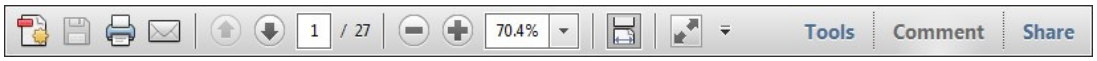
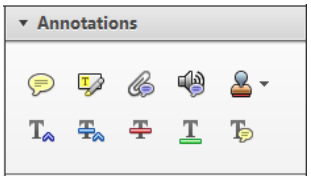
Once you have Acrobat Reader open on your computer, click on the Comment tab at the right of the toolbar:



This will open up a panel down the right side of the document. The majority of tools you will use for annotating your proof will be in the Annotations section, pictured opposite. We've picked out some of these tools below:

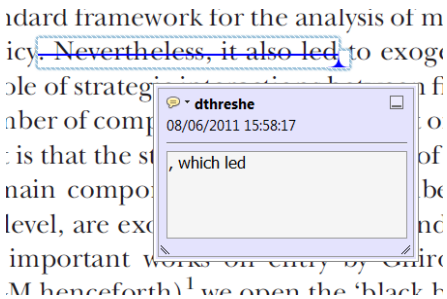


1. Replace (Ins) Tool – for replacing text.

Strikes a line through text and opens up a text box where replacement text can be entered.

How to use it

- Highlight a word or sentence.
- Click on the Replace (Ins) icon in the Annotations section.
- Type the replacement text into the blue box that appears.

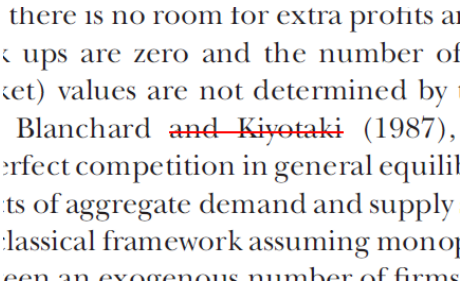


2. Strikethrough (Del) Tool – for deleting text.

Strikes a red line through text that is to be deleted.

How to use it

- Highlight a word or sentence.
- Click on the Strikethrough (Del) icon in the Annotations section.

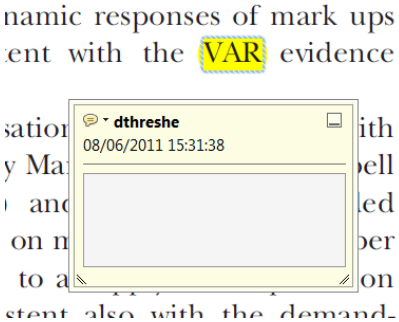


3. Add note to text Tool – for highlighting a section to be changed to bold or italic.

Highlights text in yellow and opens up a text box where comments can be entered.

How to use it

- Highlight the relevant section of text.
- Click on the Add note to text icon in the Annotations section.
- Type instruction on what should be changed regarding the text into the yellow box that appears.

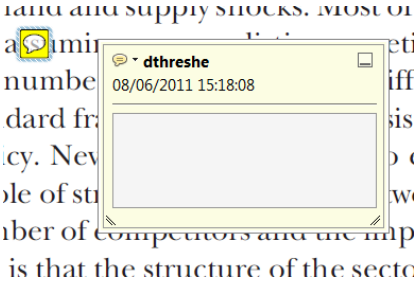


4. Add sticky note Tool – for making notes at specific points in the text.

Marks a point in the proof where a comment needs to be highlighted.

How to use it

- Click on the Add sticky note icon in the Annotations section.
- Click at the point in the proof where the comment should be inserted.
- Type the comment into the yellow box that appears.

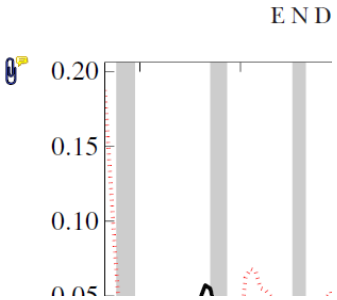


5. Attach File Tool – for inserting large amounts of text or replacement figures.

Inserts an icon linking to the attached file in the appropriate pace in the text.

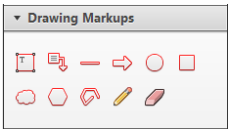
How to use it

- Click on the Attach File icon in the Annotations section.
- Click on the proof to where you'd like the attached file to be linked.
- Select the file to be attached from your computer or network.
- Select the colour and type of icon that will appear in the proof. Click OK.



