Differences in the processing of DNA ends in *Arabidopsis thaliana* and tobacco: possible implications for genome evolution

Nadiya Orel and Holger Puchta*
*Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstrasse 3, 06466 Gatersleben, Germany (*author for correspondence; e-mail puchta@bio.uka.de)*

Key words: C value paradox, deletion, DNA recombination, NHEJ, plant genome size

Abstract

Surprising species-specific differences in non-homologous end-joining (NHEJ) of genomic double-strand breaks (DSBs) have been reported for the two dicotyledonous plants *Arabidopsis thaliana* and *Nicotiana tabacum*. In *Arabidopsis* deletions were, on average, larger than in tobacco and not associated with insertions. To establish the molecular basis of the phenomenon we analysed the fate of free DNA ends in both plant species by biolistic transformation of leaf tissue with linearized plasmid molecules. Southern blotting indicated that, irrespective of the nature of the ends (blunt, 5′ or 3′ overhangs), linearized full-length DNA molecules were, on average, more stable in tobacco than in *Arabidopsis.* The relative expression of a β-glucuronidase gene encoded by the plasmid was similar in both plant species when the break was distant from the marker gene. However, if a DSB was introduced between the promoter and the open reading frame of the marker, transient expression was halved in *Arabidopsis* as compared to tobacco. These results indicate that free DNA ends are more stable in tobacco than in *Arabidopsis*, either due to lower DNA exonuclease activity or due to a better protection of DNA break ends or both. Exonucleolytic degradation of DNA ends might be a driving force in the evolution of genome size as the *Arabidopsis* genome is more than twenty times smaller than the tobacco genome.

Introduction

The elucidation of the complete sequence of the *Arabidopsis* genome revealed multiple sequence duplications, although the genome consists of only 125 Mb and is thus one of the smallest plant genomes known. This hints to the possibility of a genome duplication in the evolution of *Arabidopsis* (Arabidopsis Genome Initiative). Thus, ample amounts of DNA must have been lost after the postulated duplication event, resulting in the small size of today's *Arabidopsis* genome. Recently, programmed sequence elimination in the first generations after allopolypolydization of wheat could be demonstrated (Ozkan *et al.*, 2001; Shakled *et al.*, 2001), but further DNA elimination over longer evolutionary time periods has to be postulated to explain the amount of sequence loss (Petrov, 2001). Indications for species-specific deletions were first found in insects. Studies on non-coding retrotransposon-like sequences of insects indicated considerable DNA loss from the *Drosophila* genome (Petrov, 1996). Indeed, the rate of deletions differed drastically between different insect genomes. Cricket, which has a 11-fold larger genome than *Drosophila* has a 40-fold lower rate of DNA loss than *Drosophila*. This applies for the number as well as the size of the deletions. Based on this finding, an inverse correlation between genome size and deletions size was postulated (Petrov *et al.* 2000).

But, what kind of process might be responsible for these deletions? Deletions may occur by different mechanisms: by replication slippage (as suggested by Capy, 2000), by unequal crossover (as suggested by Smith, 1976) or by double-strand break (DSB) repair. Recently we compared DSB-induced deletions that result in a loss of function of a chromosomal marker gene in two related dicotyledonous plant species *Arabidopsis* and tobacco (Kirik *et al.*, 2000), which differ

by more than 20 times in genome size (Bennett and Leitch, 1997). Whereas the mechanism of junction formation itself was not different between the two species, the size classes of the deletions in *Arabidopsis* and tobacco differed remarkably. On average the deletions were one third larger in *Arabidopsis* than in tobacco (Kirik *et al.*, 2000). This parallels with theoretical studies on insects sustaining the hypothesis of an inverse correlation between the genome size and the average length of deletions (Petrov *et al.*, 2000). Interestingly, we were not able to find insertions associated with deletions in *Arabidopsis*, whereas this was the case in almost half of the repair events analysed in tobacco (Salomon and Puchta, 1998).

What kind of enzymatic difference between the two plant species could be responsible for such a phenomenon? A possible explanation is that the efficiency of the degradation of DNA ends might differ between the two species. This could be either due to a lower exonuclease activity or due to a better protection of DNA ends from degradation in tobacco. To test this hypothesis we performed comparative analysis of the stability of linearized plasmid molecules transferred by particle bombardement into *Arabidopsis* and tobacco cells. Southern blot analysis in combination with transient expression studies with β-glucuronidase as marker gene indicated that plasmid DNA is exonucleolytically degraded to a higher extent in *Arabidopsis* than in tobacco. The small genome size of *Arabidopsis* might therefore result from reduced stability of free DNA ends in this plant species over evolutionary time periods.

