

ORIGINAL ARTICLE

Frequent silencing of rDNA loci on the univalent-forming genomes contrasts with their stable expression on the bivalent-forming genomes in polyploid dogroses (*Rosa* sect. *Caninae*)

L Khaitová¹, G Werlemark², H Nybom² and A Kovarik¹

¹Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i. Laboratory of Molecular Epigenetics, Brno, Czech Republic and
²Balsgård-Department of Plant Breeding and Biotechnology, Swedish University of Agricultural Science, Kristianstad, Sweden

The polyploid species in *Rosa* section *Caninae* ($2n=21$, 28 or 35) are characterized by an unusual reproductive system known as odd (or asymmetric) meiosis. Only two chromosome sets form bivalents in meiosis, whereas the remaining chromosomes are transmitted as univalents through the female germline. Evolution of ribosomal rRNA genes (rDNA) does not seem to be significantly affected by interlocus homogenization in dogroses. As a consequence, most species contain several rDNA families falling into two main clades (β and γ) thought to be differentially distributed between bivalent and univalent chromosomes, respectively. Here, we have investigated expression of rRNA gene families in five pentaploid species (*R. canina*, *R. rubiginosa*, *R. dumalis*, *R. sherardii* and *R. caesia*, $2n=35$) and in one tetraploid (*R. mollis*, $2n=28$). Using extensive sequencing of

ITS clones and cleaved amplified polymorphism sequence (CAPS) analysis, we found that the β -family was constitutively expressed in all species. However, there was large variation in the expression patterns of families constituting the γ -clade. In addition, a single family can be active in one species, whereas silenced in another. The data show that the families on bivalent-forming chromosomes dominate rDNA expression in all dogrose species. We hypothesize that genes on bivalent genomes are stably expressed, whereas those on univalent genomes undergo variable levels of epigenetic silencing. Nonetheless, mosaic expression of univalent genomes could contribute to phenotypic variation between the species.

Heredity (2010) **104**, 113–120; doi:10.1038/hdy.2009.94; published online 29 July 2009

Keywords: polyploidy; rRNA genes; expression; concerted evolution; Rosaceae

Introduction

The genus *Rosa*, comprising large numbers of traditional ornamental plants, has long been a target for breeding and genetic research. The section *Caninae* DC forms a large and well-defined group of polyploid taxa known as dogroses. Pentaploids are most common, but tetraploids and hexaploids also occur (Wissemann, 2003a). All dogrose species have the peculiar 'canina' meiosis described more than 80 years ago (Tackholm, 1920; Blackburn and Heslop-Harrison, 1921). Only seven bivalents are formed in the first meiotic division. The remaining chromosomes occur exclusively as univalents and are not included in viable pollen grains, which contain only the seven divided bivalent chromosomes. All the univalents are, however, transmitted to one of the daughter cells in the female meiosis and are eventually included in the viable egg cells, which, therefore, contain 21, 28 or 35 chromosomes depending on ploidy levels.

The evolutionary significance of this strange meiotic system remains enigmatic. However, recent molecular investigations have shed some light on the origin and structure of the individual subgenomes. Specifically, microsatellite DNA markers (Nybom *et al.*, 2004, 2006), RAPD (Werlemark and Nybom, 2001) and nuclear ribosomal RNA genes (rDNA) (Kovarik *et al.*, 2008a) have shown that the bivalent-forming genomes seem to be rather similar across different dogrose species, even those occurring in different subsections. In contrast, distinctly different univalent genomes occur within single genotypes, but also show levels of differentiation that are associated with taxonomic distances between genotypes.

The rDNA genes (rDNA) coding for 18S-5.8S-26S RNA are organized on chromosomes as long tandem arrays that may comprise thousands of units in plants (Hemleben and Zentgraf, 1994). The intragenic (ITS) and intergenic spacers associated with genic regions often vary between species and may be used to detect parentage of allotetraploids (Nieto Feliner and Rossello, 2007; Volkov *et al.*, 2007). Just as other tandemly arranged sequences, the rDNA arrays undergo concerted evolution, which maintains their homogeneity (Dover, 1982). However, interlocus homogenization seems to be less efficient than that of intralocus (Schlotterer and Tautz, 1994). This is probably why some allopolyploids maintain their

Correspondence: Dr A Kovarik, Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i., Laboratory of Molecular Epigenetics, Královopolská 135, Brno CZ-612 65, Czech Republic.

E-mail: kovarik@ibp.cz

Received 4 May 2009; revised 25 June 2009; accepted 29 June 2009; published online 29 July 2009

parental genes, whereas others are characterized by extensive homogenization of repeats across the subgenomes. The pentaploid dogroses belong to the former group, having different ITS types (Wisseemann, 2003b) that fall into two phylogenetically well-separated clades (Kovarík *et al.*, 2008b) and are distributed at least on five chromosomes in *Rosa canina* (Lim *et al.*, 2005). Thus, interlocus homogenization may be limited or non-existing in *Rosa* allopolyploids. The β -family seems to be ubiquitously present in all three polyploids analyzed (*R. canina*, *R. rubiginosa* and *R. dumalis*), whereas the other families showed more restricted distribution (Kovarík *et al.*, 2008b). In addition, when the DNA was extracted from mature pollen grains (thought to contain predominantly haploid bivalent-forming genomes) the ratios were skewed toward the β -family. These data indicated that the β -rDNA family is likely to occur on bivalent-forming genomes supporting the hypothesis that similar genomes may be involved in meiotic pairing in different pentaploid species.

Epigenetic information, not directly encoded by a primary DNA sequence, seems to be important for a variety of biological processes including regulation of gene transcription and the maintenance of genome integrity and development (Colot and Rossignol, 1999; Richards, 2006). It is believed that on allopolyploidization events, epigenetic processes may help to harmonize expression of parental genomes in a unified nucleus (Adams and Wendel, 2005). Indeed, numerous studies have shown changes in DNA methylation and chromatin early after the formation of an allopolyploid nucleus (Feldman and Levy, 2005; Salmon *et al.*, 2005). One of the well-known epigenetic phenomena is nucleolar dominance (Navashin, 1934) of which the molecular basis is uniparental silencing of one parental family of rRNA genes (Preuss and Pikaard, 2007).

