Chapter 11
Evolutionary Implications of Genome and Karyotype Restructuring in *Nicotiana tabacum* L.

Ales Kovarik, Simon Renny-Byfield, Marie-Angèle Grandbastien and Andrew Leitch

Abstract *Nicotiana tabacum* is an allopolyploid that formed within the last 200,000 years from relatives of the extant diploids *N. sylvestris* and *N. tomentosiformis*, the donors of the S- and T-genomes, respectively. Here we review progress in our understanding of the divergence of *N. tabacum* subsequent to its formation, by comparing the *N. tabacum* genome with those of its diploid progenitors. We also review the data from synthetic *N. tabacum*, where there is evidence for much genetic change in early generations, including various chromosomal translocations, allopolyploid-induced retroelement mobility and loss, and reductions in the copy numbers of some tandem repeats. These observations are similar to patterns found in natural *N. tabacum*, suggesting that rapid genetic divergence is induced by allopolyploidy. The T-genome of *N. tabacum* shows the greatest number of genetic changes and appears to be less stable than the S-genome. We describe possible mechanisms that may have stimulated these genetic changes and propose that these can lead to enhanced fertility, more regular chromosome pairing, and the evolution of disomic inheritance.

A. Kovarik
Institute of Biophysics, Brno, Czech Republic

S. Renny-Byfield · A. Leitch (✉)
School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
e-mail: a.r.leitch@qmul.ac.uk

M.-A. Grandbastien
Institute Jean-Pierre Bourgin, Versailles, France

P. S. Soltis and D. E. Soltis (eds.), *Polyploidy and Genome Evolution*,
DOI: 10.1007/978-3-642-31442-1_11,

11.1 Introduction to Nicotiana

The genus *Nicotiana* is the sixth largest genus in the family *Solanaceae* and includes 76 species found naturally in the Americas, Australia and surrounding islands, and with a single species in Namibia (Goodspeed 1954; Knapp et al. 2004). However, several species, including *N. tabacum* (tobacco), now occur more widely. The genus has received much recent attention because of evidence for recurrent polyploidy, resulting in nearly half the species in the genus being chromosomally polyploid. It is likely there have been six independent polyploidy events within the genus: at about 10 million years ago (mya) (section *Suaveolentes*, *c*. 24 species), 5 mya (section *Repandae*, 4 species), 1–2 mya (section *Polydicleae*, 2 species), and within the last 200,000 years (*N. rustica*, *N. arentsii*, and *N. tabacum*) (Lim et al. 2004a). There is also reticulation at the diploid level, with at least three species likely to be of homoploid hybrid origin (*N. glauca*, *N. linearis*, and *N. spegazzinii*) (Goodspeed 1954; Kelly et al. 2010).

N. tabacum is an autotetraploid (2n = 4x = 48) derived from diploid species that most closely resemble *N. sylvestris* (2n = 2x = 24), the maternal S-genome donor, and *N. tomentosiformis* (2n = 2x = 24, section *Tomentosae*), the paternal T-genome donor. *Nicotiana sylvestris* is the only species in section *Sylvestres*, and detailed genetic analysis of multiple accessions indicates that there is little genetic diversity in this species (4.2 % polymorphisms in AFLPs—Amplified Fragment Length Polymorphisms, Petit et al. 2007). However, there is considerably more genetic diversity among accessions of *N. tomentosiformis* (31.7 % AFLP polymorphism), which form two distinct groups, one of which most closely resembles the T-genome of *N. tabacum* (Murad et al. 2002; Petit et al. 2007). These data indicate that *N. tabacum* formed subsequent to the divergence of the two groups of *N. tomentosiformis*.

Molecular clock estimates of plastid and internally transcribed spacer sequences (ITS) of nuclear ribosomal DNA (rDNA) suggest that *N. tabacum* formed less than 200,000 years ago. These estimates were calibrated using likely maximum ages of some endemic species occurring on islands of known geological ages (Clarkson et al. 2005; Leitch et al. 2008). However, given that only feral populations of *N. tabacum* have been found and no truly "wild" population exists (Knapp personal communication), it is quite possible that *N. tabacum* was born subsequent to the origin of human agriculture, within the last 10,000 years.

The purpose of this paper is to review current understanding of the genetic consequences of polyploidy in *N. tabacum*. We show that genetic change can occur rapidly within a few generations, perhaps as a consequence of the "genomic shock" of allopolyploidy (McClintock 1984), a process that generates variants from which selection favors those with enhanced fertility. This unstable phase is presumably transient; however, there remains evidence for genome dynamism over time scales of thousands of years, including the replacement of multiple copies of rDNA sequences, the turnover of retroelements, and the loss of sequences targeted at the T-genome. We discuss how such changes may promote fertility and lead to the fixation of rearranged karyotypes during polyploid species establishment.

11.2 *Nicotiana tabacum* Genome Structure

Kenton et al. (1993) were the first to use genomic in situ hybridization (GISH) to study the chromosomes of *N. tabacum*. The fluorescent probes derived from total genomic DNA of *N. sylvestris* and *N. tomentosiformis* hybridized to a separate subset of 24 chromosomes, corresponding to the chromosomes of the S- and T-genomes, respectively. However, the efficacy of the T-genome labeling was inferior (Lim et al. 2000b), likely because of a loss of T-genome sequences (Renny-Byfield et al. 2011). GISH also revealed 4-9 intergenomic translocations in all *N. tabacum* accessions (Kenton et al. 1993; Lim et al. 2004a; Moscone et al. 1996). Similar translocations (up to 3) were also observed in some synthetic *N. tabacum* lines that were only a few generations old (Skalicka et al. 2005). Perhaps these translocations have arisen as a consequence of multivalent formation (see Chap. 7), or they may represent hotspots of recombination. While some translocations are fixed (i.e. occur in all varieties), others are specific for particular accessions. There are more translocations of S-genome origin chromatin to T-genome chromosomes (T/s chromosome) than the reverse (S/t chromosome). Kenton et al. (1993) hypothesized that this could be caused by selection against S/t chromosomes. Alternatively, some of the "translocations" may actually be S-genome subtelomeric satellite sequences that now occur on T-genome chromosomes, perhaps arising in their new location via recombination-based homogenization processes (Koukalova et al. 2010). Next-generation sequencing (NGS) may enable us to distinguish among these alternative hypotheses.

