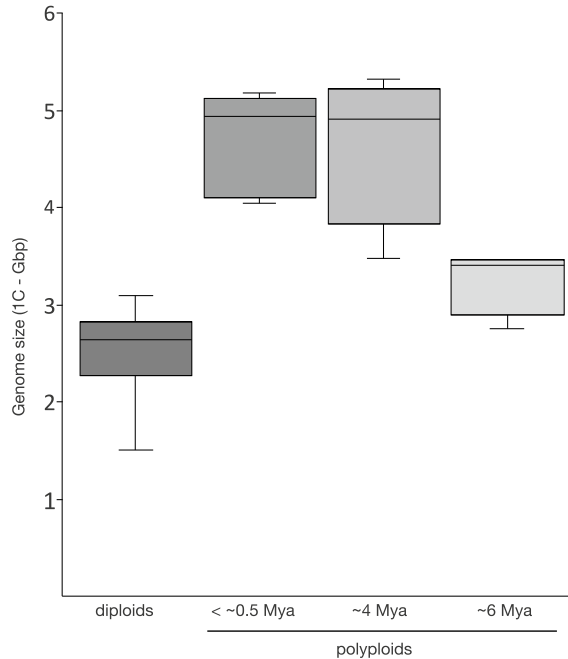
Fig. 7.2 Box plots showing the distribution of genome sizes (1C-values) across diploid and polyploid *Nicotiana* taxa. Young polyploids show near-additive and/or genome upsizing, whereas older polyploid groups tend to show genome downsizing. Analysis based on data from Leitch et al. (2008) for diploid (nine taxa), young polyploid taxa (*N. tabacum*, *N. rustica*, *N. arentsii*, *N. clevelandii*, and *N. quadrivalvis*), and the 4-Mya section *Repandae* (all four taxa), and unpublished data for the 6-Mya section *Suaveolentes* (five taxa from the core Australian group: *N. simulans*, *N. velutina*, *N. maritima*, *N. truncata*, and *N. goodspeedii*)

Alatae, and *Petunioides* (about 3.4 Gbp/1C, 14 species) of which an extinct relative or a homoploid hybrid is likely to have been the maternal parent. Such genome downsizing likely arises as part of the diploidization process that includes chromosome rearrangements and dysploidy (reduction from n = 24 to n = 15 chromosomes in some taxa).

7.4 Repeat Divergence in Diploid *Nicotiana*

In concert with GS variation, diploid *Nicotiana* genomes contain variable amounts of repetitive elements. The types, abundance, and patterns of element accumulation within the genomes of *Nicotiana* are, however, fairly consistent across the different diploid taxa thus far investigated. *Nicotiana* genomes are dominated by retrotransposons, particularly Ty3/Gypsy and Ty1/Copia

elements. The most abundant retroelements in all *Nicotiana* genomes studied are Ty3/Gypsy elements, most of which are chromoviruses, particularly Tekay elements (Fig. 7.3). Most of the retrotransposon families found in *Nicotiana* are also present in genus *Symonanthus*, which is thought to be the closest relative to *Nicotiana*, based on phylogenetic evidence available to date (Clarkson et al. 2004, 2010; Kelly et al. 2010; Särkinen et al. 2013). Nevertheless, some elements are unique to *Symonanthus* (e.g., Ty1/Copia elements of the Ikaros family) and do not appear to be present in *Nicotiana* genomes (Fig. 7.3).

Xu et al. (2017) examined repeats in four diploid *Nicotiana* species, as well as in potato (*Solanum tuberosum*) and tomato (*Solanum lycopersicum*), and suggested there was an expansion of Ty3/Gypsy retroelements in *Nicotiana* that correlated with its increased GS. In Fig. 7.3, we present further evidence that GS is
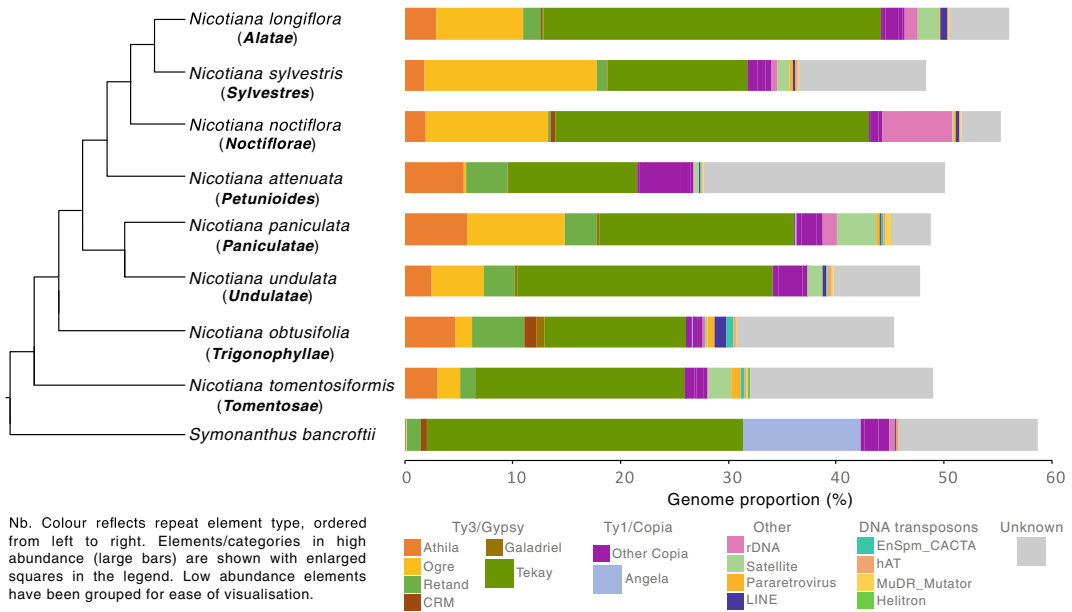
Nb. Colour reflects repeat element type, ordered from left to right. Elements/categories in high abundance (large bars) are shown with enlarged squares in the legend. Low abundance elements have been grouped for ease of visualisation.

