TERT Gene in Polyploid Plants

Subjects: Plant Sciences

Contributor: Eva Sykorova

The gene coding for the telomerase reverse transcriptase (TERT) is essential for the maintenance of telomeres. Previously we described the presence of three TERT paralogs in the allotetraploid plant Nicotiana tabacum, while a single TERT copy was identified in the paleopolyploid model plant Arabidopsis thaliana. Here we examine the presence, origin and functional status of TERT variants in allotetraploid Nicotiana species of diverse evolutionary ages and their parental genome donors, as well as in other diploid and polyploid plant species. A combination of experimental and in silico bottom-up analyses of TERT gene copies in Nicotiana polyploids revealed various patterns of retention or loss of parental TERT variants and divergence in their functions. RT–qPCR results confirmed the expression of all the identified TERT variants. In representative plant and green algal genomes, our synteny analyses show that their TERT genes were located in a conserved locus that became advantageous after the divergence of eudicots, and the gene was later translocated in several plant groups. In various diploid and polyploid species, translocation of TERT became fixed in target loci that show ancient synapomorphy.

polyploidy Nicotiana telomerase gene evolution synteny

1. Introduction

Flowering plants (angiosperms) are important for the existence of many terrestrial organisms, including humans, and a long history of plant breeding has taught us that polyploidization can be advantageous in terms of quantitative traits of crops. Gains and losses of paralogs, their neofunctionalization and sub-functionalization, have all been associated with the generation of duplicate gene copies, e.g., by whole-genome duplications (WGDs) and further rounds of genome duplication/reduction, resulting in genetic diversity upon which the fittest combinations thrived in a competitive environment [1][2][3][4]. An ancient WGD has been reconstructed at the base of seed plants, another at the base of angiosperms [5][6][7] and numerous additional, subsequent WGD events were associated with the divergence of many angiosperm lineages [3]. Polyploidy is usually associated with many genetic and epigenetic changes, including chromosomal rearrangements, expansions of transposable elements and changes in gene expression [8][9]. At the gene level, polyploids can tolerate the presence of paralogs or eliminate a copy of the spare gene. Thus, evolutionary forces result in an equilibrium defined by gene dosage [10]. Studies of model plants have mostly focused on genes important for crop production; however, genes that are critical for genome stability are extremely important for understanding repeated polyploidization events during natural selection, and these remain underexplored.

Telomerase reverse transcriptase (TERT) is involved in the maintenance of telomeres, nucleoprotein structures that are essential for genome stability [11][12][13]. Telomerase adds telomere repeats to the ends of eukaryotic

chromosomes, thereby elongating telomeres and compensating for their shortening due to incomplete end-replication. When telomerase is not active, telomeres become shortened, and their function in the protection of chromosomes is disrupted. The extreme evolutionary success of telomerase-based mechanisms of telomere maintenance is illustrated by current findings in plants (reviewed in [14]). Even among apparent exceptions in telomere sequences, in plant genera *Allium* (Asparagales) and *Cestrum* (Solanales) [15][16][17][18], recent research has revealed that novel, unusual telomere DNA sequences are synthesized by telomerase [16][18][19] and not by alternative mechanisms as had been suggested previously (reviewed in [20]). Moreover, we recently demonstrated that changes in the template region of the telomerase RNA subunit directed the observed evolutionary transitions in telomere DNA sequences [14][21][22]. In contrast to the RNA subunit, the protein subunit TERT is evolutionary well conserved and possesses a central reverse transcriptase domain essential for its catalytic function [23][24]. Plant TERTs are structurally similar to human, ciliate or yeast TERTs with a telomerase-specific T motif [25][26][27][28][29]. The gene encoding TERT is usually expressed at low mRNA levels even in telomerase-positive tissues and is maintained as a single copy gene in most eukaryotic genomes. However, the natural allotetraploid *Nicotiana tabacum* possesses three sequence variants of the *TERT* gene [30]. Various allopolyploidization events among closely and distantly related diploid parental species (Figure 1) in *Nicotiana* make the genus an ideal experimental model system to study the long-term evolution of *TERT* following natural gene duplication. The increasing number of publicly available assembled plant genomes enables the exploration of *TERT* genomic loci, gene copy numbers and gene synteny in diverse plant species for comparisons with the data from *Nicotiana* polyploids and the diploid species most closely related to their progenitors (hereafter called progenitor diploids). The *Nicotiana* genus [31][32][33][34][35] comprises relatively young polyploids (i) *N. tabacum* (section Nicotianae), *N. rustica* (sect. Rusticae), *N. arentsii* (sect. Undulatae) that formed approx. 0.4–0.6 million years ago, (ii) *N. clevelandii* and *N. quadrivalvis* (ca. 1.5 million years ago, sect. Polydicliae), (iii) four species from the 4–5 million years old section Repandae (*N. nudicaulis*, *N. repanda*, *N. nesophila* and *N. stocktonii*), and (iv) ~35 species including the model *N. benthamiana* from the oldest section Suaveolentes formed about 6 million years ago [31]. Among these species, members of sections Suaveolentes and Repandae are of interest because, with *N. tabacum*, they share an ancient genome donor, *N. sylvestris*, and these speciation events happened at different times. In *N. tabacum*, two *TERT* variants originated from the maternal *N. sylvestris* genome (*TERT*_Cs, *TERT*_D) and one from the *N. tomentosiformis* paternal genome (*TERT*_Ct). Variants *TERT*_Cs and *TERT*_Ct code for a full-length functional protein, while the *TERT*_D variant is truncated and contains several indels resulting in premature stop codons, suggesting that it is a pseudogene [30]. All three variants are nevertheless transcribed and show distinct, tissue-dependent levels of mRNA transcripts, indicating a sub-functionalization of *TERT* variants [30][36].