As with most other plant species (Heslop-Harrison and Schwarzacher 2011), a large fraction of *Nicotiana* genomes is composed of various types of retroelements. Genome sampling using NGS revealed that the major component of the genome of *N. tabacum* and its diploid relatives comprises long terminal repeat (LTR)-retrotransposons, with at least 17.1–22.5 % of the genome being *Ty3/gypsy* elements and 2.2–3.4 % being *Ty1/copia* elements. DNA transposons comprise around 1.7 % of the genome of these species (Renny-Byfield et al. 2011).

In addition, there are several unrelated satellite repeat families that have been characterized (the distributions of some are shown Fig. 11.1, Lim et al. 2000b). These repeats include: (1) The HRS60 family, which is the best-characterized tandem repeat family in *Nicotiana* and includes an interstitial repeat in *N. tomentosiformis* and *N. tabacum* (called GRS, Gazdova et al. 1995) and predominantly subtelomeric repeats in *N. sylvestris* and *N. tabacum* (called HRS60 and NSYL2, Koukalova et al. 2010; Koukalova et al. 1989); (2) NTS9 in *N. sylvestris* and *N. tabacum* (Jakowitsch et al. 1998); (3) NTRS in *N. tomentosiformis* and *N. tabacum* (Fig. 11.1, Matyasek et al. 1997), and (4) *Nic*CL3, a long 2.2 kb tandem repeat comprising ~2 % of the *N. tomentosiformis* genome and in lower abundance in *N. tabacum* (Fig. 11.2, Renny-Byfield et al. 2011 and Renny-Byfield et al. 2012). There are also a number of satellite repeat families of known origin; these are: (5) A1/A2 satellite repeats derived from the intergenic spacer (IGS) of rDNA, which have transposed and amplified to multiple locations across the genome (Lim et al. 2004b); (6) tandem repeats of

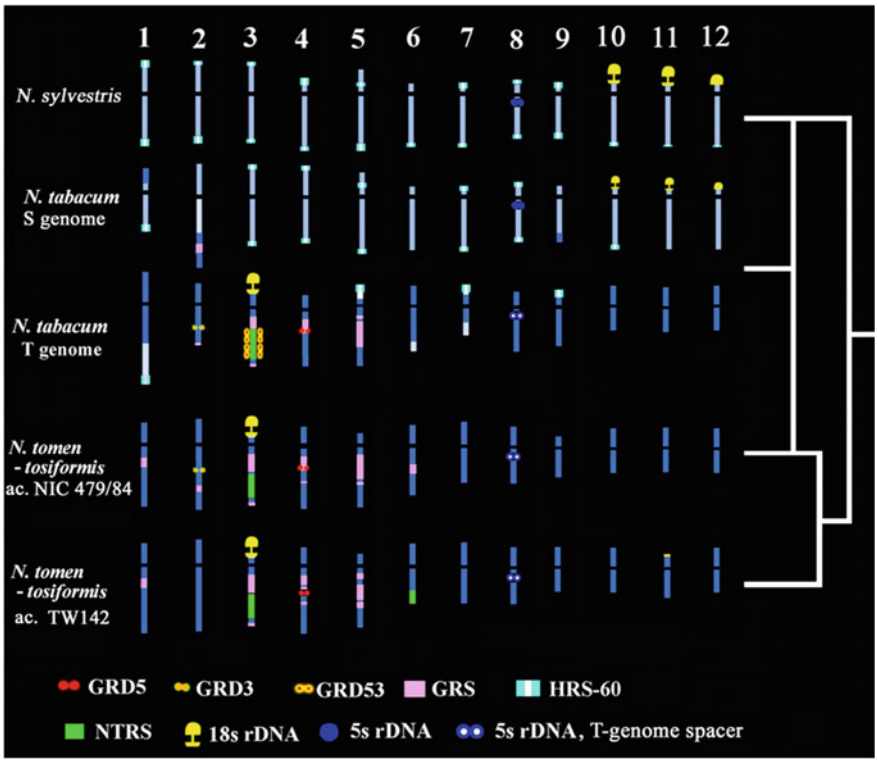

Fig. 11.1 Ideograms showing the distribution of tandem repeat sequences in *N. tabacum* and its diploid progenitors, *N. sylvestris* and *N. tomentosiformis*. Taken with permission from Murad et al. (2002). See text for a description of the repeats shown

geminivirus-related DNA (GRD), occurring in two distinct families in some accessions of *N. tomentosiformis* and inherited in *N. tabacum* (Murad et al. 2004; Murad et al. 2002), the first endogenous viruses discovered in plants (Bejarano et al. 1996); (7) sequences of pararetroviral origin (Matzke et al. 2004), including distinct variants found in *N. tomentosiformis* (NtoEPRV) and *N. sylvestris* (NsEPRV), both of which have been inherited in *N. tabacum*.