Fig. 7.3 Comparison of the repetitive DNA content in diploid *Nicotiana* species (with *Symonanthus* as the outgroup) as analyzed using high-throughput sequence data (Illumina) and RepeatExplorer2 de novo clustering (Novak et al. 2010; Novák et al. 2013). The abundance of repetitive elements is shown as a proportion of the genome and annotations are shown for major repeat types using REXdb (Neumann et al. 2019) as implemented in RepeatExplorer2. Abundant categories are shown with enlarged squares in the legend. Phylogenetic relationships are summarized and section names are shown in bold

correlated with the abundance of retroelements, particularly Ty3/Gypsy retrotransposons, across a broader set of eight *Nicotiana* diploid species (including representatives from all diploid sections). Xu et al. (2017) also suggested there was an expansion of a Solanaceae-specific subgroup of MITE elements (called DTT-NIC1) in *N. attenuata* that were enriched within a 1-Kbp region upstream of the genes involved in nicotine biosynthesis. They proposed that these elements may have been involved in the up-regulation of genes involved in nicotine biosynthesis. Other class II DNA transposons that are present in all *Nicotiana* genomes analysed to date include hAT and MuDR_Mutator elements (Fig. 7.3), while EnSpm/CACTA elements were found in the genomes of about half of the species analysed.

Using molecular and cytogenetic approaches, much research has been conducted on tandemly repeated satellite DNAs of *Nicotiana*, including well-studied families such as GRS (Gazdová et al. 1995), NTRS (Matyasek et al. 1997),

geminivirus-related DNA (GRD) (Bejarano et al. 1996), and HRS60 (e.g., Fajkus et al. 1995b; Gazdová et al. 1995) (Table 7.1). These repetitive elements were isolated predominantly from telomeric and subtelomeric domains (Fajkus et al. 1995a) and centromeric and pericentromeric domains (Shibata et al. 2013), and their overall structure and typically high levels of cytosine methylation studied (Kovarik et al. 2000). In addition, there has been considerable focus on sequence divergence, activity, cytosine methylation, and distribution of 5S (Fulneček et al. 2002; Matyasek et al. 2002) and 35S rDNA loci (Kovarik et al. 2004).

The structure of the intergenic spacer and the sub-repeats contained in 35S rDNA have been further studied because they are thought to be involved in driving homogenization of rDNA unit arrays and the regulation of rDNA expression (Borisjuk et al. 1997; Volkov et al. 1999; Lim et al. 2004b; Kovarik et al. 2008). One 135-bp sub-repeat termed A1/A2, which was isolated

Table 7.1 Repeats that characterize the chromosome domains of *Nicotiana tabacum* (common tobacco)

Repeat family	Feature	Genome proportion (%)	Chromosome localization	Subgenome specificity		Reference
				S	T	
GRD3	Endogenous geminivirus	<0.01	Interstitial	–	+	(Bejarano et al. 1996; Murad et al. 2002)
GRD5	Endogenous geminivirus	<0.01	Interstitial, subtelomeric	+	+	(Bejarano et al. 1996; Murad et al. 2002)
NsEPRV	Endogenous pararetrovirus	n.d.	Dispersed	+	–	(Matzke et al. 2004)
NtoEPRV	Endogenous pararetrovirus	n.d.	Dispersed	–	+	(Gregor et al. 2004)
Tnt1	Retroelement Ty1/Copia	<0.01	Dispersed	+	+	(Melayah et al. 2004)
NTS9	Satellite	0.075	Pericentromeric, interstitial	+	–	(Moscone et al. 1996)
HRS60	Satellite	2	Subtelomeric	+	–	(Koukalova et al. 1990, 2010; Kenton et al. 1993)
GRS	Satellite	0.15	Pericentromeric	–	+	(Lim et al. 2000b)
NTRS	Satellite	0.75	Interstitial	–	+	(Matyasek et al. 1997)
A1/A2	Satellite, related to the intergenic spacer of rDNA	0.022	Interstitial, subtelomeric	–	+	(Lim et al. 2004b)
NicCL3	Satellite	0.1	Subtelomeric	–	+	(Renny-Byfield et al. 2012)
TAS49	Dispersed	0.15	Interstitial, subtelomeric,	±	+	(Horakova and Fajkus 2000)
R8.1	Retroelement	0.3	Dispersed	–	+	(Kuhrova et al. 1991)
NtoCR	Retroelement, TY3/Gypsy	n.d.	Centromeric, pericentromeric	–	+	(Nagaki et al. 2011)
NtCR	Retroelement, TY3/Gypsy	n.d.	Centromeric, pericentromeric	+	+	(Nagaki et al. 2011)
RETS	Tandem	n.d.	Interstitial, subtelomeric	+	–	(Shibata et al. 2013)
NTA11	Tandem	n.d.	Interstitial	–	+	(Shibata et al. 2013)
HT3E06	Tandem	n.d.	Centromeric	+	–	(Shibata et al. 2013)
HT1H04	Tandem	n.d.	Centromeric, pericentromeric	+	–	(Shibata et al. 2013)
HT3G02	Tandem	n.d.	Centromeric, pericentromeric	–	+	(Shibata et al. 2013)
Telomeric repeats	TTTAGGG	0.053	Mostly telomeric, a few interstitial signals	+	+	(Kenton et al. 1993)

Subgenomes are *N. sylvestris* (S) and *N. tomentosiformis* (T); ± signs indicate presence and absence, respectively; n.d., none detected

from the 26–18S rDNA intergenic spacer of *N. tomentosiformis*, was shown to be dispersed across several chromosomes (Lim et al. 2004b). A comparison of the distribution of A1/A2 repeats in species of section *Tomentosae* showed that they were present in low copy numbers in all species except *N. tomentosiformis* and likely expanded in the lineage leading to this species.