Although the rRNA silencing has been described in many synthetic and natural allopolyploids (Volkov *et al.*, 2007), there are no reports on rRNA gene expression in systems with highly non-concerted evolution of rDNA repeats, transmitted differently through the male and female meiosis, respectively. In this study, we have addressed the question of expression of individual rRNA gene families thought to originate from bivalent and univalent *Caninae* genomes. To achieve this aim, we have carried out sequencing of ITS clones derived from RNA (cDNA) and genomic DNA (gDNA) in several pentaploid and one tetraploid species. Evidence was obtained for stable expression of a proto-canina rRNA family and variable silencing of other families likely to originate from univalent genomes.

Materials and methods

Plant material

R. canina (accession 1073), *R. rubiginosa* (0391), *R. dumalis* (8701), *R. sherardii* (1402), *R. caesia* (504) and *R. mollis* (105) plants were collected as root suckers in the wild and subsequently grown at Balsgård, South Sweden (Nybom *et al.*, 2006). Ploidy levels were determined by flow cytometry in earlier studies (Nybom *et al.*, 2004, 2006). All of these accessions were found to be pentaploid ($2n = 5x = 35$) except for *R. mollis*, which was tetraploid ($2n = 4x = 28$). Young fresh leaf (~ 100 – 200 mg

tissues) was collected and stored in RNAlater (Applied Biosystems, Foster City, CA, USA) solution at -20°C until use.

Nucleic acids isolation and treatment

RNA was extracted using an RNAeasy kit (Qiagen, Germany) according to the manufacturer's recommendation. RNA samples were treated with Turbo DNase I (Ambion, Austin, Texas, USA) ($0.1\text{ U}\mu\text{g}^{-1}$ RNA; 30 min per 37°C) to eliminate any contaminating gDNA. Quality of RNA preparations was checked with agarose gel electrophoresis. The cDNAs were prepared in a reverse transcription reaction ($20\mu\text{l}$) typically containing $1\mu\text{g}$ of total RNA, 2 pmols of random primers and 200 units of reverse transcriptase (Invitrogen Superscript II RNase H) following conditions recommended by the supplier (Invitrogen, Paisely, UK). The gDNA was extracted using a modified cetyltrimethylammonium bromide method. Briefly, the leaf tissue was washed from RNAlater with distilled water and homogenized in an extraction buffer (50 mM Tris HCl, pH 8.0, 0.8 M NaCl, 10 mM EDTA, 1.2% cetyltrimethylammonium bromide and 0.01% 2-mercaptoethanol) using a micropestle (Eppendorf, Germany). The downstream purification steps were as described in Fojtova *et al.* (1998).

The concentration of purified nucleic acids was estimated by spectrophotometry (Nanodrop, Thermo Scientific, Wilmington, DE, USA).

ITS amplification

For PCR ($25\mu\text{l}$) amplification, we used 100 ng of input cDNA or 100 ng of gDNA as template, 40 pmol of each primer, 12 nmol of each dNTP, 1.9 U of DyNAzyme II DNA polymerase (Finnzymes, Espoo, Finland). The primers were as follows: for ITS1-5.8S-ITS2, we used the 18Sfor ($5'$ -GCGCTACTGATGTATTCAACGAG- $3'$) and 26Srev ($5'$ -CTTTTCCTCCGCTTATTGATATGC- $3'$) pair; and for ITS1, the 5.8Srev ($5'$ -CGCAACTGCGTTCA AAGACTCGA- $3'$) primer was used in combination with the 18Sfor primer. Cycling conditions were as follows: initial denaturation step (92°C , 180 s) and 35 cycles (92°C , 20 s; 57°C , 30 s; 72°C , 30 s) followed by a final 72°C extension for 10 min. For each RNA isolate, we performed several control experiments involving amplification reactions on DNA templates and RNA templates before reverse transcription. No product was obtained on RNA samples without a reverse transcription step. The products were cloned into a pDrive vector (Qiagen) or analyzed directly by cleaved amplified polymorphism sequence (CAPS) (further below). About 20 clones from each sample were sequenced from both ends (MWG Eurofinn, Germany). The newly identified ribotypes were submitted to the EMBL/GenBank under the accession numbers FJ947107-FJ947110 and FJ948761.

Cleaved amplified polymorphism sequence

The PCR products from gDNA or cDNA amplification reactions were digested with an excess of *Bst*NI, *Rsa*I or *Bst*UI restriction enzymes (NEB, Beverly, Maryland, MD, USA) and the resulting fragments were separated on a 2% agarose gel. After RT-CAPS analysis, the gel was stained with ethidium bromide and visualized using ultraviolet light transillumination system (CCD camera was a Discovery model, Ultra-Lum, Claremont, CA,

USA). Images were processed by UltraQuant molecular imaging and analysis software (Ultra-Lum).

Data analysis

In total, the ITS datasets comprised ~220 DNA sequences derived from genomic and cDNA clones. The sequences were assembled and aligned using the BIOEDIT 3 program (Hall, 1999), using the CLUSTAL W algorithm. The central region from nt -168 to -48 (with respect to the 5.8S coding region) containing hypervariable sites (Kovarik et al., 2008b) was selected and used for phylogenetic analysis. The final alignment contained 120 nucleotides without indels. The phylogenetic relationships of ITS families were inferred from phylograms constructed based on the neighbor-joining approach with the setting of the Kimura two-parameter substitution model. The statistical support for branching was obtained with a bootstrap analysis: 500 repetitions were carried out using the PHYLOWIN program (Galtier et al., 1996). Only medium and strongly supported branches (bootstrap support 60–100%) were considered. The distances between individual clones were calculated with the assistance of DnaSP 4.0 software (Rozas et al., 2003) using the partial (3'-end) 18S rRNA genic and full length ITS1 sequences.