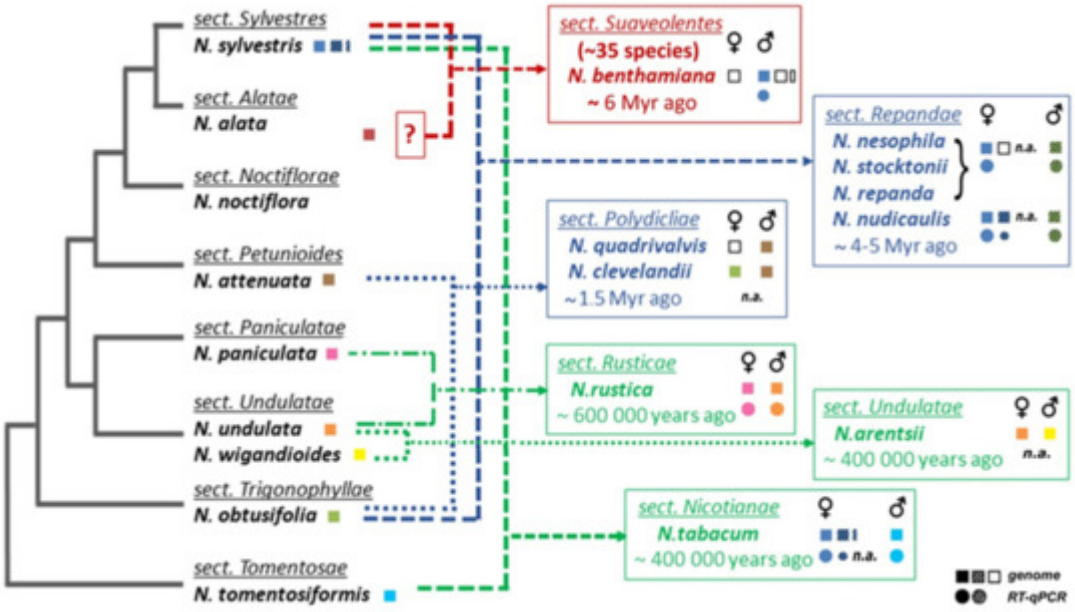

Figure 1. Overview of experimental results and illustration of phylogenetic relationships of *Nicotiana* species used in this study. Phylogeny and the proposed origin of polyploids were adapted from [31][32]. An uncertain parental genome donor for section Suaveolentes is indicated by a question mark. Summary of experimental and in silico results (squares, analyses of genomic DNA; circles, expression of *TERT* variants investigated by RT–qPCR) is shown in boxes of Nicotiana sections, the origin of *TERT* variant in polyploids is depicted by color of respective parental diploids, and variants that were not identified are depicted with open squares. *Nicotiana* accessions used in the experimental analyses are listed in Table S1, genomic assemblies and genomic/transcriptomic SRA data used for in silico analyses are listed in Material and Methods. For the purposes of this paper, we refer to a *TERT* copy that does not code for a catalytically active protein as a putative pseudogene (dashed symbols) in contrast to a functional *TERT* gene copy (open symbols), n.a. not analyzed.

Based on previously described *TERT* variants in *N. tabacum*, we explored the fate of *TERT* paralogs in other *Nicotiana* polyploids to determine whether both parental *TERT* genes are conserved in allotetraploid genomes, whether they are transcribed, present in syntenic, collinear arrays with their progenitor diploids, and whether any relationship exists between telomere lengths in polyploids and their progenitor diploids. Of particular interest in this study was to clarify the origin of the presumed pseudogene variant *TERT_D* in *N. sylvestris*, a diploid genome donor of *N. tabacum*, as well as of even older species from sections Repandae and Suaveolentes. In addition, we investigated in silico whether diploid and polyploid plants outside of the family Solanaceae sustained *TERT* paralogs/pseudogenes in their genomes, and we explored syntenic relationships of genes adjacent to *TERT* to interpret the evolutionary success of *TERT* copies after translocation.

2. Number of TERT Variants in Nicotiana Polyploids as a Case Study

At the beginning of this project, there was limited genomic sequence data available for the majority of *Nicotiana* allopolyploids and their parents. To characterize experimentally the number, identity and origin of *TERT* copies in genomes of polyploid *Nicotiana* species and representatives of their diploid progenitors, we employed several primer combinations derived from conserved *TERT* regions of the evolutionarily distant relatives *N. sylvestris* and *N. tomentosiformis* (Figure 1), designed originally for amplification of *N. tabacum TERT* variants [30][36]. These PCR primers (Figure 2A, Table S2) amplify *TERT* regions nonspecifically, i.e., all variants are produced in a single PCR. Sequencing of PCR products then identifies single nucleotide polymorphisms (SNPs) and/or indels evidencing the presence of multiple *TERT* variants. Primer positions were with respect to *Nicotiana TERT* gene structure with 13 exons (Figure 2A), which differed from the prevalent 12-exon structure of plant *TERT*s [23]. As expected, a successful amplification was achieved mostly using primers derived from the more conserved sequences at the 3′ end of *TERT* genes (Table S3). As the first screening experiment, we applied this approach to six diploid *Nicotiana* species investigated as representatives of parental genome donors, including *N. sylvestris* as a control, and to nine polyploid *Nicotiana* species (Figure 1). Among parental diploids, we detected one *TERT* variant in *N. alata*, *N. attenuata*, *N. undulata*, *N. wigandoides*, *N. paniculata* and *N. obtusifolia* (Supplementary A1), and two *TERT* variants (*TERT*_C and *TERT*_D) in *N. sylvestris* [30]. In the case of *N. attenuata* and *N. obtusifolia*, species representing parents of polyploid sections Polydicliae and Repandae, we further confirmed our results by in silico analysis using genome assemblies (GenBank accessions: GCA_001879085.1 and GCA_002018475.1, respectively). To complete the set of representative parental species, we assembled available transcriptomic SRA data of *N. noctiflora* (GenBank accession: SRR2106514) and identified one *TERT* variant. In conclusion, our results show the presence of more than one *TERT* variant in diploid *N. sylvestris* [30], an exception among parental species of *Nicotiana* polyploids.