The HRS60 family of repeats occurs at a subtelomeric location in all diploid *Nicotiana* sections examined (i.e., *Alatae*, *Sylvestres*, *Undulatae*, and *Paniculatae*), but not in section *Tomentosae* where the HRS60 repeats are interstitial; the significance of this is currently unclear (Lim et al. 2000b, 2004b, 2005, 2006). All the HRS60 sequences are clearly related, and variants have diverged and undergone sequence homogenization such that each *Nicotiana* section has its own characteristic member of this family (Koukalova et al. 2010). The predominantly subtelomeric location of the HRS60 family of sequences and rDNA loci in *Nicotiana* means that the sequences interface directly with telomeric and degenerate telomeric motifs. The colocalisation of these apparently distinct domains may have consequences for plant genome evolution, and certainly they share some interesting biology; for example, they can generate noncoding transcripts, promote large-scale chromosomal rearrangements (e.g., Robertsonian fusions), and can be associated with nucleoli (Dvořáčková et al. 2015).

The GRD tandem repeat family is of particular interest and is thought to have derived from a free-living geminivirus (with which there is much sequence similarity) that integrated into the genome. This integration is proposed to have occurred in an ancestor of diploid species in section *Tomentosae*, where it amplified in a tandem array (GRD5 sequences on homologous chromosome 2). Later, specifically in the lineage leading to *N. tomentosiformis*, but before the formation of *N. tabacum*, a related sequence integrated *de novo* at a new chromosomal locus (GRD3 on chromosome 2) where it too amplified in a tandem array. Sequence analysis suggests that a large number of synonymous changes occurred in GRD3 compared with the number of synonymous changes in GRD5, indicative of purifying selection between integration events (Murad et al. 2002, 2004). The best explanation of these data is that there was a recombination event involving GRD5 and another free-living geminivirus, which generated the new GRD3 sequence that was subsequently re-integrated into the genome in a process that may have involved a helitron transposable element (Murad et al. 2004). Thus, *N. tabacum* inherited both GRD3 and GRD5 from *N. tomentosiformis* at its formation, and subsequently a variant sequence (GRD53) integrated at a third locus in the tobacco genome, potentially also involving a helitron (Murad et al. 2004).

7.5 Divergence of Repeats in Allopolyploid *Nicotiana*

Allopolyploids have formed in *Nicotiana* over widely different timescales (<0.5–6 Mya), frequently involving the same diploid parental lineages, and creating multiple species in section *Suaveolentes*. In addition, multiple synthetic polyploid lines have been developed. This makes the *Nicotiana* genus an ideal model to determine the fate of repetitive sequences and genome restructuring subsequent to polyploidization over a range of timescales. Using sequence-specific amplified polymorphisms to characterize the distribution of Copia-like retrotransposons, Petit et al. (2007) compared *N. tabacum* with its progenitor diploids and found that, in *N. tabacum*, the different Copia-like elements studied showed unique patterns that arose subsequent to the formation of *N. tabacum*. The frequencies of losses and gains of the observed insertion sites reflect retroelement mobility and sequence losses. Such changes are also apparent in the S4 generation of synthetic tobacco polyploids, where significant amplifications of Tnt1 Copia elements have been observed (Petit et al. 2010). As first proposed by McClintock (1984), the genomic shock of polyploidy, which is thought to arise from the merger of two distinct parental genomes in the nucleus, potentially stimulated rapid and dynamic

genomic changes early in *N. tabacum* divergence (Petit et al. 2010).

In addition to changes in the mobility of retroelements, the targeted loss of tandemly repeated and dispersed repeats in *N. tabacum*, particularly from the *N. tomentosiformis*-derived subgenome, has been reported from tobacco accessions (Skalicka et al. 2005; Renny-Byfield et al. 2011) and, to a lesser extent, from synthetic tobacco lines (Skalicka et al. 2003). For example, the NicCL3 tandem repeat sequence has a high copy number in *N. tomentosiformis* as in other diploids in section *Tomentosae* (Renny-Byfield et al. 2012). In contrast, in *N. tabacum* the NicCL3 tandem repeat sequence has a much lower copy number than in *N. tomentosiformis*, indicative of its loss from the *N. tomentosiformis*-derived subgenome of *N. tabacum* (Renny-Byfield et al. 2012). Similarly, there are fewer copies of an endogenous pararetrovirus-like sequence (Matzke et al. 2004) and of A1/A2 tandem repeats (Lim et al. 2004b) in *N. tabacum* than in *N. tomentosiformis*. The loss of the repeats from the paternally derived *N. tomentosiformis* progenitor's genome is an example of asymmetric loss of DNA from parental subgenomes, a phenomenon called biased fractionation (cf. Wendel 2015). This may have arisen from incompatibilities between the maternally derived cytoplasm and the biparentally derived nucleus at the time of allopolyploid formation (Leitch et al. 2006), as proposed by the nuclear–cytoplasmic interaction hypothesis (Gill and Friebe 2013).

A comparison of six transposable elements using sequence-specific amplified polymorphisms in synthetic polyploids of *N. tabacum*, *N. rustica*, and *N. arentsii* and multiple accessions of their diploid progenitors revealed that although element losses were evident in newly formed polyploids (Mhiri et al. 1997, 2019), element mobility is not apparent until later generations, suggesting that meiosis stimulates mobility. Nevertheless, it should be noted that the dynamics of the elements were not the same in the three synthetic polyploids and that the most active elements in the diploid progenitors were those that had the greatest impact on the early genome divergence of the polyploids (Mhiri et al. 2019).