Material and methods

Cloning procedures and DNA preparation

The plasmid pGUS23 is a pUC7 derivative that carries the β-glucuronidase ORF under the control of the 35S promoter (Puchta and Hohn, 1991). A specific polylinker was inserted into the *Acc65*I site of pGUS23 between promoter and ORF of the plasmid by using the annealed oligonucleotides 5′-GTACAGCATGCGGGCCCTTAATTAACCGCGGA GGCCTGTTAACATCGATACCGGTTCTAGA-3′ and 5′-GTACTTCTAGAACCGGTATCGATGTTAACAG GCCTCCGCGGTTAATTAAGGGCCCGCATGC-3′. From the resulting clone pGUS231 plasmid DNA was prepared with the Qiagen plasmid Maxi Kit (Qiagen, Hilden, Germany) and linearized by digestion with one of the following restriction enzymes: *Xba*I, *Stu*I, *Sma*I, *Aat*II and *Sca*I (all NEB, Beverly, MA). Completion of digestion was tested by retransformation of the linearized plasmid DNA into *Escherichia coli*. For transformation experiments only plasmid preparations were used in which the number of colonies was decreased by more than 99% as compared to the unrestricted plasmid.

DNA preparation and biolistic plant transformation

Arabidopsis thaliana (ecotype Columbia-O) and *Nicotiana tabacum* (SR1) plants were grown under standard conditions on GM medium in Sigma boxes. Plant leaves (one for tobacco and eight or nine for *Arabidopsis*) were arranged to form a 3 cm diameter circle in the centre of a petri dish containing callus induction medium (CIM) and incubated for 24 h before bombardment in a growth chamber under standard conditions (Seki *et al.*, 1999).

5 μg gold particles (1 μm diameter, BioRad, Hercules, CA) were kept overnight in 1 ml absolute ethanol. After washing two times in sterile water, the particles were resuspended in 100 μl of 50% glycerol. A mixture required for four bombardments was prepared as follows: 25 μl of gold particle suspension was mixed with 10 μl of plasmid DNA and 35 μl of 1 M $Ca(NO_2)_3$. After 10 min at room temperature and brief centrifugation, the supernatant was discarded and the DNA-coated gold was re-suspended in 100 μl of absolute ethanol. Per macro carrier disc 25 μl of the plasmid/gold suspension were transferred. For each transformation about 200 ng plasmid DNA was used. After complete evaporation of the ethanol from the disc, biolistic transformation was performed. Leaves were bombarded by means of the Biolistic PDS1000-He Particle Delivery System (BioRad) with a pressure of 7.6 MPa and a gap size of 0.25″.

Plant DNA extraction and Southern analysis

Two days after transformation, total DNA was isolated from *Arabidopsis* and tobacco leaves as described (Salomon and Puchta, 1998). For Southern analysis, 5 μg of genomic DNA was digested with *Hin*dIII overnight, fractionated on a 1% agarose gel and transferred onto a nitrocellulose hybridization membrane Hybond N (Amersham, UK). Plasmid DNA of pGUS231 was labelled with a random priming labelling kit (Amersham, UK) and [α-P^{32}] dATP (Amersham, UK). Hybridization was performed according to the instructions at 65 °C.

Histological staining for GUS expression

Two days after transformation, *Arabidopsis* and tobacco leaves were infiltrated with β-glucuronidase staining solution (0.1 M sodium phosphate buffer pH 7.0 with 0.05% X-Glu (Duchefa, Haarlem, Netherlands) dissolved in dimethyl formamide, in the presence of 0.1% sodium azide) and incubated at 37 °C for two days. Blue sectors on the leaves were counted using a binocular microscope. Only those experiments were taken into account for the analysis in which a certain minimum number of sectors could be detected (at least 10 per cm^2 of tobacco leaf).