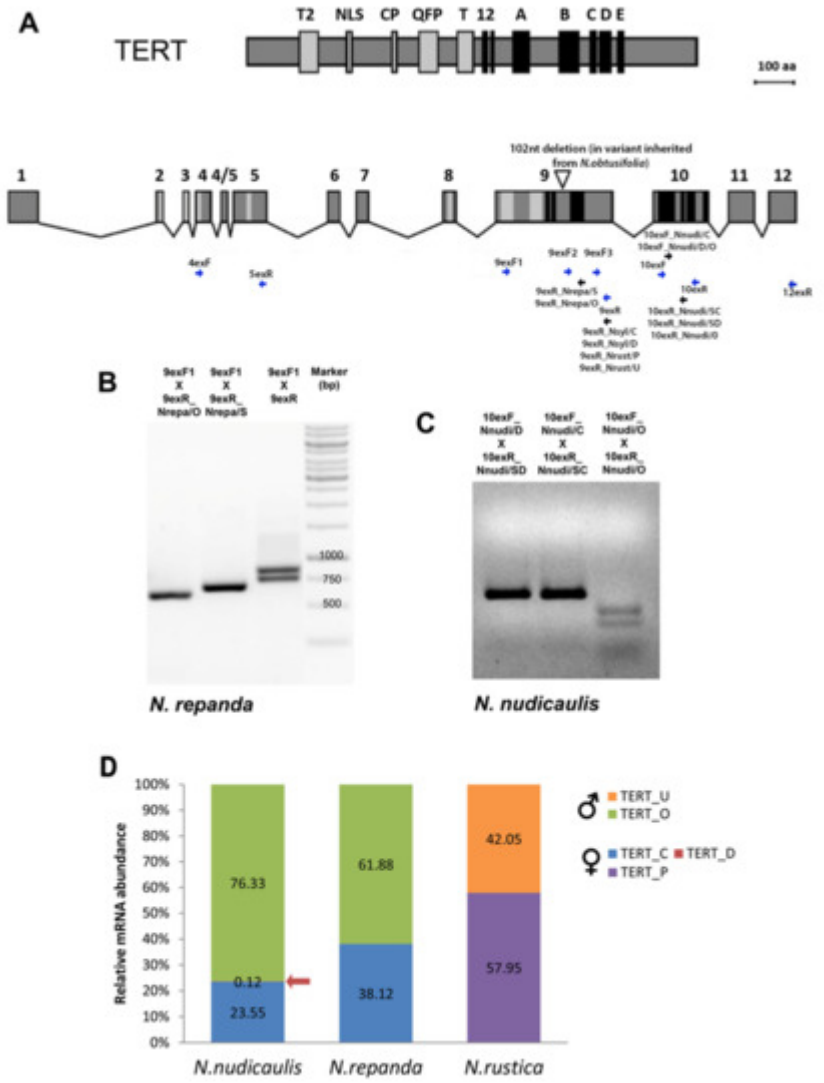

Figure 2. Experimental identification of TERT variants and analysis of gene expression in *Nicotiana* polyploids. (**A**) Conserved telomerase specific motifs (T2, NLS, CP, QFP, T) and reverse transcriptase motifs (1, 2, A–E) are highlighted in protein and mRNA of *Nicotiana* TERT (modified from [30]). Positions of primers used for screening experiments (blue arrows) and *TERT*-variant-specific primers (black arrows) are indicated at corresponding *TERT* mRNA regions (primers are listed in Table S2). The triangle within exon 9 shows the position of a 102 nt long deletion that was identified in *N. repanda*, *N. nesophila* and *N. stocktonii* and represents a specific *TERT*-variant of *N. obtusifolia* origin. (**B**,**C**) Validation of primer specificity for *TERT* variants in *N. repanda* (**B**) and *N. nudicaulis* (**C**). PCR products amplified with primers 9exF1 and 9exR1 show two bands corresponding to *TERT*_O and *TERT*_Cs variants that differ by a 102 bp long deletion. Specific amplification of *TERT*_O and *TERT*_Cs variants was demonstrated using the 9exF1 primer in combination with variant-specific reverse primers 9exR_Nrepa/O and 9exR_Nrepa/S, respectively. (**C**) For validation of qPCR primers and to distinguish three *TERT* variants in *N. nudicaulis*, the PCR products amplified with indicated qPCR primer combinations were digested with *Mse*I. A specific cut of the *TERT*_O variant that possesses the restriction site for *Mse*I within the amplified region confirmed the specificity of amplified *TERT*-variants. (**D**) Relative mRNA levels of specific *TERT* variants were determined by RT–qPCR in *N. nudicaulis*, *N. repanda* and *N. rustica*. Relative

mRNA abundance of particular parental *TERT* variants (in %) was calculated by the delta Ct method [37]. Ct values were normalized using the reaction efficiency calculated from a standard curve analysis (Table S3).

The same experimental approach applied to representative *Nicotiana* polyploids detected variant-specific SNPs and/or indels, demonstrating the presence of two *TERT* variants in 5 of 9 polyploid species investigated (*N. arentsii*, *N. rustica*, *N. repanda*, *N. nesophila*, *N. stocktonii*) and three variants were identified in *N. nudicaulis* (summarized in Figure 1, Table 1, see below for details). While PCR products obtained from *N. clevelandii, N. quadrivalvis* and *N. benthamiana* genomic DNA revealed the presence of a single copy of the *TERT* gene, our search for *TERT* variants in raw transcriptomic data from *N. clevelandii* showed the occurrence of two gene variants. To avoid possible errors in comparison of experimental and in silico data that could be caused, e.g., by possible incorrect mapping of *TERT* reads to the raw genome/transcriptome data, assembly version or allele sequence, we analyzed in detail individual SNPs in sequences from each polyploid species and its progenitor diploids (see Supplemental Text S1, Figure S1, Table S4). Results deduced from sequence similarity (in %, Table 1) and individual SNPs (Table S4) were in agreement in all cases analyzed.

Table 1. Origin of telomerase reverse transcriptase (*TERT*) variants in polyploid *Nicotiana* species determined by sequence similarity with representative progenitor diploids.