Unlike the repeats described in the previous paragraphs, fluorescent in situ hybridization to metaphase spreads of *N. tabacum*, *N. rustica*, and *N. arentsii* revealed complete additivity in the number and distribution of rDNA loci to those found in the parental diploids (Lim et al. 2004a). In some S4 generation synthetic tobacco lines, there was perfect additivity of rDNA loci and rDNA units to that found in the parents; however, in other lines, new loci had evolved and new units amplified. Furthermore, in some of these lines, new units with close similarities to those in *N. tomentosiformis* have completely overwritten parental units. This process represents astonishingly rapid homogenization that has affected thousands of genes in just a few generations (Skalicka et al. 2003). In natural allopolyploids, rDNA units show varying degrees of rDNA homogenization depending on the allopolyploid species studied. *Nicotiana arentsii* (formed ∼0.4 Mya) shows complete homogenization toward an rDNA unit that is similar to that observed in the *N. undulata* parental genome, this process having thus replaced all the units derived from the *N. wigandioides* parent. In contrast, in the older polyploid *N. rustica* (∼0.7 Mya), we see the least change in the rDNA unit structure, with the different accessions analyzed showing only different ratios of parental rDNA unit types (Dadejová et al. 2007). In *N. tabacum*, there is large-scale, but incomplete, replacement of *N. sylvestris*-type units by units that are most similar to the *N. tomentosiformis* parent (Lim et al. 2000a). More recently, an analysis of rDNA diversity in *N. tabacum* using high-throughput sequence read data revealed that sequences in the 35S rDNA genic component of the rDNA unit are most homogenous, whereas those of the intergenic spacer are more diverse, indicating a more complete homogenization of the functional domains, or less subsequent divergence in these regions following homogenization (Lunerová et al. 2017).

There is no evidence to suggest that nuclear–cytoplasmic interactions affect the direction of rDNA unit homogenization, because the direction of replacement/bias is not always toward the maternal progenitor (Volkov et al. 1999; Lim et al. 2000a). This suggests that these processes involve different mechanisms to those that affect the loss of other high-copy repeats, as reported by Renny-Byfield et al. (2011) for *N. tabacum*. Thus, as described previously in nucleolar dominance studies, there is no indication that the parental direction of the cross has any effect on homogenization processes (Chen and Pikaard 1997). Instead, based on an analysis of rRNA transcripts in *N. tabacum* and *N. rustica*, it appears that the rDNA units that are most actively transcribed are also those that were amplified (Lunerová et al. 2017). Kovarik et al. (2008) proposed that epigenetic changes triggered by polyploid formation establish patterns of nuclear dominance that affect the expression of rDNA units, rendering the active units vulnerable to homogenization and unit conversion. Conversely, silent units, which are mostly condensed at interphase, may be less vulnerable to recombination processes, that may occur in the nucleolus and are most likely to mutate and be lost. In this way, variant types that are active are prone to amplify across the rDNA loci, first across intralocus tandem arrays, then less frequently between arrays (Kovarik et al. 2008; Lunerová et al. 2017).

In older polyploids in sections *Polydicliae* and *Repandae*, which formed approximately 1 and 4 Mya, respectively, the numbers of rDNA loci have dropped to two or three, as found in the progenitor diploids (Clarkson et al. 2005). In section *Polydicliae*, subtelomeric satellite DNAs from both parents can be found on the same chromosomes, a characteristic that is not observed in younger polyploids, which tend to maintain subtelomeric satellites intact (Kenton et al. 1993; Lim et al. 2004a, 2005; Matyasek et al. 2011). The distribution in section *Polydicliae* could only have arisen if recombination-based processes occurred between both parental subgenomes (Koukalova et al. 2010). In the even older polyploid section *Repandae* (∼4 Mya), three species (*N. nesophila*, *N. stocktonii*, and *N. repanda*) have a clade-specific satellite (NNES10) that has replaced other existing satellites found in the progenitor diploids, revealing the emergence, amplification, and homogenization of sequences over time frames of a few million years (Koukalova et al. 2010; Dodsworth et al. 2017).

Genomic in situ hybridization used to identify the parental origin of the chromosomes of polyploid species in section *Repandae* largely failed in early experiments (Lim et al. 2007). This was presumed to be due to large-scale replacement of the high copy repeats across the genome. Such genome turnover, in under 5 million years, is consistent with the half-life reported for repeats in rice, of ∼6 million years (Ma et al. 2004). Modifications to genomic in situ hybridization protocols have enabled the discrimination of parental subgenomes in all *Repandae* species (Dodsworth et al. 2017), albeit with differences between taxa. These experiments probably localize the lower copy repeat fractions more effectively and, in doing so, reveal intergenomic translocations (Dodsworth et al. 2017). As noted in Sect. 7.3, three of the species in section *Repandae* (*N. nesophila, N. stocktonii*, and *N. repanda*) show genome upsizing, largely arising from the expansion of some high-copy repeats, whereas a fourth species (*N. nudicaulis*) shows genome downsizing and is sister to the rest of the species in this section (Renny-Byfield et al. 2013; Dodsworth et al. 2017). The divergence of section *Repandae* is thus associated with a changing profile of repeats (Dodsworth et al. 2017) that reflects the evolutionary history of this and other sections of *Nicotiana* (Koukalova et al. 2010).

7.6 Lag-Phase Hypothesis and *Nicotiana*

Much has been written about the advantages of polyploidy, including, for example, fixed heterozygosity, the release from selection of redundant gene duplicates that can then take up new functions (subfunctionalization/neofunctionalization), and the mix-and-match of

Table 7.2 A hypothesis for genome evolution in polyploids, with reference, where possible, to *Nicotiana* species

Early polyploidy (first few 10–100s of generations)	
Chromosome pairing and meiotic instability	• Low fertility in many synthetic polyploids, as seen in Burk's synthetic *N. tabacum*
	• Strong selection pressure to establish regular meiosis expected, as seen in Burk's synthetic *N. tabacum*
	• Fixed heterozygosity of multiple alleles on different homologs leading to heterosis, which will be maintained with disomic inheritance of homologs
"Genomic shock" causing altered epigenetic patterns and triggering mobility of (retro)transposons	• Early-generation synthetic *Nicotiana* polyploids can show novel patterns of DNA methylation and deletions and insertions of rDNA loci and (retro)transposable elements
	• Genetic and epigenetic heterogeneity in the offspring creates a great deal of genetic variation upon which selection can act
Diploidization (occurring over a few million years)	
Genome downsizing	• A general trend following polyploidy and predicted to have occurred in some but not all *Nicotiana* polyploids. It is thought to be an evolutionary response to reduce the nutrient costs associated with additional DNA arising from polyploidy, and/or help overcome disadvantages associated with larger cells and/or cell cycle times
	• Preferential loss of DNA sequences, particularly high-copy repeats, from one of the progenitor genomes (e.g., loss of repetitive DNA, especially from the *N. tomentosiformis*-derived genome in *N. tabacum*)
	• Loss of low- and middle-copy repeat sequences reflecting diploidization and downsizing
Genome upsizing	• Amplification of high-copy repeats that can lead to genome turnover
rDNA locus number reduction	• rDNA locus number drops to a diploid-like number and is fixed for species-specific alleles within a few million years
Neofunctionalization/subfunctionalization and the evolution of new biochemistry/phenology	• Seen in the evolution of nicotine, a potent insecticide, in the genus *Nicotiana*
Lag phase and species radiation (5–10 million years)	
Rapid diversification	• Observed to have occurred in section *Suaveolentes* over the last 2–3 Mya, potentially triggered through novel traits that have advantages in multiple ecological environments
Dysploidy and chromosomal rearrangements	• Chromosomal reduction processes that lead to lower chromosome numbers, as occurred in section *Suaveolentes* from 24 chromosomes (tetraploid) down to 14 chromosomes