11.3 Retroelement Response to Allopolyploid "Genomic Shock"

Nicotiana hybrids were among the first in which chromosomal changes following hybridization were demonstrated. Gerstel and Burns (1967) reported that *N. otophora* ×*N. tabacum* hybrids have genetic instabilities of two kinds. Firstly, the heterochromatin from *N. otophora* undergoes breakage causing chromatin loss; this

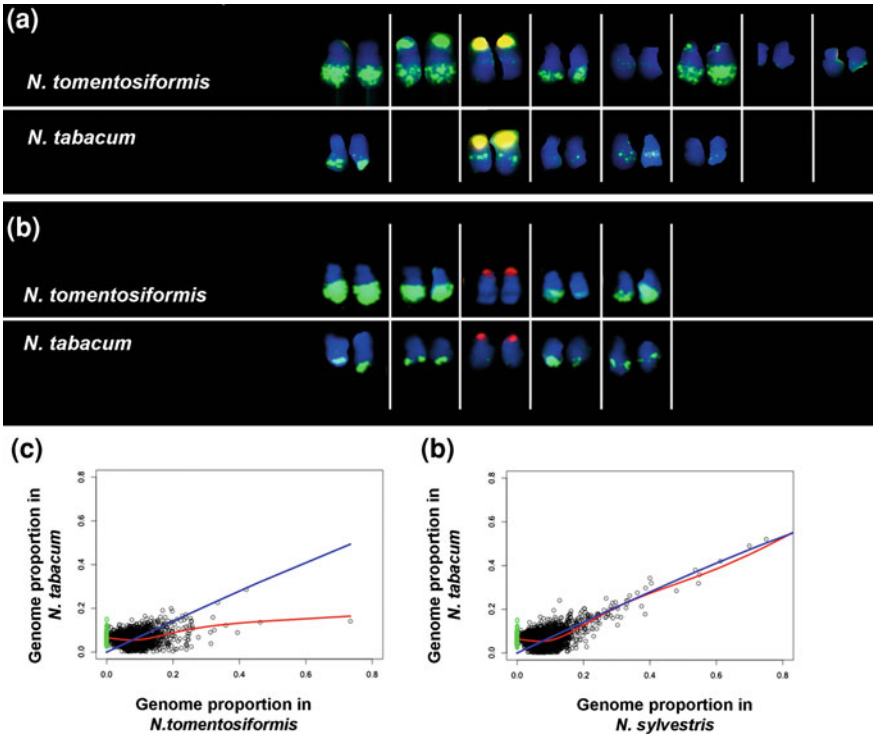

Fig. 11.2 Reduction in copy number of T-genome repeats in *N. tabacum* from expectation given their abundance in *N. tomentosiformis*. **a** Fluorescence in situ hybridization (FISH) showing the distribution of dispersed A1/A2 repeats (*green fluorescence*). The 18-5.8-26S rDNA locus on chromosome 3 (*N. tomentosiformis*) and T3 (*N. tabacum*) is labeled with *yellow fluorescence* due to the overlap of *red* and *green* rDNA fluorescence. Note that there is much reduced A1/A2 signal in *N. tabacum* (for further information see Lim et al. 2004b). **b** FISH showing the distribution of *Nic*CL3 in both species (*green fluorescence*). Note the reduced abundance of this repeat in *N. tabacum*. The 18-5.8-26S rDNA locus on chromosome 3 (*N. tomentosiformis*) and T3 (*N. tabacum*) is labeled with *red fluorescence*. For further information see Renny-Byfield et al. (2012). **c**, **d**. NGS analysis of repeat clusters in *N. tabacum* and its diploid progenitors. The graphs show the genome proportions (arcsine transformed) of repeats uniquely inherited in *N. tabacum* from **c** *N. tomentosiformis* and **d** *N. sylvestris*. If *N. tabacum* had faithfully inherited the repeats as found in the diploids, then all repeats would fall on the *blue lines*. However, in *N. tomentosiformis* **c** most repeats at higher genome proportions are underrepresented in *N. tabacum*. This is not the case for *N. sylvestris* **d**, where repeats are more or less in their expected abundance. This trend is reflected in the abundance of repeats that are biparentally inherited, i.e., the more abundant they are in *N. tomentosiformis*, the more likely they will be underrepresented in *N. tabacum*. For further information see Renny-Byfield et al. (2011)

can be observed cytologically and also phenotypically by the leaf variegation it causes. Secondly, in some cells heterochromatic blocks from *N. otophora* proliferate enormously to many times their normal length, forming "megachromosomes". This pioneering work revealed that interspecific hybridization can induce chromosomal rearrangements and rapid sequence losses and gains.

Melayah et al. (2001) were the first to show, using *N. tabacum*, that environmental stresses can induce retrotransposition. There is also evidence that allopolyploidy can stimulate transposition in *N. tabacum* and synthetic mimics of *N. tabacum* made from the diploid progenitors (Petit et al. 2007, 2010). These authors analyzed four populations of Ty1/copia LTR-retrotransposons (Tnt1-ol16, Tnt1-ol13, Tnt2, and Tto1) using sequence-specific amplification polymorphism (SSAP). They showed that the parental diploid species share essentially similar classes of retroelements. In *N. tabacum* there is evidence of retrotransposon diversification subsequent to allopolyploidy, with sequence losses concomitant with gains. Losses of retroelements were more frequent from the T-genome, while novel insertions of some populations, such as Tnt2, were shown to preferentially locate to the S-genome. However, each retrotransposon population seems to behave differently, with some populations undergoing rapid turnover, while others display relative stasis. There was little or no diversification of retrotransposons among individual accessions of natural *N. tabacum*, indicating that retroelement diversification occurred before the accessions analyzed had diverged, perhaps early in the species'evolution and arising as a consequence of allopolyploidy and induced "genomic shock" (McClintock 1984).

To test the hypothesis that allopolyploidy can induce retroelement diversification, Petit et al. (2010) compared insertion patterns of the Tnt1 family in the synthetic *N. tabacum* Th37 (Burk 1973) with its diploid progenitors. In some Th37 individuals in the fourth synthetic generation (S4), there was evidence of new Tnt1 insertions. Newly, transposed copies were amplified from elements located on the *N. sylvestris*-derived genome and were highly similar to the Tnt1A tobacco copies amplified in response to microbial factors (Grandbastien et al. 2005). Furthermore, a high proportion of parental SSAP bands was lost in Th37, particularly those from the *N. tomentosiformis*-derived genome, again as observed in natural *N. tabacum*. Together, these data indicate that retrotransposon amplification and molecular restructuring in, or around, insertion sites occur rapidly in response to allopolyploidy, and that similar sequences have responded in the same way to allopolyploidy in natural and synthetic *N. tabacum*.