Results

Rationale

To test the hypothesis that differences found in DSB repair between *Arabidopsis* and tobacco are due to differences in the processing of DNA ends, we decided to test the stability of plasmid DNA transformed into leaf cells of both organisms. For this purpose, we inserted between promoter and ORF of the β-glucuronidase gene of the plasmid pGUS23 (Puchta and Hohn, 1991) a polylinker that contained recognition sites for multiple restriction enzymes including *Xba*I and *Stu*I, resulting in the plasmid pGUS231 (Figure 1). Prior to plant transformation, the plasmid was linearized with specific restriction enzymes. Five different restriction enzymes were used: *Sca*I, resulting in linearized plasmid DNA with blunt ends, as well as *Aat*II, resulting in linearized DNA with 5′ overhangs, both enzymes cut in the pUC backbone of pGUS231; *Xba*I resulting in a linearized DNA molecule with 3′ overhangs and *Stu*I resulting in a linearized DNA molecule with blunt ends, both enzymes cut between 35S promoter and β-glucuronidase ORF; and *Sma*I resulting in a linearized DNA molecule with blunt ends, which enzyme cuts at the start of the β-glucuronidase ORF (for details see Figure 1). Particle bombardement was used to transfer the DNA into tobacco and *Arabidopsis* leaves (Seki *et al.*, 1999). To avoid variation due to differences in the DNA preparations, in all cases the same batch of the DNA-gold preparation was used for transformation of both species. For this purpose leaves of 8-week old tobacco seedlings and leaves of 6-week old *Arabidopsis* seedlings, both in a similar phase of their development, were cultured for one day on CIM medium and then bombarded in parallel with gold particles coated with DNA. Next, the leaves were

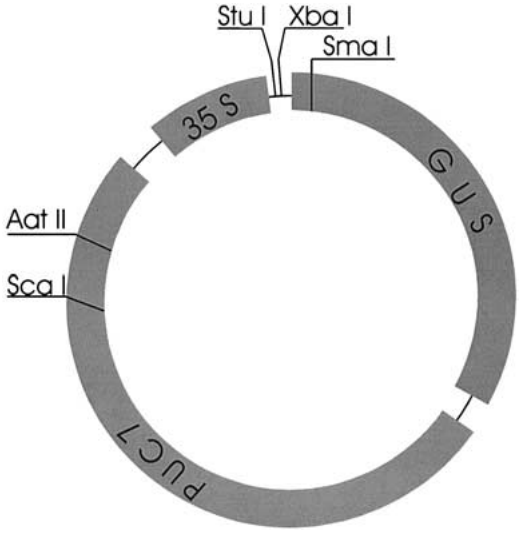

Figure 1. The plasmid pGUS231. The promoter (35S) and open reading frame (GUS) of the β-glucuronidase expression cassette are indicated as well as the recognition sites of the restriction enzymes applied for linearization of the plasmid in this study.

cultivated for two days on CIM medium before further analysis on the stability of the transferred DNA either by Southern blotting or by transient expression studies was performed.

Southern blot analysis

In a first set of experiments plasmid stability was analysed by Southern blotting. To elucidate how different types of breaks influence the stability of a DNA molecule in plant cells, pGUS231 with 3′-(*Xba*I), 5′-(*Ata*II) overhangs or blunt ends (*Sma*I, *Stu*I) was prepared. After the incubation period the leaves were harvested and total DNA was extracted and digested with the restriction enzyme *Hin*dIII, which has no recognition sequence within pGUS231. After gel electrophoresis, Southern blotting was performed and the respective membranes were hybridized with a plasmid-specific probe under highly stringent conditions. The obtained radioactive signals on the blot were quantified by the use of a phosphoimager. The quotient of signal intensity of the linearized plasmid DNA (bracket a in Figure 2) and that of the smear produced by degraded plasmid DNA, in the same lane (bracket b in Figure 2) was calculated. In pilot experiments different incubation times of the leaves after bombardment were tested. After 8, 24 and 48 h, besides degraded DNA, the linearized plasmid band was clearly visible for both plant species, whereas after 72 h the resulting band became too weak for reliable

quantification. For our further analysis we therefore chose the longest possible incubation period for which a reliable quantification could be performed (2 days). In general, the intensity of the linear band was between 1% and 5% of the total intensity of the signal in the respective lane. These values obtained for tobacco and *Arabidopsis* were then set into relation. For example, if in a particular experiment the linearized band represented 3% of total signal intensity in tobacco and 2% of total signal intensity in *Arabidopsis*, a ratio of 1.5 was established. Any value over 1 thus indicates that the full-length linear DNA was more stable in tobacco than in *Arabidopsis*. The higher the values the more linear plasmid molecules remained undegraded in tobacco as compared to *Arabidopsis*. A value of 2 would indicate that twice as many linear plasmid molecules remained undegraded in tobacco than in *Arabidopsis*. A series of experiments was performed using differently restricted plasmids. For each restriction enzyme nine independent experiments are depicted in Table 1. In spite of quite large variation between single experiments, which are typical of transient transformation assays (e.g. Puchta and Hohn, 1991), our results clearly demonstrate that, independent of the respective DNA end, in most cases (29 out of 36) the linearized DNA was less degraded in tobacco than in *Arabidopsis* (Table 1). Linearized plasmid DNA was, on average, 1.4–1.9 times more stable in tobacco than in *Arabidopsis*. No major differences between blunt and overhanging DNA ends could be detected (Table 1). However, due the fact that the subtraction of the standard deviations from the means, given in Table 1, result in a value above 1 in only two of four cases, these results can only be taken as a first indication of differences in processing of DNA molecules in the two plant species and needed to be further corroborated by transient expression analysis.