Allopolyploids	GeneBank Accessions	Sequence Similarity [%]				Analyzed Region [1]
		Maternal Parent		Paternal Parent		
SUAVEOLENTES		***N. alata***	***N. noctiflora*** [2]	***N. syl. C var.***	***N. syl. D var.***	
N. benthamiana	NbS000104	96.3	n.a.	**97.5**	n.a.	exon 4 to 5
	27g0116.1	n.a.	96.1	**97.6**	93.4	exons 10, 11, 12
REPANDAE		***N. syl. C var.***	***N. syl. D var.***	***N. obtusifolia***		
N. repanda	MG242402 [1]	95.9	91.6	**97.4**		exon 9
	MG242403 [1]	**97.9**	92.4	96.4		exon 9
N. stocktonii	MG242407 [1]	95.6	91.7	**97.6**		exon 9
	MG242408 [1]	**98.6**	93.1	97.0		exon 9
N. nesophila	MG242405 [1]	95.2	91.6	**97.0**		exon 9
	MG242406 [1]	**98.5**	92.9	96.9		exon 9
N. nudicaulis	MG242409 [1]	**98.6**	94.3	94.4		exon 10 to 12
	MG545647 [1]	92.8	**94.8**	91.6		exon 10 to 12

Allopolyploids	GeneBank Accessions	Sequence Similarity [%] Maternal Parent	Paternal Parent	Analyzed Region [1]
	MG242410 [1]	94.2 93.3	**96.3**	exon 10 to 12
POLYDICLIAE		***N. obtusifolia***	***N. attenuata***	
	MG242422 [1]	94.3	**99.3**	exon 4 to 5
N. clevelandii	var1 [2]	97.3	**98.9**	exon 9 [2]
	var2 [2]	**99.2**	97.3	exon 9 [2]
N. quadrivalvis	MG242423 [1]	94.9	**98.6**	exon 4 to 5
ARENTSII		***N. undulata***	***N. wigandiodes***	
N. arentsii	MG242418 [1]	**99.5**	98.4	exon 9
	MG242419 [1]	98.8	**99.8**	exon 9
RUSTICA		*N. paniculata*	*N. undulata*	
N. rustica	MG242413 [1]	**100.0**	98.2	exon 9
	MG242414 [1]	98.2	**99.8**	exon 9

[1] all sequences cloned in this work are in Supplementary A1, including corresponding sequences cloned from progenitor diploids; [2] regions mapped to raw NGS data or extracted from genome assembly (Supplementary A1).

3. Origin of TERT Genes in Polyploids with the Ancestral N. sylvestris Donor Genome

An *N. sylvestris* progenitor is considered to be a progenitor diploid of the allopolyploid sections Suaveolentes, Repandae and Nicotianae (Figure 1). The evolutionary history of *TERT* associated with allopolyploidy is inferred for each of these sections.

3.1. Suaveolentes

In the section Suaveolentes we used the model plant *N. benthamiana* as a representative species of the section, and *N. alata* and *N. noctiflora* as recent relatives of the putative maternal lineage originating from sections Alatae or Noctiflorae, respectively (Figure 1, [31][32][38]). We detected a single copy of the *N. benthamiana TERT* experimentally, and this result was confirmed in silico using (i) an *N. benthamiana* genome assembly based on deep sequencing (N. benthamiana Genome v1.0.1) and (ii) analysis of raw genomic NGS reads [39] by BLAST followed by read-mapping back to the query. Comparison of corresponding regions of *N. benthamiana TERT* and

2. Soltis, D.E.; Visger, C.J.; Marchant, D.B.; Soltis, P.S. Polyploidy: Pitfalls and paths to a paradigm. Am. J. Bot. 2016, 103, 1146–1166.

3. Lohaus, R.; Van de Peer, Y. Of dups and dinos: Evolution at the K/Pg boundary. Curr. Opin. Plant Biol. 2016, 30, 62–69.

4. Barker, M.S.; Husband, B.C.; Pires, J.C. Spreading Winge and flying high: The evolutionary importance of polyploidy after a century of study. Am. J. Bot. 2016, 103, 1139–1145.

5. Jiao, Y.; Wickett, N.J.; Ayyampalayam, S.; Chanderbali, A.S.; Landherr, L.; Ralph, P.E.; Tomsho, L.P.; Hu, Y.; Liang, H.; Soltis, P.S.; et al. Ancestral polyploidy in seed plants and angiosperms. Nature 2011, 473, 97–100.

6. Jiao, Y.; Leebens-Mack, J.; Ayyampalayam, S.; Bowers, J.E.; McKain, M.R.; McNeal, J.; Rolf, M.; Ruzicka, D.R.; Wafula, E.; Wickett, N.J.; et al. A genome triplication associated with early diversification of the core eudicots. Genome Biol. 2012, 13, R3.

7. Murat, F.; Armero, A.; Pont, C.; Klopp, C.; Salse, J. Reconstructing the genome of the most recent common ancestor of flowering plants. Nat. Genet. 2017, 49, 490–496.

8. Flagel, L.; Udall, J.; Nettleton, D.; Wendel, J. Duplicate gene expression in allopolyploid Gossypium reveals two temporally distinct phases of expression evolution. BMC Biol. 2008, 6, 16.

representative parental *TERT* sequences (Table 1, Table S4) revealed that the *N. benthamiana TERT* sequence (accession number NbS00010427g0016.1) was more similar to *N. sylvestris TERT*_C variant than to the *TERT* sequence cloned from *N. alata* (GenBank accession MG242421) or deduced from *N. noctiflora* SRA data (Supplementary A1). Thus, we conclude an *N. sylvestris* origin of *N. benthamiana TERT* and a loss of the second parental *TERT* copy during the evolution of *N. benthamiana*.