11.4 Loss of DNA: Targeting the T-Genome

As with many other allopolyploid species (Leitch and Bennett 2004), *N. tabacum* has a lower genome size than is expected from the sum of parental genomes (Leitch et al. 2008, *N. tabacum* genome size, 1C = 5110 Mbp; *N. sylvestris*, 1C = 2636 Mbp; *N. tomentosiformis*, 1C = 2683 Mbp (http://data.kew.org/cvalues/). Recently, (Renny-Byfield et al. 2011) used NGS analysis of *N. tabacum* and its two diploid progenitors and showed that repeats derived from *N. tomentosiformis* was underrepresented in *N. tabacum*, a trend that was not observed for repeats from *N. sylvestris* (Fig. 11.2). These observations lead to the conclusion that the T-genome of

N. tabacum has undergone extensive sequence losses, and that genetic changes have occurred more rapidly to the paternally derived T-genome than the S-genome.

N. tabacum harbors two families of endogenous pararetrovirus DNA, one from each progenitor (Jakowitsch et al. 1999). These are NsEPRV from *N. sylvestris*, which in *N. tabacum* is at the expected abundance, and NtoEPRV from *N. tomentosiformis*, which occurs in *N. tabacum* with reduced copy number. Likewise, the A1/A2 repeats (Lim et al. 2004b) and *Nic*CL3 (Renny-Byfield et al. 2012), which are found in multiple locations around the genome in *N. tomentosiformis* (Fig. 11.2), are also underrepresented in *N. tabacum*. Similarly, NGS data show that all families of *Ty3/gypsy* elements have undergone sequence loss in *N. tabacum* (Renny-Byfield et al. 2011), and SSAP analysis of Tnt retroelements (*Ty1/copia*) reveals more deletion events influencing loci from the T-genome than the S-genome of *N. tabacum* (Petit et al. 2007). There is evidence of rapid and targeted loss of repeats from the T-genome even from the earliest generations, since in synthetic *N. tabacum* lines, many individuals already have reduced copy numbers of NtoEPRV, NTRS repeats (Skalicka et al. 2005), and *Nic*CL3 (Renny-Byfield et al. 2012), whilst most losses of Tnt1 retroelements were from the T-genome (Petit et al. 2010). All these data suggest rapid, targeted changes early in allopolyploid evolution. The targeting of repeat losses from the T-genome may support the nuclear-cytoplasmic interaction hypothesis, whereby the paternally derived genome functions in a maternal cytoplasm (Gill 1991; Leitch et al. 2006), which in allopolyploids may lead to incompatibilities that preferentially destabilize the paternally derived genome.

11.5 Ribosomal DNA Homogenization is Rapid and Ongoing

In most eukaryotes, 5S and 18–5.8–26S rDNA units occur in tandem arrays at one or several loci. Each large rDNA unit contains the 18S, 5.8S, and 26S ribosomal RNA (rRNA) genes, ITS, and IGS sequences. Whilst these sequences are vital for cell functioning, they can also be highly recombinogenic and labile sequences influencing genome stability (Kobayashi 2011). The genes themselves are highly conserved; however, their spacers diverge at suitable rates for resolving species relationships within most genera, including *Nicotiana* (Chase et al. 2003). Early studies revealed that the restriction fragment length polymorphism (RFLP) patterns of the *N. tabacum* IGS is not additive of the diploid progenitors (Kovarik et al. 1996), but that *N. tabacum* has evolved its own distinct rRNA gene family(ies). Sequence analysis revealed that the tobacco-specific units arose by reorganization of *N. tomentosiformis*-inherited units followed by their subsequent amplification (Volkov et al. 1999). The sequence changes mainly involved amplification and reduction of subrepeats upstream and downstream of the transcription start site within the IGS. However, the newly evolved units still occur at the four rDNA loci that *N. tabacum* inherited from its parents (see Fig. 11.1). Thus, it is likely that the parental units of S-genome origin were overwritten by the newly amplified

N.tabacum-specific units (Lim et al. 2000a), although a few units of S genome-origin remain intact, perhaps because they are methylated and inactive (Kovarik et al. 2008). In contrast, the 5S rDNA follows a different evolutionary pattern in *N. tabacum*, with no evidence for intergenomic homogenization, although shifts in copy numbers of parental gene families have been observed (Fulnecek et al. 2002).

11.6 The Fate of Duplicated Genes

Although there has been substantial restructuring of the repetitive fraction of the *N. tabacum* genome since its formation, most genic sequences analyzed have remained in duplicate copies, e.g., genes for drug resistance (Schenke et al. 2003), putrescine N-methytransferases (Riechers and Timko 1999), a family of small GTP-binding proteins (Takumi et al. 2002), lignin forming peroxidase (Matassi et al. 1991), nitrite reductase (Kronenberger et al. 1993), nitrate reductase (Vaucheret et al. 1989), glutamine synthetase (Clarkson et al. 2010), phytochrome A (Intrieri et al. 2008), the developmental gene LEAFY (McCarthy 2010), and families of DNA methyltransferases (Fulnecek et al. 2009). The only current exception is a family of *N. tabacum* glucan endo-1,3-beta-glucosidase genes that appeared to be recombinants of both ancestral sequences (Sperisen et al. 1991). Locus additivity does not always seem to be a rule even in recently formed synthetic allopolyploids. For example, frequent deletions of homoeologs were found among 70 protein-coding loci in *Tragopogon miscellus* that formed within the last 80 years (Buggs et al. 2009, 2012; see also Chap. 14, this volume). These changes may not be random, as Buggs et al. (2012) showed that clusters of genes are repeatedly lost or retained and the likelihood of retention reflects gene ontology categories or their predicted levels of dosage sensitivity. Similarly, synthetic lines of *Brassica napus* seem to eliminate much of the parental DNA, although it is currently unknown whether coding or noncoding sequences are preferentially targeted (Song et al. 1995; Szadkowski et al. 2011). The occurrence of both homoeologs in *N. tabacum* may reflect limited sampling, or high levels of retention. Additivity is observed in synthetic lines of *Gossypium* (Liu et al. 2002; see also Chap. 10, this volume) and in synthetic wheat allopolyploids, although in the latter case there remains some controversy (Mestiri et al. 2010; Feldman and Levy 2009; see also Chap. 7, this volume). Additivity has also been demonstrated at the expression level. Alleles of phytochrome A, lignin forming peroxidases, nitrite reductase, and DNA methyltransferases are expressed from both parental homeologues. There are also reports of epigenetic silencing of one subset of the parental alleles, including rRNA genes that have escaped homogenization (Kovarik et al. 2008) and the CYP82E4 locus (Chakrabarti et al. 2007) involved in the alkaloid biosynthesis pathway.