Transient expression

To sustain the results described above and to further characterize the mode of nuclear degradation transient expression experiments were performed. A difference in DNA degradation could be due either to differences of exonucleolytic or endonucleolytic processing of the transfected DNA. To clarify the nature of nucleolytic activity we performed transient expression assays after biolistic bombardement of *Arabidopsis* and tobacco with linearized pGUS231. The rationale was that in case of endonucleolytic degradation, no difference in the relation of the expression rates of a marker

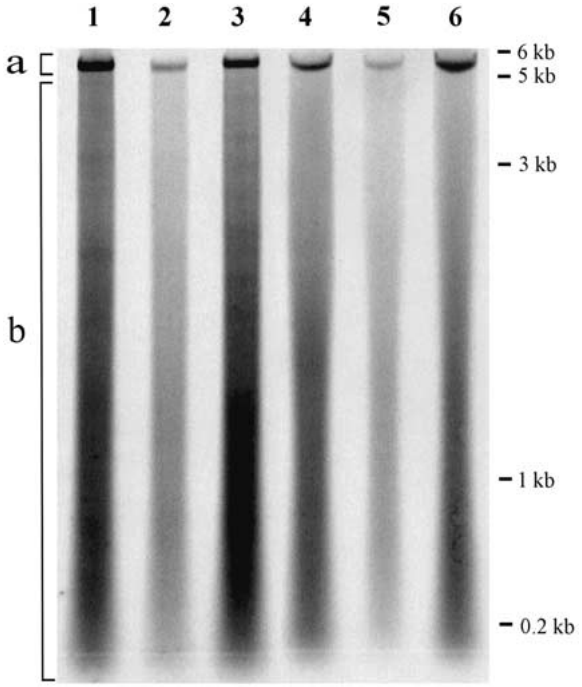

Figure 2. Representative Southern analysis of DNA extracted from leaves two days after biolistic transformation with linearized DNA of the plasmid pGUS231. After isolation the DNA was restricted with *Hin*dIII, blotted and hybridized with labelled pGUS231. The intensity of the signals was determined by the use of a phosophoimager. The linear band (a) was put into relation with the smear resulting from degraded DNA (b). In lanes 1–3 *Arabidopsis* DNA and in lanes 4–6 tobacco DNA was loaded. Lanes 1, 4: DNA of plant leaves bombarded with pGUS231 restricted with *Xba*I; lanes 2, 5: DNA of plant leaves bombarded with pGUS231 restricted with *Sma*I; lanes 3, 6: DNA of plant leaves bombarded with pGUS231 restricted with *Ata*II.

Table 1. The relative stability of linearized DNA molecules of the plasmid pGUS231 in leaf cells of tobacco in relation to *Arabidopsis.* In the bottom line the means and standard deviations are given.

*Xba*I 3′ overhang	*Stu*I blunt	*Sma*I blunt	*Ata*II 5′ overhang
2.48	2.25	0.96	2.64
1.88	0.97	1.32	0.63
1.70	2.80	2.17	0.97
1.28	0.94	1.00	1.95
1.52	1.17	1.64	1.74
1.92	2.63	2.15	1.48
0.63	3.15	1.54	1.86
1.46	1.55	1.18	1.23
1.54	1.34	1.75	0.46
1.60 ± 0.50 (9)	1.91 ± 0.92 (9)	1.52 ± 0.45 (9)	1.42 ± 0.66 (9)