4. Discussion

To test experimentally and in silico *TERT* gene balance following ancient polyploidization events, we identified and characterized *TERT* copies in genomes of polyploid *Nicotiana* species and representatives of their diploid progenitors. We also investigated the expression of *TERT* variants identified in the polyploids using RT–qPCR. We found that the *N. sylvestris* progenitor was a very successful parent of sections Suaveolentes, Repandae and Nicotianae because the *TERT*_Cs variant of *N. sylvestris* origin was identified in all polyploid genomes investigated (Figure 1), and high levels of its transcripts were detected. Moreover, an adjacent occurrence of *TERT* copies observed in *N. tabacum* and *N. nudicaulis* suggests that gene/genome duplication resulting in the formation of *TERT*_C and *TERT*_D variants in *N. sylvestris* had occurred at least before the formation of the section Repandae. The *TERT*_D transcripts were detectable but heavily under-represented in *N. nudicaulis* (Figure 2D), similar to *TERT*_D expression in *N. sylvestris* and *N. tabacum* [30][36]. In contrast to the success of the *N. sylvestris* progenitor, the *TERT*_O variant of *N. obtusifolia* origin was pseudogenized in all four polyploid species from Repandae. A 102-nt-long in-frame deletion within exon 9 would shorten the linker region between motif 2 and motif A, including protein motif GSSVP that is well-conserved in plant TERTs. This region, termed as motif 3 in human TERT, was found to be crucial for telomerase catalytic functions [40]; however, its absence is not the only problem in *TERT*_O variants. Various indels found across *TERT*_O variants from Repandae would result in out-of-frame mutations, and interestingly, a nucleotide transition found within motif C of *N. nudicaulis TERT*_O would disrupt one of three Asp residues that are essential for the catalytic function of any telomerase [28][29]. However, mRNA levels of the *TERT*_O variant revealed expression comparable to the *TERT*_Cs variant in *N. nudicaulis* and *N. repanda* (Figure 2D). Comparable transcript levels of parental *TERT* variants coding for the functional TERT protein were detected in the relatively young polyploid, *N. rustica* (*TERT*_P and *TERT*_U, Figure 2D), and similarly in *N. tabacum* [36].

Our experimental analyses were accompanied by an in silico approach to answering the question on the origin and fate of the *TERT*_D variant in the *N. sylvestris* genome and, for a wider perspective, in other polyploid plant genomes. Our experimentally estimated ratio 1:1 of *TERT*_C and *TERT*_D gene copies in five *N. sylvestris* accessions was confirmed by in silico analysis of raw data from the *N. sylvestris* genome sequencing project (Table 2). Moreover, we identified a part of the *TERT*_D variant sequence (*TERT*_12D) associated with high-copy repetitive sequences and the *MtATPO* gene within a novel genomic locus in the *N. sylvestris* genome. An unplaced genomic scaffold arranged similarly to the *TERT*_12D locus was identified in *N. tabacum*, suggesting that an ancestral split of the *TERT*_D copy had occurred at least before the formation of *N. tabacum*. There is no information about a species-specific WGD event or an additional genome donor in *N. sylvestris*, but the increase in

9. Parisod, C.; Mhiri, C.; Lim, K.Y.; Clarkson, J.J.; Chase, M.W.; Leitch, A.R.; Grandbastien, M.A. Differential dynamics of transposable elements during long-term diploidization of Nicotiana section Repandae (Solanaceae) allopolyploid genomes. PLoS ONE 2012, 7, e50352.
10. Birchler, J.A.; Veitia, R.A. The gene balance hypothesis: From classical genetics to modern genomics. Plant Cell 2007, 19, 395–402.
11. McClintock, B. The fusion of broken chromosome ends of sister half-chromatids following chromatid breakage at meiotic anaphases. Mo. Agric. Exp. Stn. Res. Bull. 1938, 290, 1–48.
12. Blackburn, E.H.; Gall, J.G. Tandemly Repeated Sequence at Termini of Extrachromosomal Ribosomal-Rna Genes in Tetrahymena. J. Mol. Biol. 1978, 120, 33–53.
13. Greider, C.W.; Blackburn, E.H. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell 1985, 43, 405–413.
14. Peska, V.; Garcia, S. Origin, Diversity, and Evolution of Telomere Sequences in Plants. Front. Plant Sci. 2020, 11, 117.
15. Sykorova, E.; Lim, K.Y.; Kunicka, Z.; Chase, M.W.; Bennett, M.D.; Fajkus, J.; Leitch, A.R. Telomere variability in the monocotyledonous plant order Asparagales. Proc. R. Soc. Lond. B Biol. Sci. 2003, 270, 1893–1904.
16. Fajkus, P.; Peska, V.; Sitova, Z.; Fulneckova, J.; Dvorackova, M.; Gogela, R.; Sykorova, E.; Hapala, J.; Fajkus, J. Allium telomeres unmasked: The unusual telomeric sequence (CTCGGTTATGGG)n is synthesized by telomerase. Plant J. 2016, 85, 337–347.
17. Sykorova, E.; Lim, K.Y.; Chase, M.W.; Knapp, S.; Leitch, I.J.; Leitch, A.R.; Fajkus, J. The absence of Arabidopsis-type telomeres in Cestrum and closely related genera Vestia and Sessea (Solanaceae): First evidence from eudicots. Plant J. 2003, 34, 283–291.
18. Peska, V.; Fajkus, P.; Fojtova, M.; Dvorackova, M.; Hapala, J.; Dvoracek, V.; Polanska, P.; Leitch, A.R.; Sykorova, E.; Fajkus, J. Characterisation of an unusual telomere motif (TTTTTTAGGG)n in the plant Cestrum elegans (Solanaceae), a species with a large genome. Plant J. 2015, 82, 644–654.
19. Peska, V.; Sitova, Z.; Fajkus, P.; Fajkus, J. BAL31-NGS approach for identification of telomeres de novo in large genomes. Methods 2017, 114, 16–27.
20. Fajkus, J.; Sykorova, E.; Leitch, A.R. Telomeres in evolution and evolution of telomeres. Chromosome Res. 2005, 13, 469–479.
21. Fajkus, P.; Peska, V.; Zavodnik, M.; Fojtova, M.; Fulneckova, J.; Dobias, S.; Kilar, A.; Dvorackova, M.; Zachova, D.; Necasova, I.; et al. Telomerase RNAs in land plants. Nucleic Acids Res. 2019, 47, 9842–9856.