Duplicate copies arising through polyploidy can be retained through a number of mechanisms (Doyle et al. 2008): (1) one copy can evolve a new function (neofunctionalization) that can become fixed through a selective advantage

(Lynch et al. 2001); (2) duplicate copies can diverge through complementary loss of function at a particular point in the development or in particular tissues (subfunctionalization), a process that can occur through the action of drift alone; and (3) selection can occur against gene loss because that loss would compromise appropriate levels of gene product in relation to another gene(s) (gene balance hypothesis, cf. Birchler and Veitia 2010). In *N. tabacum*, a mechanism termed "nonfunctionalization" has been proposed which considers a combination of degenerative mutation and epigenetic silencing (Chakrabarti et al. 2007). Here, "non-functionalization" has altered alkaloid metabolism through reduced conversion of nicotine to nornicotine. Perhaps this mutation has been important to the "success" of *N. tabacum* as a recreational drug.

11.7 Genome Revolution and Allopolyploid Establishment

Until the early 1990s, it was commonly thought that chromosomal changes establish isolation barriers among populations because of reduced fertility in heterozygotes. Rearranged karyotypes may, then, become fixed through inbreeding or meiotic drive, particularly in small populations (King 1993). However, it is difficult to distinguish between karyotype changes that drive speciation events and those arising by genetic drift after speciation. Certainly, many rearrangements are deleterious and can cause embryo lethality or hybrid breakdown prior to reproduction (Rieseberg 1997). Only neutral or advantageous recombinants will pass the bottleneck of selection during the early phase of speciation. Furthermore, evolutionary geneticists pointed out that genetic mutations at the DNA level are more common than karyotype change and are therefore likely to be more important in speciation processes (Butlin 1993). However, recent evidence from a number of sources, e.g., *Anopheles gambiae* (Turner and Hahn 2010) and *Helianthus* homoploid hybrids (Strasburg et al. 2009), indicates that chromosomal inversions can, indeed, have a role in establishing barriers between populations and drive speciation. Strasburg et al. (2009) suggested that inversions may become foci for the accumulation of adaptive genes, particularly at the junctions between collinear and inverted regions, where gene flow is likely to be most impeded. Likewise, computer modeling has shown that when recombination is eliminated at an inverted region, and in the presence of strong selection and minimal or no gene flow, then species isolation can be driven by such a rearrangement (Feder and Nosil 2009).

In the context of early polyploid evolution, karyotype rearrangement may drive enhanced fertility and potentially be fundamental to allopolyploid species establishment. Certainly, high levels of chromosomal change have been observed in young polyploids, e.g., karyotype variability (dysploidy and intergenomic translocations) within and among populations of *Tragopogon* polyploids that formed within the last 80 years (Chester et al. 2010, 2012; Lim et al. 2008), and large-scale chromosomal deletions occurring in individual plants of early generation

synthetic allopolyploids of *B. napus* (Gaeta et al. 2007). Typically, a major hurdle that a newly formed polyploid must overcome is reduced fertility, often arising through multivalent formation, where chromosomes can segregate aberrantly leading to aneuploidy and reduction in fertility. We can envisage several allopolyploid-induced processes that may act to reduce multivalent formation and be favored by selection, these are:

- *Reduced chiasma frequency*. Selection may favor the formation of fewer and/or focussed chiasma at recombination hotspots. This would result in quadrivalents formed in prophase I falling into two bivalents in metaphase I and then segregating normally (Fig. 11.3a). Computer modeling has shown that reduced chiasma frequency results in the evolution of disomic inheritance (Le Comber et al. 2010), although we are unaware of empirical evidence to support this assertion.
- *Structural and epigenetic divergence of homeologues*. Allopolyploidy can trigger extensive karyotype and molecular restructuring in early generations (Gaeta et al. 2007; Renny-Byfield et al. 2011). This may result in genetic divergence of the parental genomes, e.g., through the loss or rearrangement of sequences shared by homeologues and recognized by the homolog recognition machinery in meiosis (Fig. 11.3b). Such sequences are unlikely to be those that are highly repeated across the genome since they are not chromosome specific. Nevertheless, genetic changes may introduce regions among homeologues that do not recombine, forming "islands of divergence", as in the establishment of homoploid hybrid species (Strasburg et al. 2009). Without recombination, the local regions around these islands will continue to diverge, leading to regular homolog pairing, and bivalent formation. Similarly, allopolyploidy can trigger epigenetic changes across the genome (Parisod et al. 2009). Epigenetic changes, which are thought to alter patterns of ectopic recombination (Colot et al. 1995), may also influence patterns of meiotic recombination and initiate islands of divergence. In addition, newly formed epi-alleles may establish tissue or temporal patterns of gene expression (subfunctionalization, Adams and Wendel 2005). Subfunctionalization is likely to favor the evolution of regular disomic inheritance and to select against multivalent formation. This is because multivalents result in the inappropriate segregation of homeoalleles and the generation of individuals with reduced fitness, because they have aberrant patterns of gene expression (Le Comber et al. 2010).
- *Intergenomic translocations*. In the context of a diploid, translocations among heterologues can promote quadrivalent formation, leading to chromosome loss, and reduced fertility. Similarly, in the context of polyploids, translocations and aneuploidy are associated with low fertility, as observed in synthetic *B. napus* lines (Xiong et al. 2011). However, it is possible that a more complex dynamic can occur (Fig. 11.3c). If homeologues already form multivalents, then it is conceivable that the gain of genetic material from a heterologous chromosome may promote bivalent formation, disomic inheritance, and increased fertility. Alternatively, complex rearrangements can result through cascades of induced