Results

Analysis of ITS from gDNA clones

We cloned and sequenced ITS from *R. caesia*, *R. sherardii* and *R. mollis* from gDNA. About 20 clones from each species were sequenced. The sequences were aligned and checked for polymorphic sites (for the alignment matrix, see Supplementary Figure S1). In *R. caesia* and *R. sherardii*, essentially the same mutation hot spots (8–9) were identified as found in earlier ITS studies of *R. canina*, *R. rubiginosa* and *R. dumalis* (Kovarik et al., 2008b). In addition, four sites (-48, -119, -129, -156) were found to be highly polymorphic in the tetraploid *R. mollis*, whereas they were apparently monomorphic in the pentaploid species. The high-copy polymorphisms arising from several mutation hot spots were used to define gene families. Low-copy number mutations occurring once (singletons) in an alignment matrix were not considered as separate ribotypes. The consensus sequences for individual gene families and phylogenetic relationships are shown in Table 1 and Figure 1, respectively. One well-supported branch was formed

by α - and β -families. The second, more diversified clade, contained several related families (γ , δ , ϵ , ω and their derivatives). *Rosa caesia* and *R. sherardii* shared the gene families (α , β , γ and ϵ) with other pentaploid dogroses. In contrast, *R. mollis* had two, possibly three, additional unique ITS types (ω , ϵ' , ϵ'') falling within the γ -family branch. The overall divergence of ITS clones was also relatively high in *R. mollis* (Table 2). Quantitative representation of individual gene families among the species is shown in Supplementary Table S1 and Figure 2 ('gDNA' columns).

Analysis of ITS from cDNA clones

Using the RT-PCR approach, we analyzed expression of rRNA genes in leaf material from five pentaploid species (*R. canina*, *R. rubiginosa*, *R. dumalis*, *R. sherardii* and *R. caesia*) and one tetraploid (*R. mollis*). We obtained ~120 ITS sequences that were aligned and analyzed as described above for the gDNA. Diversity data for comparisons between clones were calculated (Tables 2 and 3). Evidently, overall heterogeneity of cDNA clones was slightly lower compared with gDNA clones (with exception of *R. mollis*) indicating that only particular subsets of gene families had been expressed. Relatively homogeneous ITS transcripts were found in *R. rubiginosa*, whereas both *R. caesia* and *R. dumalis* displayed higher diversities. These findings are consistent with a more heterogenous spectrum of gene families in the latter two species. The contribution of each family to the total rRNA pool is shown in Figure 2 ('cDNA' columns).

As hundreds of rRNA genes may be transcribed in the cell, we used CAPS analyses to obtain a more statistically valid overview of rRNA gene expression patterns. The restriction cleavage profiles of amplification products from cDNA and gDNA are shown in Figure 3. In general, there were fewer bands in cDNA profiles compared with those of gDNA.

There are two *Bst*NI sites in the ITS1-5.8S-ITS2 region of the β -clade, whereas there is a single *Bst*NI site in most members of the γ -clade. Consequently, the ~500 bp band represents the β -clade families, whereas the heavier, ~700 bp, band comprise mostly the γ -clade families (Figure 3b). It is evident that the ~500 bp band had similar relative intensity among species, and also among lanes loaded with cDNA or gDNA, respectively. In contrast, intensity of the ~700 bp band varied across the species and also between lanes loaded with cDNA and gDNA from the same species. For example, in lanes

Table 1 Classification of major rRNA gene families according to the ITS nucleotide sequence at diagnostic sites

Position/family	-168	-156 ^a	-149	-138	-137	-130	-129 ^a	-120	-119 ^a	-112	-111	-102	-56	-48 ^a
α	G	T	C	T	C	C	C	C	A	T	C	G	G	G
β	G	T	C	T	C	C	C	C	A	C	T	G	G	G
β'	G	T	C	T	C	C	C	C	A	C	T	A	G	G
γ	A	T	T	G	T	C	C	T	A	C	T	G	T	G
γ'	A	T	T	G	T	T	C	T	A	C	T	G	G	G
δ	A	T	T	G	T	C	C	T	A	C	T	G	G	G
ϵ	A	T	T	G	T	C	C	C	A	C	C	G	G	G
ϵ'	A	T	T	G	C	T	C	C	A	C	C	G	G	G
ϵ''	A	T	T	G	C	T	C	C	A	T	C	G	G	G
ω	A	C	T	G	C	C	C	C	G	C	T	G	G	G
ω'	A	C	T	G	C	C	T	C	G	C	T	G	A	A

^aSites polymorphic in tetraploid species only.