gene should be detectable between *Arabidopsis* and tobacco, irrespective where the break was induced in the plasmid molecule. However in case of exonucleolytic degradation, different positions of the break resulting in a linear plasmid molecule are expected to reveal differences between the two species. The closer the break site is to a DNA sequence required for expression the higher the efficiency of marker destruction by exonucleolitic degradation resulting in a reduced expression activity is expected to be. Therefore we compared the *Xba*I and *Stu*I restriction sites of pGUS231 lying between promoter and ORF of the β-glucuronidase gene with the *Sca*I site located within the pUC backbone (Figure 1). Whereas in case of *Xba*I and *Stu*I sites exonucleolytic degradation of less than 100 bp at either end would result in a loss of function of the β-glucuronidase gene, in case of the *Sca*I site only degradation of more than 1 kb at one end or more than 2 kb at the other would be required to result in a loss of function of the enzyme. Since *Stu*I as well as *Sca*I digestion results in blunt ends, not the nature of the break but its position differs between the respective DNA molecules. Conversely, *Xba*I-digested ($3'$ overhanging ends) and *Stu*I-digested (blunt ends) DNA molecules differ in respect to the DNA ends but hardly in the position of the break site. However, for the latter DNAs a religation step (Gorbunova and Levy, 1997) to fuse the promoter to the ORF is required for transcription of a functional mRNA. Bombardement with circular plasmid DNA was included as transcription control. After histochemical staining and detection of blue sectors the average number of spots per cm^2 leaf was determined. The data are presented in Table 2. In general independent of plant species the highest numbers of spots were detected after shooting with circular pGUS231, followed by the plasmid cut outside of the marker gene. Religation of pGUS231 harbouring sticky ends occurred quite efficiently, depending on experiment in the range of 10–30% of the molecules the ORF was restored, whereas religation of the blunt-ended plasmid gave in most experiments lower rates. All values were above 1%, the putative maximal 'background' value, that could arise due to traces of undigested plasmid DNA in the preparation, as determined by retransformation of the respective linearized plasmid preparations into *E. coli*. To directly compare the results obtained from tobacco and *Arabidopsis* the values obtained with the linearized plasmid molecules were set into relation to the circular ones (Table 2, second last row). By this procedure variations between tobacco and *Arabidopsis* due to transcriptional differences were excluded from the calculation. These values were then set into relation to one another (Table 2, last row). For example, a value of 2 indicates the presence of twice as many functional β-glucuronidase genes in tobacco than in *Arabidopsis* after bombardement with the respective DNA. In repeated experiments (for statistical evaluation, see Table 3) this relation was found to be around 1 in case of the *Sca*I-digested pGUS231 and around 2 for both *Xba*I- and *Stu*I-digested pGUS231. This indicates that in case of a break between promoter and ORF, on average, twice more functional β-glucuronidase expression cassettes were restored in tobacco than in *Arabidopsis* irrespective of the nature of the DNA ends. The calculated standard deviations indicate that the detected differences are highly significant.

Discussion

Double-strand breaks (DSBs) are critical lesions in genomes. Efficient repair of DSBs is necessary for the survival of all organisms. In principle, DSBs can be repaired via non-homologous end joining (NHEJ) or via homologous recombination. Although homologous recombination is used for DSB repair in plants (Chiurazzi *et al.*, 1996; Shalev and Levy, 1997; Puchta, 1999; Xiao and Peterson, 2000; Siebert and Puchta, 2002) NHEJ seems to be the main mode of repair (for reviews, see Puchta and Hohn, 1996; Vergunst and Hooykaas, 1999; Gorbunova and Levy, 1999; Mengiste and Paszkowski, 1999). Error-prone DSB repair may result either in deletions, insertions or various other kinds of genomic rearrangements (Pipiras *et al.*, 1998; Salomon and Puchta, 1998; Jasin, 2000). In plants genomic alterations in somatic cells of meristems can be transferred to the offspring (Walbot, 1996) and therefore relevant for evolution. Thus, somatic DSB repair might have in evolutionary terms an impact on genome size and genome organization.

The reason for the large differences in the nuclear DNA content of eukaryotes, particularly plants, known as 'C value paradox', has been a matter of debate since long (Cavalier-Smith, 1985; Dove and Flavell, 1988). Even closely related species with similar phenotype may significantly differ as to their diploid genome size. One mechanism responsible for these differences could be due to a species-specific increase/reduction of repetitive sequences. In principle, genomes may become larger via duplications and insertions or smaller via deletions. Species-specific

Table 2. Expression of β-glucuronidase after bombardment of the plasmid pGUS231 in leaf cells of tobacco and *Arabidopsis*.