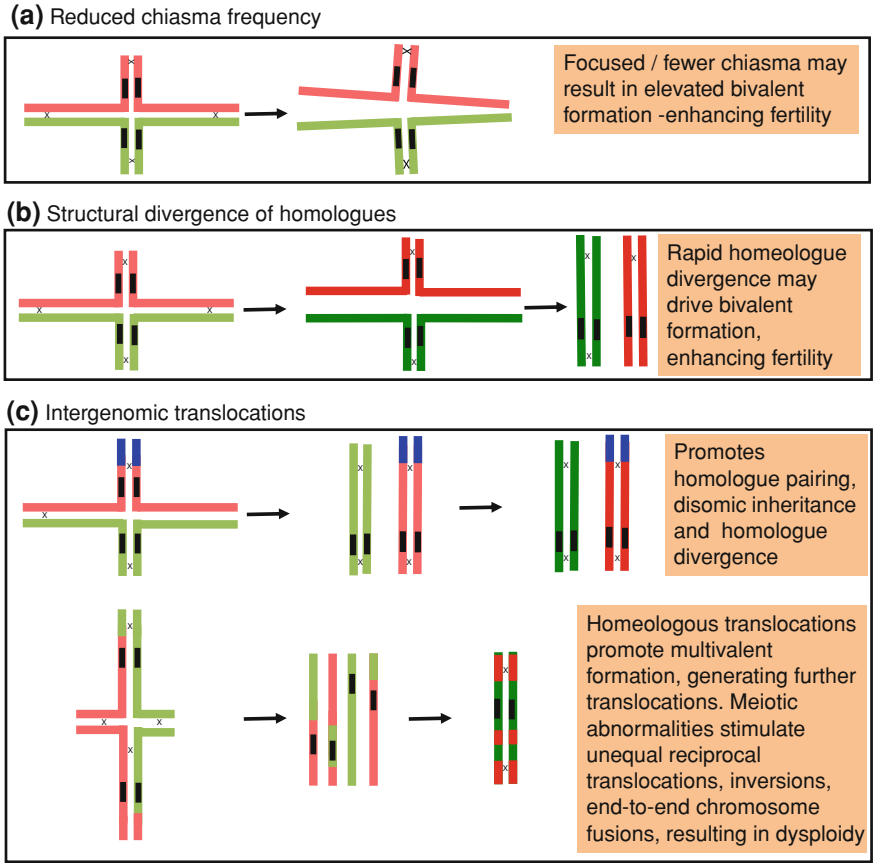

Fig. 11. 3 Fast molecular and chromosome evolution may enhance bivalent formation, driving disomic inheritance, and fertility in young polyploids. Homologs are shown in the same color (*pink* or *pale green*). Homeologues diverge over time, this is reflected in their colors becoming more distinct (*red* and *green*). A translocation from a heterologue is shown in *blue*. Crossovers among chromosomes are shown with crosses

pairing problems resulting from recurrent multivalent formation (unequal reciprocal translocations, inversions, end-to-end chromosome fusions), potentially generating novel chromosomes with segments from multiple ancestral chromosomes. Such a process may be associated with a reduction in chromosome number to a diploid-like level, as reported in the divergence of *Arabidopsis* karyotypes (Mandakova et al. 2010).

From *Nicotiana* allopolyploids, we have evidence for several of the proposed modes of chromosome evolution. First, in synthetic *N. tabacum*, we have observed plants that were homozygous for an intergenomic translocation that is similar to translocations observed in natural *N. tabacum* (Lim et al. 2004a). Perhaps this translocation is selected because it enhances fertility, and that in the synthetic

N. tabacum evolution is repeating itself. Second, the allotetraploids of sections *Polydicliae* and *Repandae*, which formed *c*. 1 mya and 5 mya, respectively, have chromosome numbers that are additive of the progenitor diploids ($2n = 4x = 48$). However, the sequence organization along their chromosomes has diverged considerably, particularly in the latter. Indeed, Lim et al. (2007) suggested that in some species of section *Repandae* there has been near complete genome turnover. Perhaps these changes contributed to the establishment of disomic inheritance. Finally, in section *Suaveolentes*, several species show evidence of chromosome number reduction from expectation (chromosome numbers range from $2n = 4x = 32$ to 46, depending on species, Goodspeed 1954; Knapp et al. 2004). We anticipate that this reduction is associated with karyotype restructuring, as in the polyploids of *Arabidopsis* (Mandakova et al. 2010).

11.8 Advantages of the *Nicotiana* System and Future Perspectives

Nicotiana provides many opportunities for studying allopolyploid genome divergence and addressing key questions facing polyploidy researchers [outlined in Soltis et al. (2010)]. It is important, even vital for the future of humankind, to know how allopolyploid genomes interact together. This is because many of our most important crop species are recognizably polyploid based on their chromosome numbers, and recent evidence is emerging that all seed-bearing plants have undergone at least one round of polyploidy in their ancestry (Jiao et al. 2011). The allopolyploid species of *Nicotiana* and the synthetic allopolyploids provide a unique opportunity to study snap shots of polyploid divergence over 10 million years of evolution.

Acknowledgments We thank NERC and the Czech Science Foundation for support.