induced redundancy in biochemical pathways allows for the generation of new biochemistry (Soltis and Soltis 2000; Wendel 2015). In the polyploid events ancestral to Solanaceae and *Nicotiana*, there were duplications of the polyamine and nicotinamide adenine dinucleotide pathways that led to gene redundancy, reducing

selection pressure on these pathways, and the evolution of a new alkaloid, nicotine, which characterizes the genus (Xu et al. 2017).

Despite the advantages of polyploidy, there may also be both severe and subtle disadvantages. For example, in the early stages of polyploid formation, there is a potential for reduced

fitness, perhaps associated with nuclear–cyto-plasmic interactions and the aberrant segregation of chromosomes at meiosis, the latter seen in the synthetic tetraploid *N. tabacum* (Burk, 1973; Leitch et al. 2006) and in other well-studied tetraploids such as *Arabidopsis arenosa* (Yant et al. 2013). It has also been proposed that the extra burden of DNA associated with polyploidy may lead to resource limitation, particularly nitrogen and phosphorus, because nucleic acids are particularly demanding for these elements and many soils contain only small amounts of these nutrients (Šmarda et al. 2013; Guignard et al. 2016). Furthermore, increased GS is associated with larger cell sizes, longer cell cycle times, and negative impacts on photosynthesis (Greilhuber and Leitch 2013), all of which may act as a selection pressure against larger GSs, potentially leading to genome downsizing.

Given the high occurrence of polyploidy in the ancestry of most angiosperm lineages at a variety of nested phylogenetic levels (Wood et al. 2009), it has been proposed that the advantages likely outweighed disadvantages, especially from a long-term evolutionary viewpoint. Nonetheless, in the short term, polyploids are likely to show higher extinction rates (Mayrose et al. 2011), and, possibly, higher diversification rates than diploids (Soltis et al. 2014; Mayrose et al. 2015). Further analyses of diversification rates among angiosperms suggest nested shifts in diversification that do not correlate perfectly with predicted paleo-opolyploidy events, instead increased diversification rates appear to follow a lag phase (Tank et al. 2015; Landis et al. 2018). The lag-phase hypothesis (Schranz et al. 2012) posits that over several millions or tens of millions of years, polyploid species can undergo rapid radiations and speciation depending on subsequent evolutionary events. These events may include the diploidization process itself (e.g., genome downsizing and gene neo- and sub-functionalization) that occurs subsequent to polyploidy (Dodsworth et al. 2016). Potentially, such diploidization processes also facilitate adaptation to new environments and the radiation of diploidized taxa into new niches. This phenomenon is certainly occurring in section *Suaveolentes*, where, despite

the origin of the section being ~ 6 Mya (Clarkson et al. 2017), most species have evolved more recently (2–3 Mya; Chase et al. 2018), concomitant with moves into new habitats across the arid zones of Australia. A summary of the events that are likely to have occurred in that lag phase from a genomic viewpoint is given in Table 7.2, including an outline of the genomic processes affecting repeat dynamics over shorter timescales in polyploid *Nicotiana*.

References

Bejarano ER, Khashoggi A, Witty M, Lichtenstein C (1996) Integration of multiple repeats of geminiviral DNA into the nuclear genome of tobacco during evolution. Proc Natl Acad Sci 93(2):759–764

Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA, Martin GB (2012) A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. Mol Plant Microbe Interact 25(12):1523–1530

Borisjuk NV, Davidjuk YM, Kostishin SS, Miroshnichenco GP, Velasco R, Hemleben V, Volkov RA (1997) Structural analysis of rDNA in the genus *Nicotiana*. Plant Mol Biol 35(5):655–660

Burk LG (1973) Partial self fertility in a theoretical amphiploid progenitor of *N. tabacum*. J Hered 64 (6):348–350

Chase MW, Christenhusz MJ, Conran JG, Dodsworth S, Medeiros de Assis FN, Felix LP, Fay MF (2018) Unexpected diversity of Australian tobacco species (*Nicotiana* section Suaveolentes, Solanaceae) Curtis's Bot Mag 35(3):212–227

Chen ZJ, Pikaard CS (1997) Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. Proc Natl Acad Sci 94(7):3442–3447

Clarkson JJ, Dodsworth S, Chase MW (2017) Time-calibrated phylogenetic trees establish a lag between polyploidisation and diversification in *Nicotiana* (Solanaceae). Plant Syst Evol 303:1001–1012

Clarkson JJ, Kelly LJ, Leitch AR, Knapp S, Chase MW (2010) Nuclear glutamine synthetase evolution in *Nicotiana*: phylogenetics and the origins of allotetraploid and homoploid (diploid) hybrids. Mol Phylogenet Evol 55(1):99–112

Clarkson JJ, Knapp S, Garcia VF, Olmstead RG, Leitch AR, Chase MW (2004) Phylogenetic relationships in *Nicotiana* (Solanaceae) inferred from multiple plastid DNA regions. Mol Phylogenet Evol 33(1): 75–90