Experiment no.	Plasmid	Species	Spots per cm^2	Relation to circular pGUS231	Relation tobacco/*Arabidopsis*
1	circular	tobacco	260		
		Arabidopsis	44,2		
	*Xba*I	tobacco	70	0.27	1.40
		Arabidopsis	8.4	0.19	
	*Sca*I	tobacco	240	0.92	1.03
		Arabidopsis	39.4	0.89	
2	circular	tobacco	500		
		Arabidopsis	131		
	*Xba*I	tobacco	122.2	0.24	2.00
		Arabidopsis	15.2	0.12	
	*Sca*I	tobacco	290	0.58	1.14
		Arabidopsis	67.4	0.51	
3	circular	tobacco	157.5		
		Arabidopsis	52.7		
	*Xba*I	tobacco	24.6	0.16	1.60
		Arabidopsis	5.4	0.10	
	*Sca*I	tobacco	73	0.46	0.50
		Arabidopsis	48.5	0.92	
4	circular	tobacco	103		
		Arabidopsis	40		
	*Xba*I	tobacco	30.3	0.29	1.38
		Arabidopsis	8.5	0.21	
	*Sca*I	tobacco	86.4	0.83	0.71
		Arabidopsis	47	1.17	
5	circular	tobacco	185.7		
		Arabidopsis	34.7		
	*Xba*I	tobacco	28.1	0.15	2.10
		Arabidopsis	2.29	0.07	
6	circular	tobacco	260		
		Arabidopsis	44.2		
	*Xba*I	tobacco	73	0.28	1.50
		Arabidopsis	8.4	0.19	
	*Sca*I	tobacco	120	0.46	0.50
		Arabidopsis	39.4	0.89	
7	circular	tobacco	292.8		
		Arabidopsis	132.7		
	*Sca*I	tobacco	158	0.54	1.23
		Arabidopsis	58.6	0.44	
	*Stu*I	tobacco	25	0.085	2.43
		Arabidopsis	4.6	0.035	
8	circular	tobacco	67.7		
		Arabidopsis	9.3		
	*Stu*I	tobacco	19.3	0.29	2.4
		Arabidopsis	1.17	0.12	
9	circular	tobacco	368.5		
		Arabidopsis	72		
	*Stu*I	tobacco	15.5	0.040	2.35
		Arabidopsis	1.2	0.017	
10	circular	tobacco	185.7		
		Arabidopsis	34.7		
	*Stu*I	tobacco	28.1	0.152	2.15
		Arabidopsis	2.29	0.0706	
11	circular	tobacco	258		
		Arabidopsis	55		
	*Stu*I	tobacco	168.5	0.654	2.18
		Arabidopsis	17	0.300	
12	circular	tobacco	280		
		Arabidopsis	120		
	*Stu*I	tobacco	11.5	0.0410	1.97
		Arabidopsis	2.5	0.0208	

spread of retrotransposons was postulated as a main route enlarging plant genomes (SanMiguel *et al.*, 1996; SanMiguel *et al.*, 1998; Bennetzen and Kellog, 1997). Alternatively, deletions might reduce genome size and counterbalance enlargements (Petrov, 2001). Recently, an elegant theoretical study (Petrov *et al.*, 2000) has demonstrated that deletions of significantly different extension within retroelements yielded species-specific genome size alterations in related insect species over evolutionary time periods. We were able to demonstrate experimentally that the mean size of deletions obtained from DSB repair is on average a third larger in *Arabidopsis* than in tobacco (Kirik *et al.*, 2000). Thus, for plants and insects an inverse correlation between genome size and deletion size was found. However, one has to be careful to draw final conclusion from the comparison of only two species each, but it is tempting to speculate that the phenomenon might be a general principle of genome evolution.

During DSB repair the size of a deletion depends on the processing of DNA ends. If broken ends are not religated directly the processing of such ends might result in the loss of DNA at the break site (Salomon at al., 1998; Kirik *et al.*, 2000). Depending on the efficiency of DNA degradation more or less information will be lost. We were able to detect differences as to the stability of linearized plasmid DNA in tobacco and *Arabidopsis* cells. Although the results obtained by Southern blotting indicate that on average plasmid molecules transformed into *Arabidopsis* cells are less stable than in tobacco cells, we of course cannot exclude that this difference might not be caused by degradation in the nucleus, as plasmid molecules are also transferred into other compartments of the cell. However, as two other assays, that are clearly nuclear-based (chromosomal DSB repair [Kirik *et al.*, 2000] and the transient expression assay described in this report) demonstrate a similar difference using unrelated methodologies, we are tempted to speculate that this is caused by the same kind of phenomenon.