References

Adams KL, Wendel JF (2005) Allele-specific, bidirectional silencing of an alcohol dehydrogenase gene in different organs of interspecific diploid cotton hybrids. Genetics 171:2139–2142

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 Nat Acad Sci USA 93:759–764

Birchler JA, Veitia RA (2010) The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution. New Phytol 186:54–62

Buggs RJ, Doust AN, Tate JA, Koh J, Soltis K, Feltus FA, Paterson AH, Soltis PS, Soltis DE (2009) Gene loss and silencing in *Tragopogon miscellus* (*Asteraceae*): comparison of natural and synthetic allotetraploids. Heredity 103:73–81

Buggs RJA, Chamala S, Wu W, Tate JA, Schnable PS, Soltis DE, Soltis PS, Barbazuk WB (2012) Rapid, repeated, and clustered loss of duplicate genes in allopolyploid plant populations of independent origin. Curr Biol 22:248–252

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

Butlin RK (1993) Species evolution—the role of chromosome change—King, M. Nature 366:27

Chakrabarti M, Meekins KM, Gavilano LB, Siminszky B (2007) Inactivation of the cytochrome P450 gene CYP82E2 by degenerative mutations was a key event in the evolution of the alkaloid profile of modern tobacco. New Phytol 175:565–574

Chase MW, Knapp S, Cox AV, Clarkson JJ, Butsko Y, Joseph J, Savolainen V, Parokonny AS (2003) Molecular systematics, GISH and the origin of hybrid taxa in *Nicotiana* (*Solanaceae*). Ann Bot 92:107–127

Chester M, Gallagher JP, Symonds VV, Cruz da Silva AV, Mavrodiev EV, Leitch AR, Soltis PS and Soltis DE (2012) Extensive chromosomal variation in a recently formed natural allopolyploid species, Tragopogon miscellus (Asteraceae). Proceedings of the National Academy of Sciences of the United States of America, 109:1176-1181

Chester M, Sykorova E, Fajkus J, Leitch AR (2010) Single integration and spread of a Copia-like sequence nested in rDNA intergenic spacers of *Allium cernuum* (*Alliaceae*). Cytogenet Genome Res 129:35–46

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 55:99–112

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 168:241–252

Colot V, Goyon C, Faugeron G, Rossignol JL (1995) Methylation of repeated DNA sequences and genome stability in Ascobolus immersus. Can J Bot-Rev Canadienne De Botanique 73:S221–S225

Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF (2008) Evolutionary genetics of genome merger and doubling in plants. Ann Rev Genet 443–461

Feder JL, Nosil P (2009) Chromosomal inversions and species differences: when are geens affecting adaptive divergence and reproductive isolation expected to reside within inversions? Evolution 63:3061–3075

Feldman M, Levy AA (2009) Genome evolution in allopolyploid wheat-a revolutionary reprogramming followed by gradual changes. J Genet Genomics 36:511–518

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

Fulnecek J, Matyasek R, Kovarik A (2009) Faithful inheritance of cytosine methylation patterns in repeated sequences of the allotetraploid tobacco correlates with the expression of DNA methyltransferase gene families from both parental genomes. Mol Genet Genomics 281:407–420

Gaeta RT, Pires JC, Iniguez-Luy F, Leon E, Osborn TC (2007) Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. Plant Cell 19:3403–3417

Gazdova B, Siroky J, Fajkus J, Brzobohaty B, Kenton A, Parokonny A, Heslop-Harrison JS, Palme K, Bezdek M (1995) Characterization of a new family of tobacco highly repetitive DNA, GRS, specific for the *Nicotiana tomentosiformis* genomic component. Chromosome Res 3:245–254

Gerstel DU, Burns JA (1967) Phenotypic and chromosomal abnormalities associated with the introduction of heterochromatin from *Nicotiana otophora* into *N. tabacum*. Genetics 56:483–502

Gill BS (1991) Nucleocytoplasmic interaction (NCI) hypothesis of genome evolution and speciation in polyploid plants. In: Sasakuma T, KinoshitaT (ed.) Proceedings of the Kihara

memorial international symposium on cytoplasmic engineering in wheat, Kihara Memorial Foundation, Yokohama, Japan, pp 48–53
Goodspeed TH (1954) The genus *Nicotiana* Massachusetts. Chronica Botanica Company, USA
Grandbastien M, Audeon C, Bonnivard E, Casacuberta JM, Chalhoub B, Costa APP, Le QH, Melayah D, Petit M, Poncet C, Tam SM, Van Sluys MA, Mhiri C (2005) Stress activation and genomic impact of Tnt1 retrotransposons in *Solanaceae*. Cytogenetic Genome Res 110:229–241
Heslop-Harrison JS, Schwarzacher T (2011) Organisation of the plant genome in chromosomes. Plant J 66:18–33
Intrieri MC, Muleo R, Buiatti M (2008) *Phytochrome* A as a functional marker of phyletic relationships in *Nicotiana* genus. Biol Plant 52:36–41
Jakowitsch J, Mette MF, van der Winden J, Matzke MA, Matzke AJM (1999) Integrated pararetroviral sequences define a unique class of dispersed repetitive DNA in plants. Proc Nat Acad Sci USA 96:13241–13246
Jakowitsch J, Papp I, Matzke MA, Matzke AJM (1998) Identification of a new family of highly repetitive DNA, NTS9, that is located predominantly on the S9 chromosome of tobacco. Chromosome Res 6:649–659
Jiao YN, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang HY, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphilis CW (2011) Ancestral polyploidy in seed plants and angiosperms. Nature 473:97–113
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 27:781–799
Kenton A, Parokonny AS, Gleba YY, Bennett MD (1993) Characterization of the *Nicotiana tabacum* L. genome by molecular cytogenetics. Mol Gen Genet 240:159–169
King M (1993) Species evolution the role of chromosome change. Cambridge University Press, Cambridge
Knapp S, Chase MW, Clarkson JJ (2004) Nomenclatural changes and a new sectional classification in *Nicotiana* (*Solanaceae*). Taxon 53:73–82
Kobayashi T (2011) Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. Cell Mol Life Sci 68:1395–1403
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 186:148–160
Koukalova B, Reich J, Matyasek R, Kuhrova V, Bezdek M (1989) A BamHI family of highly repeated DNA sequecnes of *Nicotiana tabacum*. Theor Appl Genet 78:77–80
Kovarik A, Dadejova M, Lim K, Chase M, JJ C, Knapp S, Leitch A (2008) Evolution of rDNA in Nicotiana allopolyploids: a potential link between rDNA homogenization and epigentics. Ann Bot 101:815–823
Kovarik A, Fajkus J, Koukalova B, Bezdek M (1996) Species-specific evolution of telomeric and rDNA repeats in the tobacco composite genome. Theor Appl Genet 92:1108–1111
Kronenberger J, Lepingle A, Caboche M, Vaucheret H (1993) Cloning and expession of distinct nitrite reductases in tobacco-leaves and roots. Mol Gen Genet 236:203–208
Le Comber SC, Ainouche ML, Kovarik A, Leitch AR (2010) Making a functional diploid: from polysomic to disomic inheritance. New Phytol 186:113–122
Leitch AR, Lim KY, Skalicka K, Kovarik A (2006) Nuclear cytoplasmic interaction hypothesis and the role of translocations in *Nicotiana allopolyploids*. In: Cigna AAD, M. Yerevan (eds.) Radiation risk estimates in normal and emergency situations book series: NATO security through science series B: Physics and Biophysics (ARMENIA) pp 319–326
Leitch IJ, Bennett MD (2004) Genome downsizing in polyploid plants. Biol J Linn 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 101:805–814