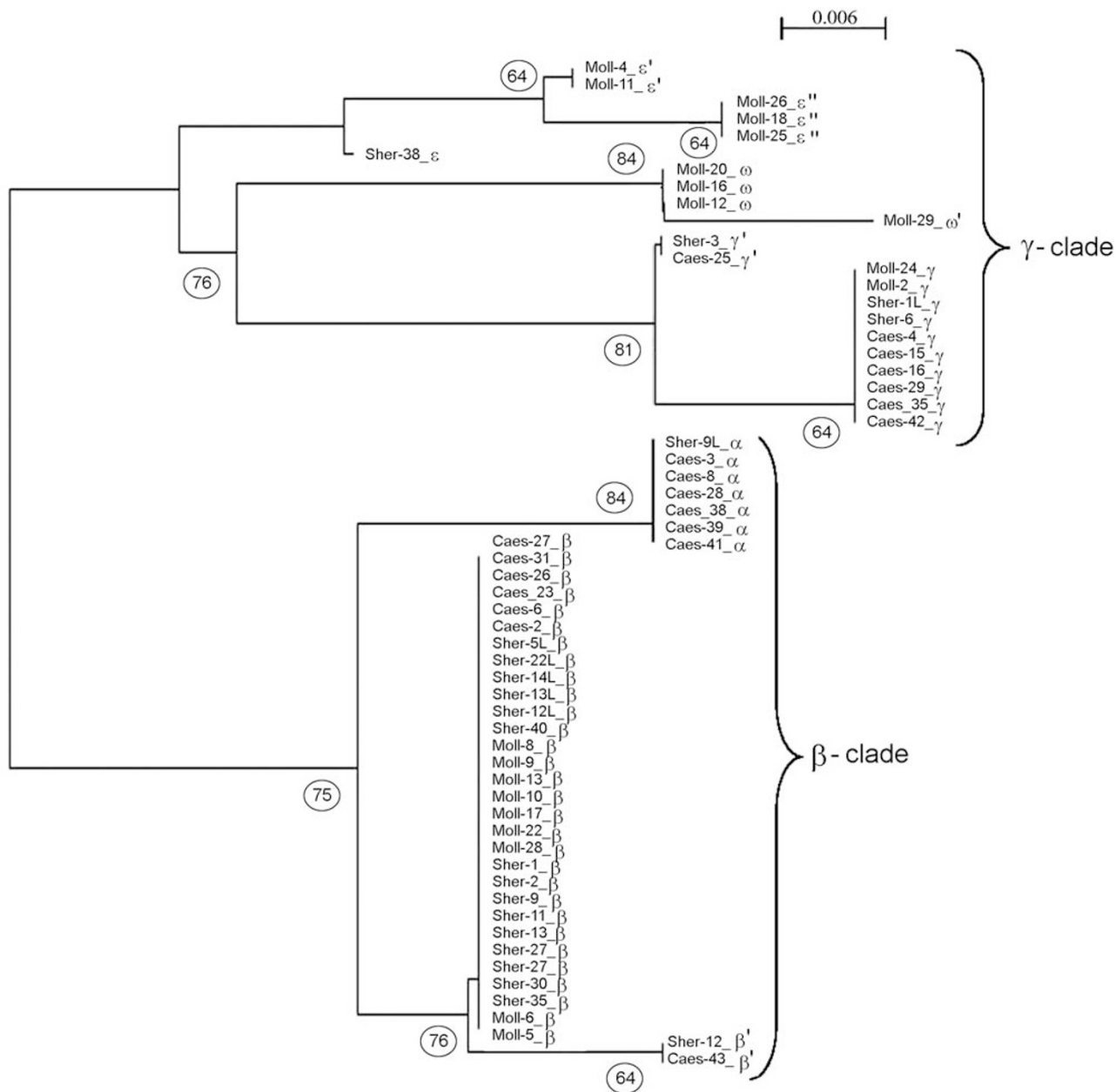


Figure 1 Phylogenetic relationships between ITS1 clones. The clones were annotated as follows: moll- *R. mollis*, sher- *R. sherardii*, caes- *R. caesia*. The tree has been constructed using a neighbor-joining method (Kimura distance model). Bootstrap values of >60% are indicated. Scales indicate the base substitutions per site.

Table 2 Divergences of ITS1 sequences among the species

Species	<i>gDNA</i> (Pi) ^a	<i>s.d.</i>	<i>cDNA</i> (Pi)	<i>s.d.</i>
<i>R. caesia</i>	0.03268	0.00362	0.02666*	0.00556
<i>R. sherardii</i>	0.01831	0.00544	0.01156	0.00564
<i>R. rubiginosa</i>	0.00296	0.00036	0.00268	0.00198
<i>R. canina</i>	0.01279	0.00148	0.01156*	0.00138
<i>R. dumalis</i>	0.03230	0.00234	0.02629*	0.00604
<i>R. mollis</i>	0.03378	0.00313	0.03851*	0.00572

Abbreviation: *s.d.*, standard deviation.

^aPi—average nucleotide diversity per site

*Statistically significant difference between the genomic and cDNA clones, $P < 0.01$.

loaded with amplified *R. canina*, *R. mollis* and *R. rubiginosa* cDNA, the 700 bp signal was absent or at least much weaker than in the lane loaded with gDNA. On the other hand, the 700 bp band was intensive in a cDNA lane containing *R. dumalis* PCR product lane, which is consistent with relatively high proportion of the γ -clade transcripts in this species (Figure 2). Compared with the *Bst*NI profiles, the *Rsa*I digestion showed more sequence polymorphisms (Figure 3c) revealing subtle differences between individual ITS types. In *R. canina* and *R. dumalis*, five or six bands were found in lanes loaded with gDNA, whereas only four or five clear bands showed up in lanes loaded with corresponding products of cDNA