transposable elements and repeats was reported[41]. Moreover, activation of transposable elements was also observed as a stress response to genome instability that may have been caused by polyploidization event or environmental stress[41][42]. We presume, therefore, that the ancestral *TERT*_D locus (including *TERT*_12exD and *MtATPO*) originated as a result of gene/segment duplication of the *TERT*_C (plus *MtATPO*) locus or vice versa (Supplemental Text S1, Figure S2). Both loci were pseudogenized—the *TERT*_C locus within the *MtATPO* region and the *TERT*_D locus within the *TERT* region—and later, the *TERT*_D locus was split and translocated by Ogre/SD1-I. Currently, the mutual positioning of *TERT*_C, *TERT*_D and *TERT*_12exD within the genome of *N. sylvestris* is not known; however, similar scenarios could have resulted in pseudogenization and/or neofunctionalization of an additional *TERT* gene copy that we found in diploid species *Populus trichocarpa*, *Vigna radiata* and *Mimulus guttatus*. These *TERT-like* sequences may illustrate possible scenarios leading to the formation of *TERT* pseudogene variants in *Nicotiana* and the progression of gene elimination after gene/genome duplication: (i) A large-scale segment/genome duplication event had created an additional *TERT* locus, presumably encoding a *TERT* pseudogene on chromosome 1 in *Populus*, (ii) two *TERT* copies placed on the same scaffold in *Mimulus*. (iii) A completely different arrangement comprising an additional *TERT* variant of *Vigna radiata* that is formed by two adjacent inverted copies of exon 9 of *TERT*, and this *TERT*-like sequence was annotated as ncRNA (summarized in Figure S2). Multiple *TERT* copies were present in some, but not all polyploid species investigated, and toleration of more *TERT* copies after young polyploidization events is obvious (Figure 4).

Table 2. Number of *TERT* gene copies in *Nicotiana* species determined in silico.

Species/Genome Dataset Accession	Total No. of *TERT* Reads	Expected Genome Coverage (Depth)	No. of Detected *TERT* Variants	Read Counts Corresponding to Known *TERT* Variants			Ratio of *TERT* Variants in Genome
N. tabacum SRX338107	1259	35×	3	*NtTERT_Cs*	*NtTERT_D*	*NtTERT_Ct*	1:1:1
				425	424	410	
N. sylvestris ERX248848	644	26×	2	*NsTERT_C*	*NsTERT_D*		1:1
				332	312		
N. tomentosiformis ERX248865	203	15×	1	*NtomTERT*			-
				203			
N. benthamiana (raw data from [39])	286	20×	1	*NbenTERT*			-

Regarding the origin and evolution of the *TERT* loci in eudicots, comparison of eudicot phylogeny relationships [43] with the occurrence of syntenic loci that adopted the *TERT* gene demonstrated ancient synapomorphies, i.e., loci preserved in current genomes are assumed to have been present in their most recent common ancestor (nodes are depicted in Figure 4). The eudicot-like synteny locus emerged in early eudicots (*Amborella*) and adopted the *TERT* gene later in the ancestral parent of *Nelumbo*. The original *Amborella TERT* locus was probably fragmented. Another translocation of the *TERT* gene into novel loci grouped in all investigated malvids and Malpighiales (in fabids), and further translocations to other loci, took place later on. Interestingly, in several cases, we detected a translocation into loci that had already existed in ancestral genomes for a long time, e.g., the locus with the *Citrus*-like synteny originated in early eudicots, as assumed from the locus synapomorphy. The first *TERT* translocation from a locus with eudicot-type synteny was not caused by locus fragmentation because

22. Peska, V.; Mátl, M.; Mandáková, T.; Vitales, D.; Fajkus, P.; Fajkus, J.; Garcia, S. Human-like telomeres in Zostera marina reveal a mode of transition from the plant to the human telomeric sequences. J. Exp. Bot. 2020, 71, 5786–5798.
23. Sykorova, E.; Fajkus, J. Structure-Function relationships in telomerase genes. Biol. Cell 2009, 101, 375–392.
24. Belfort, M.; Curcio, M.J.; Lue, N.F. Telomerase and retrotransposons: Reverse transcriptases that shaped genomes. Proc. Natl. Acad. Sci. USA 2011, 108, 20304–20310.
25. Fitzgerald, M.S.; Riha, K.; Gao, F.; Ren, S.; McKnight, T.D.; Shippen, D.E. Disruption of the telomerase catalytic subunit gene from Arabidopsis inactivates telomerase and leads to a slow loss of telomeric DNA. Proc. Natl. Acad. Sci. USA 1999, 96, 14813–14818.
26. Oguchi, K.; Liu, H.; Tamura, K.; Takahashi, H. Molecular cloning and characterization of AtTERT, a telomerase reverse transcriptase homolog in Arabidopsis thaliana. FEBS Lett. 1999, 457, 465–469.
27. Harrington, L.; Zhou, W.; McPhail, T.; Oulton, R.; Yeung, D.S.; Mar, V.; Bass, M.B.; Robinson, M.O. Human telomerase contains evolutionarily conserved catalytic and structural subunits. Genes Dev. 1997, 11, 3109–3115.
28. Lingner, J.; Hughes, T.R.; Shevchenko, A.; Mann, M.; Lundblad, V.; Cech, T.R. Reverse
2 … Harley, … ence
3 … nes in
3 … lag … 017,
32. Leitch, I.J.; Hanson, L.; Lim, K.Y.; Kovarik, A.; Chase, M.W.; Clarkson, J.J.; Leitch, A.R. The ups and downs of genome size evolution in polyploid species of Nicotiana (Solanaceae). Ann. Bot. 2008, 101, 805–814.
33. Clarkson, J.J.; Lim, K.Y.; Kovarik, A.; Chase, M.W.; Knapp, S.; Leitch, A.R. Long-Term genome diploidization in allopolyploid Nicotiana section Repandae (Solanaceae). New Phytol. 2005, 168, 241–252.
34. Knapp, S.; Lughadha, E.N.; Paton, A. Taxonomic inflation, species concepts and global species lists. Trends Ecol. Evol. 2005, 20, 7–8.