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 109:161–172

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 175:756–763

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

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

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 166:1935–1946

Lim KY, Soltis DE, Soltis PS, Tate J, Matyasek R, Srubarova H, Kovarik A, Pires JC, Xiong ZY, Leitch AR (2008) Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (*Asteraceae*). Plos One, 3

Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF (2002) Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome 44:321–330

Lynch M, O'Hely M, Walsh B, Force A (2001) The probability of preservation of a newly arisen gene duplicate. Genetics 159:1789–1804

Mandakova T, Joly S, Krzywinski M, Mummenhoff K, Lysak MA (2010) Fast diploidization in close mesopolyploid relatives of *Arabidopsis*. Plant Cell 22:2277–2290

Matassi G, Melis R, Macaya G, Bernardi G (1991) Compositional bimodality of the nuclear genome of tobacco. Nucleic Acids Res 19:5561–5567

Matyasek R, Gazdova 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 paragretroviruses of polyploid *Nicotiana tabacum* and its diploid progenitors, *N. sylvestris* and *N. tomentosiformis*. Biol J Linn Soc 82:627–638

McCarthy EW (2010) The effect of molecular and spectral reflectance evolution on *Nicotiana* polyploids of different ages. In School of Biological and Chemical Sciences. London: Ph.D. thesis from Queen Mary University of London, p 259

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

Melayah D, Bonnivard E, Chalhoub B, Audeon C, Grandbastien MA (2001) The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. Plant Journal, 28:159-168

Mestiri I, Chague V, Tanguy AM, Huneau C, Huteau V, Belcram H, Coriton O, Chalhoub B, Jahier J (2010) Newly synthesized wheat allohexaploids display progenitor-dependent meiotic stability and aneuploidy but structural genomic additivity. New Phytol 186:101–186

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

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 92:352–358

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 89:921–928

Parisod C, Salmon A, Zerjal T, Tenaillon M, Grandbastien MA, Ainouche M (2009) Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in *Spartina*. New Phytol 184:1003–1015

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

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

Renny-Byfield S, Chester M, Kovarik A, Le Comber S, Grandbastien M, Deloger M, Nichols R, Macas J, Novak P, Chase M, Leitch A (2011) Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs Molecular Biology and Evolution 28:2843–2854

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

Riechers DE, Timko MP (1999) Structure and expression of the gene family encoding putrescine N-methyltransferase in *Nicotiana tabacum*: new clues to the evolutionary origin of cultivated tobacco. Plant Mol Biol 41:387–401

Rieseberg LH (1997) Hybrid origins of plant species. Annu Rev Ecol Syst 28:359–389

Schenke D, Sasabe M, Toyoda K, Inagaki Y, Shiraishi T, Ichinose Y (2003) Genomic structure of the NtPDR1 gene, harboring the two miniature inverted-repeat transposable elements, NtToya1 and NtStowaway101. Genes Genet Syst 78:409–418

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 166:291–303

Soltis DE, Buggs RJA, Doyle JJ, Soltis PS (2010) What we still don't know about polyploidy. Taxon 59:1387–1403

Song KM, Lu P, Tang KL and Osborn TC (1995) Rapid genome change in synthetic polyploids of *Brassica* and its implications for polyploid evolution. Proceedings of the National Academy of Sciences of the United States of America, 92:7719-7723

Sperisen C, Ryals J, Meins F (1991) Comparison of cloned genes provides evidence for intergenomic exchange of DNA in the evolution of a tobacco glucan endo-1,3 beta-glucosidase gene family. Proc Nat Acad Sci USA 88:1820–1824

Strasburg JL, Scotti-Saintagne C, Scotti I, Lai Z, Rieseberg LH (2009) Genomic patterns of adaptive divergence between chromosomally differentiated sunflower species. Mol Biol Evol 26:1341–1355

Szadkowski E, Eber F, Huteau V, Lode M, Coriton O, Jenczewski E and Chevre AM (2011) Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus*. New Phytologist, 191:884-894

Takumi S, Ida M, Haisa Y, Ando S, Nakamura C (2002) Genomic structure and homoeologous relationship of the two alpha-subunit genes of a heterotrimeric GTP-binding protein in tobacco. Genome 45:626–633

Turner TL, Hahn MW (2010) Genomic islands of speciation or genomic islands and speciation? Mol Ecol 19:848–850

Vaucheret H, Vincentz M, Kronenberger J, Caboche M, Rouze P (1989) Molecular cloning and characterization of the 2 homologous genes coding for nitrate reductase in tobacco. Mol Gen Genet 216:10–15

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 16:311–320

Xiong ZY, Gaeta RT, Pires JC (2011) Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. Proc Nat Acad Sci USA 108:7908–7913