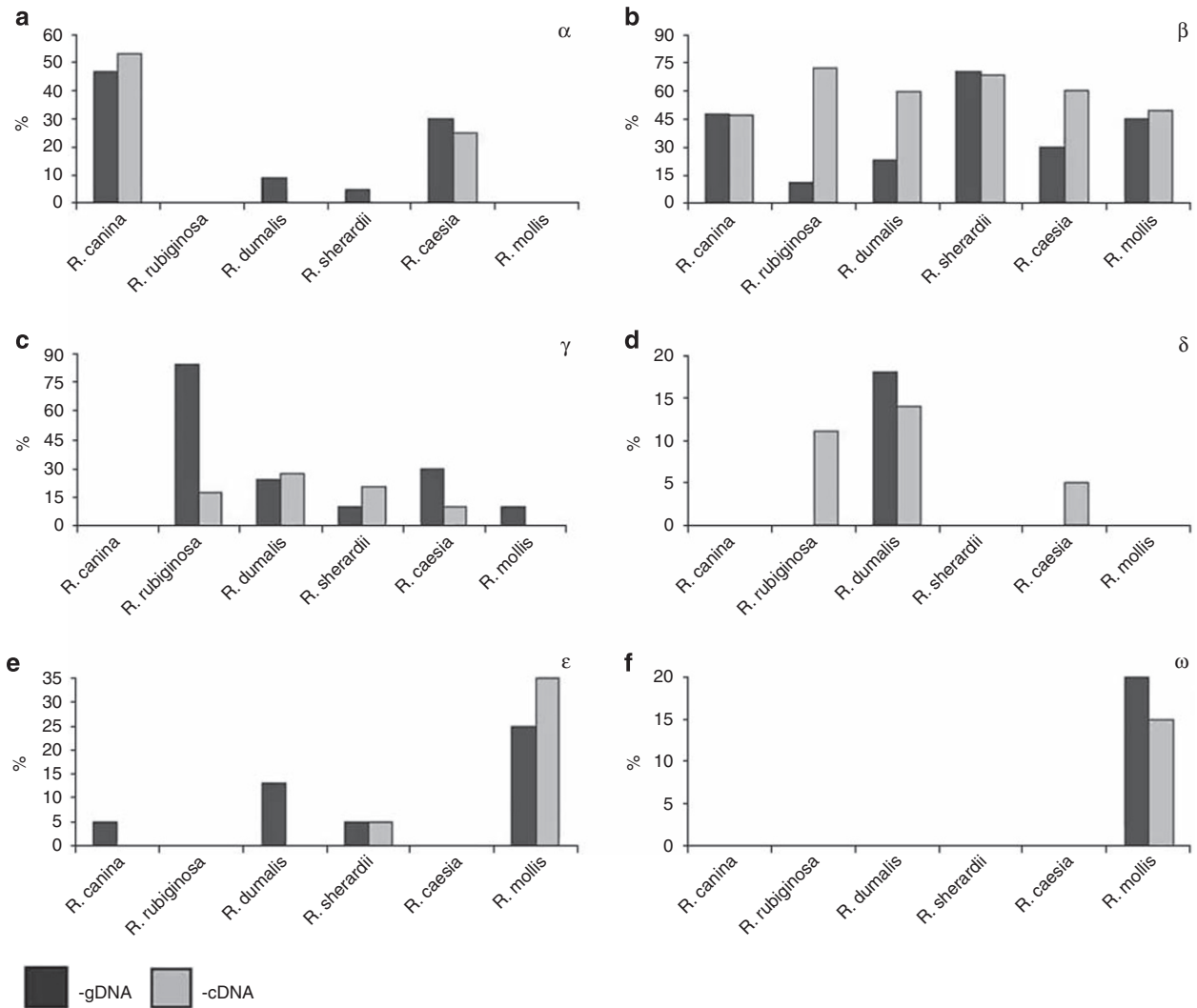


Figure 2 Analysis of ITS1 clones in gDNA and cDNA. Y axis, relative content of gene families in percentages. Scales were proportionally adjusted according to the largest family in each graph. The total numbers of clones and families are given in Supplementary Table S1. The ϵ and ω panels included all the derivatives (ϵ' , ϵ'' and ω'). All species except *R. mollis* were pentaploids ($2n = 5x = 35$); *R. mollis* is a tetraploid ($2n = 4x = 28$).

Table 3 Comparison of divergences of genic and intragenic regions

Region/source	18S (3'-end, 250 bp)			ITS1 (261 bp)		
	N^a	Pi	$s.d.$	N^a	Pi	$s.d.$
gDNA	91	0.00373	0.00073	98	0.02186*	0.00166
CDNA	105	0.00316	0.00082	113	0.01715*	0.00117

The data were collected from sequencing of cloned 18S-5.8S-26S subregions in six species. The symbols are the same as in Table 2.

^a N —number of sequences within the datasets (Supplementary Table S1). Some sequences containing incomplete reads were not included in the analysis.

*Statistically significant difference between the 18S and ITS clones, $P < 0.01$.

amplification. The *Bst*UI was used to discriminate between the ω -family and the rest of ITSs classes (Figures 3a and d). There were three bands in the *R. mollis* 'gDNA' lane, whereas there were two bands in that of *R. rubiginosa* and other pentaploid species (not shown).

The extra ω -specific band was not visible after the digestion of the cDNA product confirming low expression of the ω -family in *R. mollis*.

Discussion

Polyploid dogroses contain both conserved and divergent rDNA types

Our earlier study revealed the presence of multiple ITS families in three pentaploid species, *R. canina*, *R. rubiginosa* and *R. dumalis* (Kovarik et al., 2008b). In this study, we extended the analysis to two pentaploid (*R. sherardii*, *R. caesia*) and one tetraploid (*R. mollis*) species. Whereas *R. sherardii* had a relatively homogeneous spectrum of ITS families and a prevalence of the β -family, *R. caesia* contained three families (α , β , γ) in approximately equal proportions. A similar spectrum of ITS families was found in *R. dumalis* confirming close phylogenetic relationships between these two species; *R. caesia* is a synonym of *R. coriifolia* (Wissemann, 2003a),