The results presented here indicate that the difference in the processing of DNA ends between the two plant species is at least mainly due to an enhanced exonucleolytic degradation of DNA in *Arabidopsis*. We were able to show that if a break was induced between the promoter and the open reading frame of a marker gene the expression level was reduced by a factor of two in *Arabidopsis* relative to tobacco. This is in contrast to the similar expression levels obtained for both species when the break was introduced distantly from the marker gene. Whereas in the first case exonucleolytic digest of few nucleotides would result in destruction of the expression cassette, in the latter case the removal of more than thousand base pairs would be a prerequisite for the destruction of the ORF. Because endonucleolytic degradation would destroy all information on a plasmid at the same rate independent of a pre-existing break site, no differences should be found in presence of such a nuclease activity. However, since the restoration of the β-glucuronidase ORF in case of the *Stu*I- and *Xba*I-restricted plasmid requires a ligation step, the data of the expression assay alone do not suffice to discriminate between a more efficient exonucleolytic destruction of the marker gene or a less efficient ligation reaction in *Arabidopsis* cells. As we were able to detect by Southern analysis differences in the degradation of plasmid DNA linearized with the exactly the same enzymes between *Arabidopsis* and tobacco, we think that we are allowed to speculate that at least most of the effect is due a more efficient exonucleolytic degradation of DNA in *Arabidopsis*. However, one question remains: what is the cause of the enhanced DNA degradation in *Arabidopsis*? This could be either due to less active DNA exonucleases or due to the better protection of DNA ends from degradation, for example by DNA end binding factors (Liang *et al.*, 1996) in tobacco.

Purposely, we did not test the stability of circular plasmid DNA in both plant species, as this approach would not have allowed us to discriminate between exo- and endonucleolytic degradation. Linearization of the plasmid DNA by an endogenous endonuclease would be a prerequisite of exonucleolitic degradation, however if the plasmid is linearized, exonuclease could immediately work on it. Thus, it would be hard to discriminate to which extent what kind of enzyme activity would contribute to DNA degradation

Table 3. Statistical evaluation of the data presented in Table 2. In the bottom line the means and standard deviations are given.

*Stu*I blunt in ORF	*Sca*I blunt out of ORF	*Xba*I 3′ overhang in ORF
2.43	1.03	1.40
2.40	1.14	2.00
2.35	0.50	1.60
2.15	0.71	1.38
2.18	1.23	2.10
1.97	0.50	1.50
2.25±0.18 (6)	0.85±0.32 (6)	1.66±0.31 (6)

in *Arabidopsis* and tobacco. Using linearized molecules we were working with a more defined system as endonucleases could work on all DNA molecules immediately after transformation and independent of a second reaction.

Theoretical calculations on the evolution of genome sizes have to take into account the rate of duplication and the rate of loss of genomic sequences (Petrov, 2001). DSB repair is one mechanism that may result in sequence deletions. To evaluate the possible role of DSB repair in this process one has to take into account two factors: the number of breaks induced and the loss of sequences information per repair event. The loss of sequence information is strongly influenced by the stability of the broken DNA ends. The more efficient the ends are recessed by exonucleolitic degradation the more the genome size is supposed to shrink over an evolutionary time period. In line with this argumentation we found indications that in a plant species with a small genome free DNA ends are less stable than in a plant species with a more than twenty times larger genome.

Acknowledgements

We would like to thank Ingo Schubert and two anonymous referees for useful criticism on the manuscript and Angela Kirik and Frank Hartung for help and discussions. The study was partly funded by a grant of the Deutsche Forschungsgemeinschaft to H.P.

References

Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815.

Bennett, M.D. and Leitch, I.J. 1997. Nuclear DNA amounts in angiosperms: 583 new estimates. Annal. Bot. 80: 169–196.

Bennetzen, J.L. and Kellog, E.A. 1997. Do plants have a one-way ticket to the genomic obesity? Plant Cell 9: 1509–1514.

Capy, P. 2000. Perspectives: evolution. Is bigger better in cricket? Science 287: 985–986.

Cavalier-Smith, T. 1985. The Evolution of Genome Size. John Wiley, Chichester, UK.

Chiurazzi, M., Ray, A., Viret., J.-F., Perera, R., Wang, X.-H., Lloyd, A. and Signer, E.R.. 1996. Enhancement of somatic intrachromosomal homologous recombination in *Arabidopsis* by HO-endonuclease. Plant Cell 8: 2057–2066.