Clarkson JJ, Lim KY, Kovarik A, Chase MW, Knapp S, Leitch AR (2005) Long-term genome diploidization in

allopolyploid *Nicotiana* section *Repandae* (Solanaceae). New Phytol 168(1):241–252

Dadejová M, Lim KY, Soucková-Skalická K, Matyášek R, Grandbastien M-A, Leitch A, Kovařík A (2007) Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in *Nicotiana* allotetraploids. New Phytol 174(3):658–668

Dodsworth S, Chase MW, Leitch AR (2016) Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? Bot J Linn Soc 180(1):1–5

Dodsworth S, Jang T-S, Struebig M, Chase MW, Weiss-Schneeweiss H, Leitch AR (2017) Genome-wide repeat dynamics reflect phylogenetic distance in closely related allotetraploid *Nicotiana* (Solanaceae). Plant Syst Evol 303(8):1013–1020

Dvořáčková M, Fojtová M, Fajkus J (2015) Chromatin dynamics of plant telomeres and ribosomal genes. Plant J 83(1):18–37

El Baidouri M, Panaud O (2013) Comparative genomic paleontology across plant kingdom reveals the dynamics of TE-driven genome evolution. Genome Biol Evol 5(5):954–965

Fajkus J, Kovařík A, mKrálovics R, Bezděk M (1995a) Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. Mol Gen Genet MGG 247(5):633–638

Fajkus J, Královics R, Kovařík A, Fajkusová L (1995b) The telomeric sequence is directly attached to the HRS60 subtelomeric tandem repeat in tobacco chromosomes. FEBS Lett 364(1):33–35

Fulneček J, Lim KY, Leitch AR, Kovařík A, Matyášek R (2002) Evolution and structure of 5S rDNA loci in allotetraploid *Nicotiana tabacum* and its putative parental species. Heredity 88:19–25

Gazdová B, Široký J, Fajkus J, Brzobohatý B, Kenton A, Parokonny A, Heslop-Harrison JS, Palme K, Bezděk M (1995) Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. Chromosom Res 3(4):245–254

Gill BS, Friebe B (2013) Nucleocytoplasmic interaction hypothesis of genome evolution and speciation in polyploid plants revisited: polyploid species-specific chromosomal polymorphisms in wheat. In: Chen AJ, Birchler JA (eds) Polyploid and hybrid genomics, pp 213–221

Gregor W, Mette MF, Staginnus C, Matzke MA, Matzke AJM (2004) A distinct endogenous pararetrovirus family in *Nicotiana tomentosiformis*, a diploid progenitor of polyploid tobacco. Plant Physiol 134(3):1191–1199

Greilhuber J, Leitch IJ (2013) Genome size and the phenotype. In: Leitch IJ, Greilhuber J, Doležel J, Wendel JF (eds) Plant genome diversity, vol 2, Physical structure, behaviour and evolution of plant genomes. Springer, Wien, pp 323–344

Guignard MS, Nichols RA, Knell RJ, Macdonald A, Romila C-A, Trimmer M, Leitch IJ, Leitch AR (2016) Genome size and ploidy influence angiosperm species' biomass under nitrogen and phosphorus limitation. New Phytol 210:1195–1206

Horakova M, Fajkus J (2000) TAS49-a dispersed repetitive sequence isolated from subtelomeric regions of *Nicotiana tomentosiformis* chromosomes. Genome 43(2):273–284

Ibarra-Laclette E, Lyons E, Hernandez-Guzman G, Perez-Torres CA, Carretero-Paulet L, Chang T-H, Lan T, Welch AJ, Juarez MJA, Simpson J et al (2013) Architecture and evolution of a minute plant genome. Nature 498(7452):94–98

Kejnovsky E, Hawkins JS, Feschotte C (2012) Plant transposable elements: biology and evolution. In: Wendel JF, Greilhuber J, Doležel J, Leitch IJ (eds) Plant genome diversity, vol 1, Plant genomes, their residents, and their evolutionary dynamics. Springer-Verlag, Wien, pp 17–34

Kelly LJ, Leitch AR, Clarkson JJ, Hunter RB, Knapp S, Chase MW (2010) Intragenic recombination events and evidence for hybrid speciation in *Nicotiana* (Solanaceae). Mol Biol Evol 27(4):781–799

Kenton A, Parokonny AS, Gleba YY, Bennett MD (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Gen Genet 240(2):159–169

Knapp S, Chase MW, Clarkson JJ (2004) Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). Taxon 53(1):73–82

Koukalova B, Moraes AP, Renny-Byfield S, Matyasek R, Leitch AR, Kovarik A (2010) Fall and rise of satellite repeats in allopolyploids of *Nicotiana* over c. 5 million years. New Phytol 186(1):148–160

Koukalova B, Reich J, Bezdek M (1990) A BamHI family of tobacco highly repeated DNA—a study about its species-specificity. Biol Plant 32(6):445–449

Kovarik A, Dadejova M, Lim YK, Chase MW, Clarkson JJ, Knapp S, Leitch AR (2008) Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann Bot 101(6):815–823

Kovarik A, Koukalova B, Lim KY, Matyasek R, Lichtenstein CP, Leitch AR, Bezdek M (2000) Comparative analysis of DNA methylation in tobacco heterochromatic sequences. Chromosom Res 8(6):527–541

Kovarik A, Matyasek R, Lim KY, Skalicka K, Koukalova B, Knapp S, Chase M, Leitch AR (2004) Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. Biol J Linn Soc 82(4):615–625

Kuhrova V, Bezdek M, Vyskot B, Koukalova B, Fajkus J (1991) Isolation and characterization of two middle repetitive DNA sequences of nuclear tobacco genome. Theor Appl Genet 81:740–744

Landis JB, Soltis DE, Li Z, Marx HE, Barker MS, Tank DC, Soltis PS (2018) Impact of whole-genome duplication events on diversification rates in angiosperms. Am J Bot 105(3):348–363

Leitch AR, Leitch IJ (2012) Ecological and genetic factors linked to contrasting genome dynamics in seed plants. New Phytol 194(3):629–646