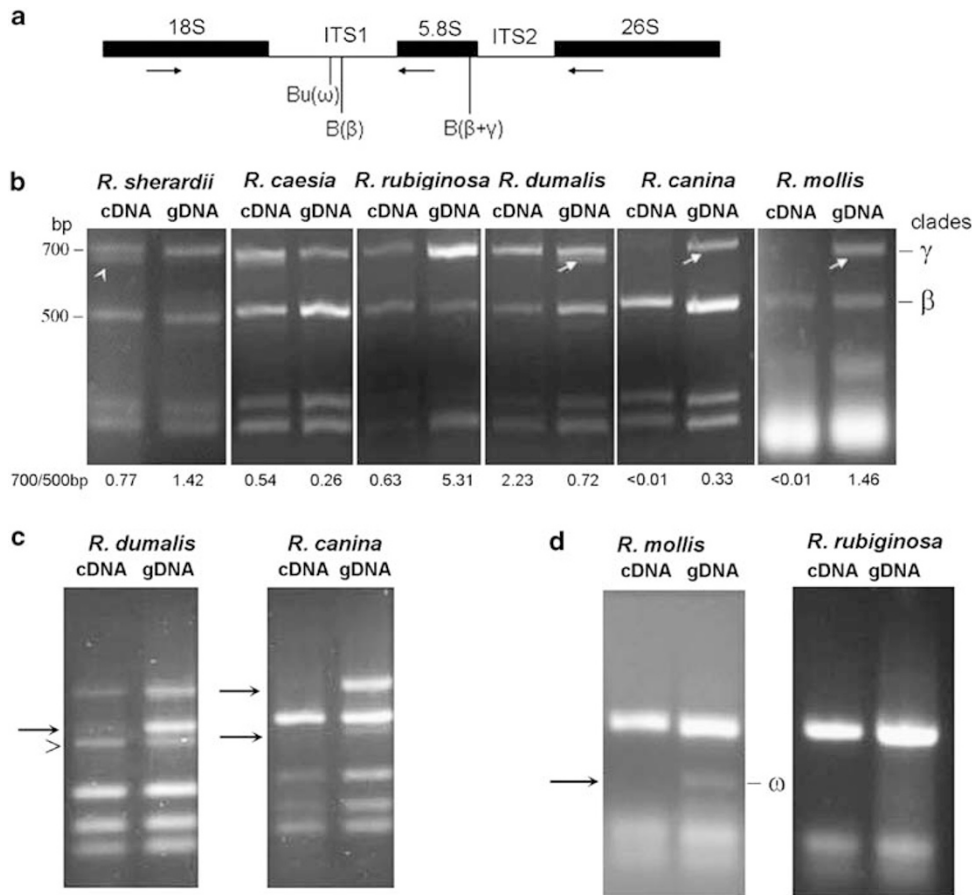


Figure 3 Cleaved amplified polymorphism analysis of ITS sequences in gDNA (lanes 'gDNA') and cDNA. (a) Positions of primers and diagnostic restriction sites along the amplified sequences. The polymorphic *Bst*NI site (B) discriminating between the β - and γ -clade families is indicated. (b) gCAPS and cCAPS profiles of products cleaved with *Bst*NI. Note that the ratio between the 700 and 500 bp signals varies between the samples. PCR products digested with *Rsa*I (c) and *Bst*UI (d) The bands revealing differences between the gDNA and cDNA profiles are indicated by arrows (mostly present in gDNA) and arrowheads (mostly present in cDNA). The fluorescence intensity ratio between the 700 and 500 bp *Bst*NI bands (representing the γ/β -clades) is shown below each lane.

which is often treated as *R. dumalis* subsp. *coriifolia* (Nilsson, 1967). Earlier sequencing of clones revealed two major ITS families (α , β) in *R. canina* (Kovarík *et al.*, 2008b). However, the results of CAPS analysis are more consistent with the presence of additional families including those of the γ -clade in this species. Perhaps, our screens of clones did not reach saturation levels despite considerable number of sequences obtained (~ 60) and minor variations could have been missed. Tetraploid species could be expected to contain a more homogenous spectrum of ITS types than those of pentaploid species because of the lower number of univalent genomes. However, this does not seem to hold true for *R. mollis*, which contained several divergent families in addition to the dominant β -family. The unique ω -family and derivatives of the ε -family have not been found earlier and may originate from a distantly related species (the ω -family sequence was nearly identical to the ITS of *R. rugosa*).

Intralocus but not interlocus concerted evolution of rDNA repeats

The phylogenetic tree (Figure 1) showed more diversification of the γ -clade clones compared with those of the β -clade. An explanation is that the closely related β -clade

families were acquired by all allopolyploid dogroses from a common proto-canina ancestor, whereas the γ -clade families were inherited from divergent progenitors. However, it is also possible that homogenization pressures operating on rDNAs differ between genetic compartments of the allopolyploid. If so, the repeats on bivalent chromosomes would be more homogeneous compared with those on univalent chromosomes, as concerted evolution of repeated sequences seems to depend on the number of meiotic recombination events (Smith, 1974). Certainly, members of the highly homogeneous β -family are predominantly located on the bivalent chromosomes in at least three dogrose species (Kovarík *et al.*, 2008b). It is also intriguing that, in our experiments, recombinants between the two divergent ITS clades (β and γ) seem to be very few. Rare chimeric clones (usually $< 5\%$) probably represent PCR artifacts (Cronn *et al.*, 2002) rather than true genomic recombinants. In contrast, recombinant ITS types have been observed in allopolyploid systems with regular meiosis (Buckler *et al.*, 1997; Franzke and Mummenhoff, 1999; Nieto Feliner *et al.*, 2004; Kovarík *et al.*, 2005). The situation here may resemble that of paeonia hybrids, which frequently undergo vegetative reproduction and whose ITSs have retained much of the parental features (Sang *et al.*, 1995). Together, the absence or reduced

frequency of extensive interlocus homogenization is likely to be a consequence of the non-symmetrical meiosis in dogroses, which prevents meiotic recombination of univalent chromosomes.

Despite the overall intergenomic ITS heterogeneity, the intra-family homogeneity of units was high in both clades. For example, diversity between the members of the γ -family (>300 clones isolated from different species) was <0.005 (Figure 1; Kovarik *et al.*, 2008b). Given that the γ -family is likely located on the univalent chromosomes (based on the differential analysis of leaf and pollen clones), it seems that there is little evidence for degeneracy of univalent genome sequences. Perhaps, insufficient time has elapsed, as the allopolyploidy event to accumulate mutations in the array. However, it is also possible that some kind of rDNA homogenization may be occurring in the absence of regular meiosis. In support, rDNAs on non-recombining B-chromosomes of *Crepis* were found to be relatively homogeneous within the arrays (Leach *et al.*, 2005). Further, well-controlled experiments in a parthenogenic freshwater crustacean, *Daphnia*, revealed variation in arrays sizes over generations (McTaggart *et al.*, 2007), and it has been proposed that cross overs between sister chromatids could account for these unexpected rDNA dynamics (Eickbush and Eickbush, 2007). Thus, mitotic non-homologous recombination, organization of rDNA interphase (Kovarik *et al.*, 2008a), chromatin and DNA modification (Dadejova *et al.*, 2007) might be the factors contributing to the intra-array homogeneity of units. Certainly, it will be interesting to study homogeneity of rDNA in allopolyploid genera differing in the extent/frequency of sexual reproduction.