35. Kelly, L.J.; Leitch, A.R.; Clarkson, J.J.; Knapp, S.; Chase, M.W. Reconstructing the complex evolutionary origin of wild allopolyploid tobaccos (Nicotiana section Suaveolentes). Evolution 2013, 67, 80–94.
36. Jureckova, J.F.; Sykorova, E.; Hafidh, S.; Honys, D.; Fajkus, J.; Fojtova, M. Tissue-Specific expression of telomerase reverse transcriptase gene variants in Nicotiana tabacum. Planta 2017, 245, 549–561.
37. Pfaffl, M.W. Quantification strategies in real-time PCR. In A-Z of Quantitative PCR; Bustin, S.A., Ed.; International University Line: La Jolla, CA, USA, 2004; pp. 87–112.
38. Clarkson, J.J.; Knapp, S.; Garcia, V.F.; Olmstead, R.G.; Leitch, A.R.; Chase, M.W. Phylogenetic relationships in Nicotiana (Solanaceae) inferred from multiple plastid DNA regions. Mol. Phylogenet. Evol. 2004, 33, 75–90.
39. Bombarely, A.; Rosli, H.G.; Vrebalov, J.; Moffett, P.; Mueller, L.A.; Martin, G.B. A draft genome sequence of Nicotiana benthamiana to enhance molecular plant-microbe biology research. Mol. Plant Microbe Interact. 2012, 25, 1523–1530.
40. Xie, M.; Podlevsky, J.D.; Qi, X.; Bley, C.J.; Chen, J.J. A novel motif in telomerase reverse transcriptase regulates telomere repeat addition rate and processivity. Nucleic Acids Res. 2010, 38, 1982–1996.
41. Renny-Byfield, S.; Chester, M.; Kovarik, A.; Le Comber, S.C.; Grandbastien, M.A.; Deloger, M.; Nichols, R.A.; Macas, J.; Novak, P.; Chase, M.W.; et al. Next generation sequencing reveals genome downsizing in allotetraploid Nicotiana tabacum, predominantly through the elimination of paternally derived repetitive DNAs. Mol. Biol. Evol. 2011, 28, 2843–2854.
42. Madlung, A.; Tyagi, A.P.; Watson, B.; Jiang, H.; Kagochi, T.; Doerge, R.W.; Martienssen, R.; Comai, L. Genomic changes in synthetic Arabidopsis polyploids. Plant J. 2005, 41, 221–230.
43. Byng, J.W.; Chase, M.W.; Christenhusz, M.J.M.; Fay, M.F.; Judd, W.S.; Mabberley, D.J.; Sennikov, A.N.; Soltis, D.E.; Soltis, P.S.; Stevens, P.F.; et al. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot. J. Linn. Soc. 2016, 181, 1–20.
44. Huang, S.; Ding, J.; Deng, D.; Tang, W.; Sun, H.; Liu, D.; Zhang, L.; Niu, X.; Zhang, X.; Meng, M.; et al. Draft genome of the kiwifruit Actinidia chinensis. Nat. Commun. 2013, 4, 2640.

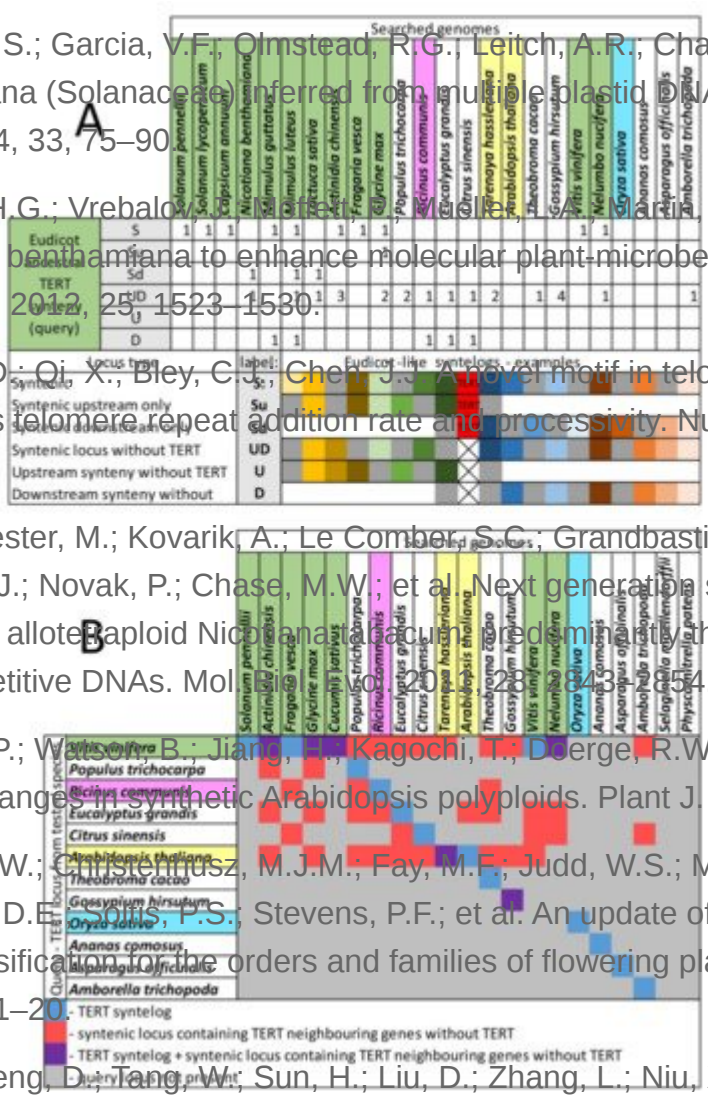

Figure 3. Occurrence of *TERT* syntelogs in Angiosperms. (**A**) Detailed analysis of the eudicot-like type of synteny (represented by *Vitis* syntelog as a query) in indicated genomes shows the presence of syntenic regions with/without the *TERT* gene in the majority of investigated eudicots and *Amborella*. The number of syntelogs and synteny categories are shown for each species. (**B**) The occurrence of conserved syntenic regions corresponding to the species-specific *TERT* query was investigated in representative genomes. Co-occurrence of syntelogs in more species suggests an ancient origin of target loci that accommodated *TERT* in current species. Analyses were

carried out using CoGe, GEvo and SynFind. SynFind parameters—algorithm: last; Gene window size: 30; minimum number of genes: 5.