Dove, G.A. and Flavell, R.B. 1988. Genome Evolution. Academic Press, London.

Gorbunova, V. and Levy, A.A. 1999. How plants make ends meet: DNA double-strand break repair. Trends Plant Sci. 4: 263–269.

Gorbunova, V. and Levy, A.A. 1997. Non-homologous DNA end joining in plant cells is associated with deletions and filler DNA insertions. Nucl. Acids Res. 25: 4650–4657.

Jasin, M. 2000. Chromosome breaks and genomic instability. Cancer Invest. 18: 78–86.

Kirik, A., Salomon, S. and Puchta, H. 2000. Species-specific double-strand break repair and genome evolution in plants. EMBO J. 19: 5562–5566.

Liang, F., Romanienko, P.J., Weaver, D.T., Jeggo, P.A. and Jasin, M. 1996. Chromosomal double-strand break repair in Ku80-deficient cells. Proc. Natl. Acad. Sci. USA 93: 8929–8933.

Mengiste, T. and Paszkowski, J. 1999. Prospects for the precise engineering of plant genomes by homologous recombination. J. Biol. Chem. 380: 749–758.

Ozkan, H., Levy, A.A. and Feldman, M. 2001. Allopolyploidy-induced rapid genome evolution in the wheat (*Aegilops-Triticum*) group. Plant Cell 13: 1735–1747.

Petrov, D.A. 2001. Evolution of genome size: new approaches to an old problem. Trends Genet. 17: 23–28.

Petrov, D.A. 1997. Slow but steady: Reduction of genome size through biased mutation. Plant Cell 9: 1900–1901.

Petrov, D.A., Lozovskaya, E.R and Hartl, D.L. 1996. High intrinsic rate of DNA loss in *Drosophila*. Nature 384: 346–349.

Petrov, D.A., Sangster, T.A., Johnston, J.S., Hartl, D.L. and Shaw, K.L. 2000. Evidence for DNA loss as a determinant of genome size. Science 287: 1060–1062.

Pipiras, E., Coquelle, A., Bieth, A. and Debatisse, M. 1998. Interstitial deletions and intrachromosomal amplification initiated from a double-strand break targeted to a mammalian chromosome. EMBO J. 17: 325–333.

Puchta, H. 1999. DSB-induced recombination between ectopic homologous sequences in somatic plant cells. Genetics 152: 1173–1181.

Puchta, H. and Hohn, B. 1991. A transient assay in plant cells reveals a positive correlation between extrachromosomal recombination rates and length of homologous overlap. Nucl. Acids Res. 19: 2693–2700.

Puchta, H. and Hohn, B. 1996. From centiMorgans to basepairs: homologous recombination in plants. Trends Plant Sci. 1: 340–348.

Salomon, S. and Puchta, H. 1998. Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J. 17: 6086–6095.

SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z. and Bennetzen, J.L. 1996. Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765–768.

SanMiguel, P., Gaut, B.S., Tikhonov, A., Nakajima, Y. and Bennetzen, J.L. 1998. The paleontology of intergene retrotransposons of maize. Nature Genet. 20: 43–45.

Seki, M., Iida, A. and Morikawa, H. 1999. Transient expression of the β-glucuronidase gene in tissues of *Arabidopsis thaliana* by bombardment-mediated transformation. Mol. Biotechnol. 11: 251–255.

Shaked, H., Kashkush, K., Ozkan, H., Feldman, M. and Levy, A.A. 2001. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. Plant Cell 13: 1749–1759.

Shalev, G. and Levy, A.A. 1997. The maize transposable element *Ac* induces recombination between the donor site and an homologous ectopic sequence. Genetics 146: 1143–1151.

Siebert, R. and Puchta, H. 2002. Efficient repair of genomic double-strand breaks via homologous recombination between directly

repeated sequences in the plant genome. Plant Cell 14: 1121–1131.

Smith, GP. 1976. Evolution of repeated DNA sequences by unequal crossover. Science 191: 528–535.

Vergunst, A.C. and Hooykaas, P.J.J. 1999. Recombination in the plant genome and its application in biotechnology. Crit. Rev. Plant. Sci. 18: 1–31.

Walbot, V. 1996. Sources and consequences of phenotypic and genotypic plasticity in flowering plants. Trends Plant Sci. 1: 27–33.

Xiao, Y.L, Peterson, T. 2000. Intrachromosomal homologous recombination in *Arabidopsis* induced by a maize transposon. Mol. Gen. Genet. 263: 22–29.