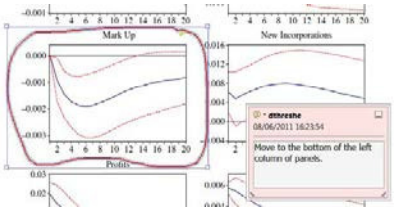
6. Drawing Markups Tools – for drawing shapes, lines and freeform annotations on proofs and commenting on these marks.

Allows shapes, lines and freeform annotations to be drawn on proofs and for comment to be made on these marks.



How to use it

- Click on one of the shapes in the Drawing Markups section.
- Click on the proof at the relevant point and draw the selected shape with the cursor.
- To add a comment to the drawn shape, move the cursor over the shape until an arrowhead appears.
- Double click on the shape and type any text in the red box that appears.



New Phytologist

Research

When fathers are instant losers: homogenization of rDNA loci in recently formed *Cardamine* × *schulzii* trigenomic allopolyploid

Judita Zozomová-Lihová[1]*, Terezie Mandáková[2]*, Alena Kovaříková[3], Andreas Mühlhausen[4], Klaus Mummenhoff[4], Martin A. Lysak[2] and Aleš Kovařík[3]

[1]Institute of Botany, Slovak Academy of Sciences, Bratislava, Slovakia; [2]RG Plant Cytogenomics, CEITEC, Masaryk University, Brno, Czech Republic; [3]Department of Molecular Epigenetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic; [4]Department of Biology, University Osnabrueck, Osnabrueck, Germany

Author for correspondence:
Aleš Kovařík
Tel: +420 541517178
Email: kovarik@ibp.cz

Received: *25 February 2014*
Accepted: *29 April 2014*

New Phytologist (2014)
doi: 10.1111/nph.12873

Key words: Brassicaceae, concerted evolution, hybridization, hybrids, polyploidy, rDNA.

Summary

• Recently formed allopolyploids represent an excellent system to study the impacts of hybridization and genomic duplication on genome structure and evolution. Here we explored the *35S* rRNA genes (rDNA) in the *Cardamine* × *schulzii* allohexaploid that was formed by two subsequent hybridization events within the past *c*. 150 yr.

• The rDNA loci were analyzed by cloning, next generation sequencing (NGS), RT-PCR and FISH methods.

• The primary *C.* × *insueta* triploid hybrid derived from *C. rivularis* (♀) and *C. amara* (♂) had gene ratios highly skewed towards maternal sequences. Similarly, *C.* × *schulzii,* originating from the secondary hybridization event involving *C.* × *insueta* (♀) and *C. pratensis* (♂), showed a reduction in paternal rDNA homeologs despite an excess of chromosomes inherited from *C. pratensis*. We also identified novel rDNA loci in *C.* × *schulzii,* suggesting that lost loci might be slowly reinstalled by translocation (but not recombination) of genes from partner genomes.

• Prevalent clonal propagation of allopolyploids, *C.* × *insueta* and *C.* × *schulzii*,indicates that concerted evolution of rDNA may occur in the absence of extensive meiotic cycles. Adoption of NGS in rDNA variant analysis is highly informative for deciphering the evolutionary histories of allopolyploid species with ongoing homogenization processes.

Introduction

Allopolyploidy resulting from the merging of two or more similar yet diverged genomes may induce a number of genomic changes that can occur rapidly within a few generations (Doyle *et al.*, 2008; Leitch & Leitch, 2008). This 'genomic shock' (McClintock, 1984) generates variants, from which selection favors those with enhanced fertility (Xiong *et al.*, 2011; Birchler, 2012; Veitia *et al.*, 2013). Although most allopolyploids are ancient (Soltis & Soltis, 2009), there are a few polyploids that have formed relatively recently, that is, within the last 200 yr, such as *Spartina anglica* (Ainouche *et al.*, 2004), *Senecio cambrensis* and *S. eboracensis* (Abbott & Lowe, 2004), *Tragopogon mirus* and *T. miscellus* (Soltis *et al.*, 2004), *Mimulus peregrinus* (Vallejo-Marin & Lye, 2013) and *Cardamine* × *insueta*/*C.* × *schulzii* (Urbanska *et al.*, 1997). These neoallopolyploid systems differ in their histories, origins, genomic and karyotypic stabilities, and evolutionary successes (Weiss-Schneeweiss *et al.*, 2013).