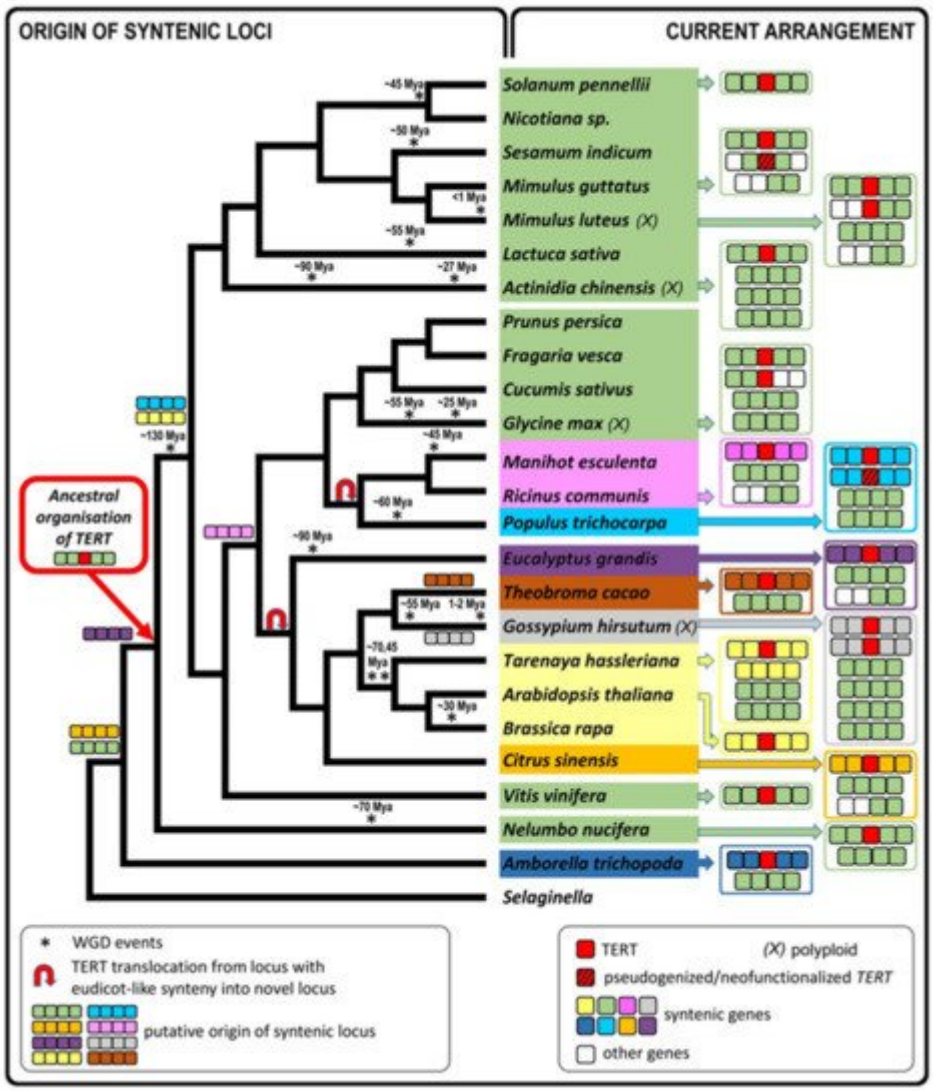

Figure 4. Origin of syntenic *TERT* loci in angiosperms. An ancestral locus with eudicot-like synteny that is present in the *Amborella* genome adopted the TERT gene in basal eudicots (*Nelumbo*). This ancestral locus with eudicot-like synteny occurs within the genomes of eudicots with the exception of the model plant *Arabidopsis thaliana* (for simplicity, diagrams on right panels show occurrence and arrangement of loci with eudicot-type synteny and with a specific synteny if present). The TERT gene was translocated several times into novel loci with a conserved synteny observed in current species (termed here as *Citrus*-like, *Populus*-like, *Eucalyptus*-like, *Ricinus*-like, *Arabidopsis*-like syntelogs) that had already occurred in ancestors (nodes depicting synapomorphic relationships of specific synteny-types and thus the putative origin of ancestral syntenic loci are shown above respective phylogeny nodes). As an exception, Malpighiales (*Theobroma*, *Gossypium*) show the TERT gene translocated into novel species-specific loci. These genomes nevertheless still contain the ancient loci with conserved synteny (details in Figure 3). *TERT* is mostly maintained as a single copy gene, but polyploid species can tolerate more copies (*M. luteus*, *G. max*, *G. hirsutum* are shown as representatives). Copies of genomic loci with the original synteny remain present after *TERT* gene elimination, e.g., in *Actinidia chinensis*, where it is difficult

to distinguish which of the ohnologous loci (ohnologs = paralogs derived by WGD) have lost their TERT gene copy (see Supplemental Text S1, Figure S2). Phylogeny was adapted from APG IV [43], WGDs were mapped according to [7] in eudicots, and according to [44] in *Actinidia*.

In conclusion, our results show that natural *Nicotiana* polyploids tolerate more *TERT* copies and, similarly to other polyploid genomes investigated, retention of various copies is obvious in species formed by young polyploidization events. A comparison of *TERT* locus arrangement in current genomes suggests that the *TERT* gene was placed in a conserved locus that became advantageous following the emergence of basal eudicots (Figure 4). The gene was relocated later in several plant groups where only a narrow syntenic relationship restricted to closely related species could be found. Various evolutionary scenarios took place in ancestral genomes with multiple *TERT* copies resulting in elimination, pseudogenization and/or fragmentation, and neofunctionalization of novel *TERT* copies that could also illustrate the origin and fate of *N. sylvestris* and polyploid *Nicotiana TERT* variants (Figure S2).