Leitch AR, Lim KY, Skalicka K, Kovarik A (2006) Nuclear cytoplasmic interaction hypothesis and the role of translocations in *Nicotiana* allopolyploids. NATO Security through Science Series 319–326

Leitch IJ, Bennett MD (2004) Genome downsizing in polyploid plants. Biol J Lin Soc 82:651–663

Leitch IJ, Hanson L, Lim KY, Kovarik A, Chase MW, Clarkson JJ, Leitch AR (2008) The ups and downs of genome size evolution in polyploid species of *Nicotiana* (Solanaceae). Ann Bot 101(6):805–814

Li W, Zhang P, Fellers JP, Friebe B, Gill BS (2004) Sequence composition, organization, and evolution of the core Triticeae genome. Plant J 40(4):500–511

Lim KY, Kovarik A, Matyasek R, Bezdek M, Lichtenstein CP, Leitch AR (2000a) Gene conversion of ribosomal DNA in *Nicotiana tabacum* is associated with undermethylated, decondensed and probably active gene units. Chromosoma 109(3):161–172

Lim KY, Kovarik A, Matyasek R, Chase MW, Clarkson JJ, Grandbastien MA, Leitch AR (2007) Sequence of events leading to near-complete genome turnover in allopolyploid *Nicotiana* within five million years. New Phytol 175(4):756–763

Lim KY, Kovarik A, Matyasek R, Chase MW, Knapp S, McCarthy E, Clarkson JJ, Leitch AR (2006) Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section *Alatae*. Plant J 48(6):907–919

Lim KY, Matyasek R, Kovarik A, Fulnecek J, Leitch AR (2005) Molecular cytogenetics and tandem repeat sequence evolution in the allopolyploid *Nicotiana rustica* compared with diploid progenitors *N. paniculata* and *N. undulata*. Cytogenet Genome Res 109(1–3):298–309

Lim KY, Matyasek R, Kovarik A, Leitch AR (2004a) Genome evolution in allotetraploid *Nicotiana*. Biol J Lin Soc 82(4):599–606

Lim KY, Matyasek R, Lichtenstein CP, Leitch AR (2000b) Molecular cytogenetic analyses and phylogenetic studies in the *Nicotiana* section *Tomentosae*. Chromosoma 109(4):245–258

Lim KY, Skalicka K, Koukalova B, Volkov RA, Matyasek R, Hemleben V, Leitch AR, Kovarik A (2004b) Dynamic changes in the distribution of a satellite homologous to intergenic 26-18S rDNA spacer in the evolution of *Nicotiana*. Genetics 166(4):1935–1946

Lunerová J, Renny-Byfield S, Matyášek R, Leitch AR, Kovařík A (2017) Concerted evolution rapidly eliminates sequence variation in rDNA coding regions but not in intergenic spacers in *Nicotiana tabacum* allotetraploid. Plant Syst Evol 303(8):1043–1060

Ma JX, Devos KM, Bennetzen JL (2004) Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. Genome Res 14(5):860–869

Matyasek R, Fulnecek J, Leitch AR, Kovarik A (2011) Analysis of two abundant, highly related satellites in the allotetraploid *Nicotiana arentsii* using double-strand conformation polymorphism analysis and sequencing. New Phytol 192(3):747–759

Matyasek R, Fulnecek J, Y Lim K, Leitch A, Kovarik A (2002) Evolution of 5S rDNA unit arrays in the plant genus *Nicotiana* (Solanaceae)

Matyasek R, Gazdová B, Fajkus J, Bezdek M (1997) NTRS, a new family of highly repetitive DNAs specific for the T1 chromosome of tobacco. Chromosoma 106:369–379

Matzke M, Gregor W, Mette MF, Aufsatz W, Kanno T, Jakowitsch J, Matzke AJM (2004) Endogenous pararetroviruses of allotetraploid *Nicotiana tabacum* and its diploid progenitors, *N. sylvestris* and *N. tomentosiformis*. Biol J Linn Soc 82(4):627–638

Mayrose I, Zhan SH, Rothfels CJ, Arrigo N, Barker MS, Rieseberg LH, Otto SP (2015) Methods for studying polyploid diversification and the dead end hypothesis: a reply to Soltis et al. (2014). New Phytol 206(1):27–35

Mayrose I, Zhan SH, Rothfels CJ, Magnuson-Ford K, Barker MS, Rieseberg LH, Otto SP (2011) Recently formed polyploid plants diversify at lower rates. Science 333(6047):1257

McClintock B (1984) The significance of responses of the genome to challenge. Science 226(4676):792–801

Melayah D, Lim KY, Bonnivard B, Chalhoub B, Dorlac de Borne F, Mhiri C, Leitch AR, Grandbastien MA (2004) Distribution of the Tnt1 retrotransposon family in the amphidiploid tobacco (*Nicotiana tabacum*) and its wild *Nicotiana* relatives. Biol J Linn Soc 82:639–649

Mhiri C, Morel J-B, Vernhettes S, Casacuberta JM, Lucas H, Grandbastien M-A (1997) The promoter of the tobacco Tnt1 retrotransposon is induced by wounding and by abiotic stress. Plant Mol Biol 33(2):257–266

Mhiri C, Parisod C, Daniel J, Petit M, Lim KY, Dorlhac de Borne F, Kovarik A, Leitch AR, Grandbastien M-A (2019) Parental transposable element loads influence their dynamics in young *Nicotiana* hybrids and allotetraploids. New Phytol 221(3):1619–1633

Moscone EA, Matzke MA, Matzke AJ (1996) The use of combined FISH/GISH in conjunction with DAPI counterstaining to identify chromosomes containing transgene inserts in amphidiploid tobacco. Chromosoma 105(5):231–236

Murad L, Bielawski JP, Matyasek R, Kovarik A, Nichols RA, Leitch AR, Lichtenstein CP (2004) The origin and evolution of geminivirus-related DNA sequences in *Nicotiana*. Heredity 92(4):352