Silencing of rRNA genes in univalent, but not bivalent genomes

Theoretical predictions based on secondary structure models have suggested that most ITS1 types are potentially functional in dogroses (Ritz *et al.*, 2005). Although the GC content of the γ -ITS clade is somewhat lower (55–56%) than that of the β (~59%) clade, this difference is unlikely to mark the γ -clade families as pseudogenes. Indeed, we show that members of both clades are expressed. However, there were differences in the stability of expression patterns. For example, the γ - and δ -families showed contrasting and, in some cases, even reciprocal expression patterns: the γ -family units were largely silenced in *R. rubiginosa*, but not in *R. dumalis*. In contrast, the δ -family was partially silenced in *R. dumalis*, but expressed in *R. rubiginosa* despite its very low abundance in this species. A noteworthy fact is that even minor families may be highly transcriptionally active as indicated by enhanced signals in CAPS analysis (Figure 3b, panel *R. sherardii* and Figure 3c, panel *R. dumalis*). The phenomenon of low-copy gene-family expression dominance has also been observed in other allopolyploid systems (Joly *et al.*, 2004; Matyasek *et al.*, 2007). With all evidence taken together, it seems that the β -family, residing on bivalent-forming genomes (Kovarik *et al.*, 2008b) is stably expressed across the species, whereas families residing on univalent genomes tend to be suppressed. However, suppression does not seem to be absolute, as a family that is suppressed in one species can be active in other species.

Highly divergent families, perhaps including pseudogenes, are expressed in grasshopper species in which rDNA apparently escapes concerted evolution, as frequent mutations were found in the genic regions (Keller *et al.*, 2006). In dogroses, the 18S genic part of rDNA units did not contain mutation hot spots and the overall sequence divergence was ~5 fold lower than that of ITS1 (Table 3). Rare mutations were restricted to low frequency singletons (Supplementary Figure S1). It seems that selection has acted to maintain most of units within the array functional (and homogeneous) despite non-concerted evolution across the arrays. However, the functionality of individual transcripts in ribosomes remains to be determined. Apparently, functional copies are almost completely silenced immediately after formation of the allopolyploid nucleus in synthetic hybrids of *Nicotiana* (Dadejova *et al.*, 2007) and some *Arabidopsis* tetraploids (Beaulieu *et al.*, 2009). Obviously, there is no strict consensus concerning the gene-copy number, repeat homogeneity and transcription.

Epigenetic variation at univalent genomes—a hypothesis

The question arises as to the relevance of our findings for the general epigenetic landscape of dogrose genomes. In the *Arabidopsis suecica* allopolyploid, uniparental silencing of rRNA genes seems to correlate with an overall lower expression of the *A. thaliana* subgenome (Wang *et al.*, 2006). Hypothetically, in section *Caninae*, genes on univalent genomes may be more vulnerable to silencing compared with those located on bivalent genomes. Active constitutive expression of bivalent genomes may be needed for basal metabolism and successful transmission of genetic material in gametes. On the other hand, a variable mosaic expression of univalent genomes might be responsible for phenotypic variation and expression of species-specific traits. Clearly, expression analysis of other homeologous genes will be needed to support the hypotheses.

Acknowledgements

We thank Blazena Koukalova (Academy of Sciences) for helpful discussions and critical reading of the paper. This research was funded by the Grant Agency of the Czech Republic (521/07/0116, GACR204/09/H502, MSMT LC06004), Academy of Sciences of the Czech Republic (AVOZ50040507 and AVOZ50040702) and by the Philip-Sørensen Foundation.

References

- Adams KL, Wendel JF (2005). Polyploidy and genome evolution in plants. *Curr Opin Plant Biol* 8: 135–141.
- Beaulieu J, Jean M, Belzile F (2009). The allotetraploid *Arabidopsis thaliana*-*Arabidopsis lyrata* subsp. *petraea* as an alternative model system for the study of polyploidy in plants. *Mol Genet Genomics* 281: 421–435.
- Blackburn K, Heslop-Harrison JW (1921). The status of the British rose forms as determined by their cytological behaviour. *Ann Bot* 35: 159–188.
- Buckler ES, Ippolito A, Holtsford TP (1997). The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145: 821–832.
- Colot V, Rossignol JL (1999). Eukaryotic DNA methylation as an evolutionary device. *BioEssays* 21: 402–411.

- Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF (2002). PCR-mediated recombination in amplification products derived from polyploid cotton. *Theor Appl Genet* **104**: 482–489.
- Dadejova M, Lim KY, Souckova-Skalicka K, Matyasek R, Grandbastien MA, Leitch A *et al.* (2007). Transcription activity of rRNA genes correlates with a tendency towards intergenomic homogenization in Nicotiana allotetraploids. *New Phytol* **174**: 658–668.
- Dover G (1982). Molecular drive: cohesive mode of species evolution. *Nature* **9**: 111–116.
- Eickbush TH, Eickbush DG (2007). Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**: 477–485.
- Feldman M, Levy AA (2005). Allopolyploidy—a shaping force in the evolution of wheat genomes. *Cytogenet Genome Res* **109**: 250–258.
- Fojtova M, Kovarik A, Votruba I, Holy A (1998). Evaluation of the impact of S-adenosylhomocysteine metabolic pools on cytosine methylation of the tobacco genome. *Eur J Biochem* **252**: 347–352.
- Franzke A, Mummenhoff K (1999). Recent hybrid speciation in *Cardamine* (Brassicaceae)—conversion of nuclear ribosomal ITS sequences in statu nascendi. *Theor Appl Genet* **98**: 831–834.
- Galtier N, Gouy M, Gautier C (1996). SEAVIEW and PHYLO-WIN: two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci* **12**: 543–548.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41**: 95–98.
- Hemleben V, Zentgraf U (1994). Structural organisation and regulation of transcription by RNA polymerase I of plant nuclear ribosomal genes. In: Nover L (ed). *Results and Problems in Cell Differentiation 20: Plant Promoters and Transcription Factors*. Springer-Verlag: Berlin/Heidelberg. pp 3–24.
- Joly S, Rauscher JT, Sherman-Broyles SL, Brown AH, Doyle JJ (2004). Evolutionary dynamics and preferential expression of homeologous 18S-5.8S-26S nuclear ribosomal genes in natural and artificial glycine allopolyploids. *Mol Biol Evol* **21**: 1409–1421.
- Keller I, Chintauan-Marquier IC, Veltsos P, Nichols RA (2006). Ribosomal DNA in the grasshopper *Podisma pedestris*: escape from concerted evolution. *Genetics* **174**: 863–874.
- Kovarik A, Dadejova M, Lim YK, Chase MW, Clarkson JJ, Knapp S *et al.* (2008a). Evolution of rDNA in nicotiana allopolyploids: a potential link between rDNA homogenization and epigenetics. *Ann Bot (Lond)* **101**: 815–823.
- Kovarik A, Pires JC, Leitch AR, Lim KY, Sherwood AM, Matyasek R *et al.* (2005). Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* **169**: 931–944.
- Kovarik A, Werlemark G, Leitch AR, Souckova-Skalicka K, Lim YK, Khaitova L *et al.* (2008b). The asymmetric meiosis in pentaploid dogroses (*Rosa* sect. *Caninae*) is associated with a skewed distribution of rRNA gene families in the gametes. *Heredity* **101**: 359–367.
- Leach CR, Houben A, Field B, Pistrick K, Demidov D, Timmis JN (2005). Molecular evidence for transcription of genes on a B chromosome in *Crepis capillaris*. *Genetics* **171**: 269–278.
- Lim KY, Werlemark G, Matyasek R, Bringloe JB, Sieber V, El Mokadem H *et al.* (2005). Evolutionary implications of permanent odd polyploidy in the stable sexual, pentaploid of *Rosa canina* L. *Heredity* **94**: 501–506.
- Matyasek R, Tate JA, Lim YK, Srubarova H, Koh J, Leitch AR *et al.* (2007). Concerted evolution of rDNA in recently formed *Tragopogon* allotetraploids is typically associated with an inverse correlation between gene copy number and expression. *Genetics* **176**: 2509–2519.
- McTaggart SJ, Dudycha JL, Omilian A, Crease TJ (2007). Rates of recombination in the ribosomal DNA of apomictically propagated *Daphnia obtusa* lines. *Genetics* **175**: 311–320.
- Navashin M (1934). Chromosomal alterations caused by hybridization and their bearing upon certain general genetic problems. *Cytologia* **5**: 169–203.
- Nieto Feliner G, Gutierrez Larena B, Fuertes Aguilar J (2004). Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in *Armeria* (Plumbaginaceae). *Ann Bot (Lond)* **93**: 189–200.
- Nieto Feliner G, Rossello JA (2007). Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Mol Phylogenet Evol* **44**: 911–919.
- Nilsson Ö (1967). Drawings of Scandinavian plants 3–4 *Rosa* L. *Bot Notiser* **120**: 137–143.
- Nybom H, Esselink GD, Werlemark G, Leus L, Vosman B (2006). Unique genomic configuration revealed by microsatellite DNA in polyploid dogroses, *Rosa* sect. *Caninae*. *J Evol Biol* **19**: 635–648.
- Nybom H, Esselink GD, Werlemark G, Vosman B (2004). Microsatellite DNA marker inheritance indicates preferential pairing between two highly homologous genomes in polyploid and hemisexual dog-roses, *Rosa* L. Sect. *Caninae* DC. *Heredity* **92**: 139–150.
- Preuss S, Pikaard CS (2007). rRNA gene silencing and nucleolar dominance: insights into a chromosome-scale epigenetic on/off switch. *Biochim Biophys Acta* **1769**: 383–392.
- Richards EJ (2006). Inherited epigenetic variation—revisiting soft inheritance. *Nat Rev Genet* **7**: 395–401.
- Ritz CM, Schmutz H, Wissemann V (2005). Evolution by reticulation: European dogroses originated by multiple hybridization across the genus *Rosa*. *J Hered* **96**: 4–14.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003). DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.
- Salmon A, Ainouche ML, Wendel JF (2005). Genetic and epigenetic consequences of recent hybridization and polyploidy in *Spartina* (Poaceae). *Mol Ecol* **14**: 1163–1175.
- Sang T, Crawford DJ, Stuessy TF (1995). Documentation of reticulate evolution in peonies (*Paeonia*) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proc Natl Acad Sci USA* **92**: 6813–6817.
- Schlotterer C, Tautz D (1994). Chromosomal homogeneity of *Drosophila* ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution. *Curr Biol* **4**: 777–783.
- Smith GP (1974). Unequal crossover and the evolution of multigene families. *Cold Spring Harb Symp Quant Biol* **38**: 507–513.
- Tackholm G (1920). On the cytology of the genus *Rosa*. *Svensk Bot Tidskrift* **14**: 300–311.
- Volkov RA, Komarova NY, Hemleben V (2007). Ribosomal DNA in plant hybrids: inheritance, rearrangement, expression. *Syst Biodivers* **5**: 261–276.
- Wang J, Tian L, Lee HS, Wei NE, Jiang H, Watson B *et al.* (2006). Genomewide nonadditive gene regulation in Arabidopsis allotetraploids. *Genetics* **172**: 507–517.
- Werlemark G, Nybom H (2001). Skewed distribution of morphological character scores and molecular markers in three interspecific crosses in *Rosa* sect. *Caninae*. *Heredity* **134**: 1–13.
- Wissemann V (2003a). Conventional taxonomy (wild roses). In: Roberts AV, Debener T, Gudin S (eds). *Encyklopedia of Rosa Science*. Elsevier Academic Press: Oxford. pp 111–117.
- Wissemann V (2003b). Hybridization and the evolution of the nrITS spacer region. In: SAAs A (ed). *Plant Genome: Biodiversity and Evolution*, vol. 1 part A. Science Publishers: Enfield.

Supplementary Information accompanies the paper on Heredity website (<http://www.nature.com/hdy>)