Murad L, Lim KY, Christopodulou V, Matyasek R, Lichtenstein CP, Kovarik A, Leitch AR (2002) The origin of tobacco's T genome is traced to a particular lineage within *Nicotiana tomentosiformis* (Solanaceae). Am J Bot 89(6):921–928

Nagaki K, Shibata F, Suzuki G, Kanatani A, Ozaki S, Hironaka A, Kashihara K, Murata M (2011) Coexistence of NtCENH3 and two retrotransposons in tobacco centromeres. Chromosom Res 19(5):591–605

Neumann P, Novák P, Hošťáková N, Macas J (2019) Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mobile DNA 10(1):1

Novak P, Neumann P, Macas J (2010) Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinform 11(1):378

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J (2013) RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics 29(6):792–793

Pellicer J, Fay MF, Leitch IJ (2010) The largest eukaryotic genome of them all? Bot J Linn Soc 164(1):10–15

Pellicer J, Hidalgo O, Dodsworth S, Leitch IJ (2018) Genome size diversity and its impact on the evolution of land plants. Genes 9(2):88

Petit M, Guidat C, Daniel J, Denis E, Montoriol E, Bui QT, Lim KY, Kovarik A, Leitch AR, Grandbastien MA et al (2010) Mobilization of retrotransposons in synthetic allotetraploid tobacco. New Phytol 186(1):135–147

Petit M, Lim K, Julio E, Poncet C, de Borne Dorlhac, Fo Kovarik A, Leitch AR, Grandbastien M-Al, Mhiri C (2007) Differential impact of retrotransposon populations on the genome of allotetraploid tobacco (Nicotiana tabacum). Mol Genet Genomics 278(1):1–15

Renny-Byfield S, Chester M, Kovařík A, Le Comber SC, Grandbastien M-A, Deloger M, Nichols RA, Macas J, Novák P, Chase MW et al (2011) Next generation sequencing reveals genome downsizing in allotetraploid Nicotiana tabacum, predominantly through the elimination of paternally derived repetitive DNAs. Mol Biol Evol 28(10):2843–2854

Renny-Byfield S, Kovarik A, Chester M, Nichols RA, Macas J, Novak P, Leitch AR (2012) Independent, rapid and targeted loss of highly repetitive DNA in natural and synthetic allopolyploids of Nicotiana tabacum. PLoS ONE 7(5):e36963

Renny-Byfield S, Kovarik A, Kelly LJ, Macas J, Novak P, Chase MW, Nichols RA, Pancholi MR, Grandbastien MA, Leitch AR (2013) Diploidization and genome size change in allopolyploids is associated with differential dynamics of low- and high-copy sequences. Plant J 74(5):829–839

Särkinen T, Bohs L, Olmstead RG, Knapp S (2013) A phylogenetic framework for evolutionary study of the nightshades (Solanaceae): a dated 1000-tip tree. BMC Evol Biol 13(1):214

Schiavinato M, Marcet-Houben M, Dohm JC, Gabaldón T, Himmelbauer H (2020) Parental origin of the allotetraploid tobacco. Plant J

Schranz EM, Mohammadin S, Edger PP (2012) Ancient whole genome duplications, novelty and diversification: the WGD radiation lag-time model. Curr Opin Plant Biol 15(2):147–153

Shibata F, Nagaki K, Yokota E, Murata M (2013) Tobacco karyotyping by accurate centromere identification and novel repetitive DNA localization. Chromosom Res 21(4):375–381

Skalicka K, Lim KY, Matyasek R, Koukalova B, Leitch AR, Kovarik A (2003) Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. Am J Bot 90(7):988–996

Skalicka K, Lim KY, Matyasek R, Matzke M, Leitch AR, Kovarik A (2005) Preferential elimination of repeated DNA sequences from the paternal, Nicotiana tomentosiformis genome donor of a synthetic, allotetraploid tobacco. New Phytol 166(1):291–303

Šmarda P, Hejcman M, Březinová A, Horová L, Steigerová H, Zedek F, Bureš P, Hejcmanová P, Schellberg J (2013) Effect of phosphorus availability on the selection of species with different ploidy levels and genome sizes in a long-term grassland fertilization experiment. New Phytol 200:911–921

Soltis DE, Segovia-Salcedo MC, Jordon-Thaden I, Majure L, Miles NM, Mavrodiev EV, Mei W, Cortez MB, Soltis PS, Gitzendanner MA (2014) Are polyploids really evolutionary dead-ends (again)? A critical reappraisal of Mayrose et al. (2011). New Phytol 202(4):1105–1117

Soltis PS, Soltis DE (2000) The role of genetic and genomic attributes in the success of polyploids. Proc Natl Acad Sci (USA) 97(13):7051–7057

Tank DC, Eastman JM, Pennell MW, Soltis PS, Soltis DE, Hinchliff CE, Brown JW, Sessa EB, Harmon LJ (2015) Nested radiations and the pulse of angiosperm diversification: increased diversification rates often follow whole genome duplications. New Phytol 207:454–467

Volkov RA, Borisjuk NV, Panchuk II, Schweizer D, Hemleben V (1999) Elimination and rearrangement of parental rDNA in the allotetraploid Nicotiana tabacum. Mol Biol Evol 16(3):311–320

Wendel JF (2015) The wondrous cycles of polyploidy in plants. Am J Bot 102(11):1753–1756

Wood TE, Takebayashi N, Barker MS, Mayrose I, Greenspoon PB, Rieseberg LH (2009) The frequency of polyploid speciation in vascular plants. Proc Natl Acad Sci (USA) 106(33):13875–13879

Xu S, Brockmöller T, Navarro-Quezada A, Kuhl H, Gase K, Ling Z, Zhou W, Kreitzer C, Stanke M, Tang H et al (2017) Wild tobacco genomes reveal the evolution of nicotine biosynthesis. Proc Natl Acad Sci 114(23):6133–6138

Yant L, Hollister Jesse D, Wright Kevin M, Arnold Brian J, Higgins James D, Franklin FChris H, Bomblies K (2013) Meiotic adaptation to genome duplication in Arabidopsis arenosa. Curr Biol 23(21):2151–2156