The origin of *Cardamine* allopolyploids in the Alpine valley of Urnerboden (Switzerland) was associated with changes in land-use management at the end of the 19th Century, of which, continuous farming, intensive deforestation, drainage and formation of human-made hay meadows and pastures (Zimmerli, 1986; Urbanska *et al.*, 1997) were the most prominent. Thus, ecological processes apparently facilitated the formation and establishment of *C.* × *insueta* (RRA), a triploid ($2n = 24$) hybrid of diploid *C. rivularis* auct. non Schur (further referred to as *C. rivularis*; maternal genome donor) and *C. amara* (both $2n = 16$) (Urbanska-Worytkiewicz, 1977). *Bona fide* hexaploid plants ($2n = 48$, RRRRAA, *C.* × *schulzii*) were assumed to have evolved from *C.* × *insueta* by autopolyploidization shortly after the origin of *C.* × *insueta,* when habitat changes further intensified (Urbanska & Landolt, 1999). Recent studies re-evaluated the origin of *C.* × *schulzii* and showed the involvement of a third species, *C. pratensis*, which is phylogenetically closely related to *C. rivularis.* Indeed, *C.* × *schulzii* was shown to be a hypohexaploid trigenomic hybrid of *C.* × *insueta* (maternal genome donor, RRA) and *C. pratensis* ($2n = 30$, PPPP) with 46 chromosomes and a PPPPRA genomic constitution (Mandakova *et al.*, 2013; Zozomová-Lihová *et al.*, 2014).

In plants, the *35S* rDNA locus is composed of highly repetitive units encoding *18S*, *5.8S* and *26S* ribosomal RNA (rRNA) and is cytogenetically recognized as the nucleolar organizer region (NOR). Each 8–14 kb-long unit contains coding regions, internally transcribed spacers (ITS) and intergenic spacers (IGS) (Hemleben *et al.*, 1988). As the spacers diverge at relatively high

*These authors contributed equally to this work.

rates, sequence analysis of these regions is sufficient for resolving species relationships within most genera (Alvarez & Wendel, 2003; Nieto Feliner & Rossello, 2007). In plant genomes there may be several thousand units, whose integrity and homogeneity is maintained by a process called concerted evolution (Brown *et al.*, 1972; Dover, 1982; Eickbush & Eickbush, 2007; Nieto Feliner & Rossello, 2012). However, plants usually exhibit multiple ITS families and this variability has been attributed to frequent hybridization events and the presence of multiple rDNA loci on chromosomes. For unknown reasons, allopolyploid repeats have reduced complexity across parental genomes, and it is not uncommon to find a single repeat in advanced allopolyploids that formed several thousand years ago (Wendel *et al.*, 1995; Volkov *et al.*, 1999; Kotseruba *et al.*, 2003; Kovarik *et al.*, 2004; Rauscher *et al.*, 2004; Bao *et al.*, 2010; Weiss-Schneeweiss *et al.*, 2012). Patterns of concerted evolution have also been observed in recently formed polyploids. Populations of *Tragopogon mirus* and *T. miscellus* allotetraploids that formed < 80 yr ago already show a reduction in gene copy number from one of the parents, although complete homogenization has not yet been observed (Kovarik *et al.*, 2005). In *S. cambrensis*, only one ITS repeat was identified in one population, whereas other populations have retained both homeologs (Lowe & Abbott, 1996). These observations, together with studies on synthetic allopolyploid lines (Skalicka *et al.*, 2003; Pontes *et al.*, 2004; Malinska *et al.*, 2010; Ksiazczyk *et al.*, 2011), suggest that rDNA re-arrangements can occur from the time of formation of the allopolyploid nucleus.

In *C.* × *schulzii* and *C.* × *insueta*, ITS sequence variation has been previously investigated in a few individuals by direct sequencing and by inspecting additive patterns and intra-individual polymorphic sites (Franzke & Mummenhoff, 1999). In the hybrids, most of the nucleotide positions which distinguish the assumed parents (*C. rivularis* and *C. amara*; *C. pratensis* was not considered) were converted to the maternal (*C. rivularis*) type, suggesting homogenization of rDNA. However, the tri-parental nature of the *C.* × *schulzii* genome complicates the interpretation of these initial findings. With this caveat in mind, we analyzed the rDNA repeats in both *Cardamine* hybrids and their progenitor genomes. We employed multiple molecular, genomic and cytogenetic approaches to determine intragenomic heterogeneity, the expression and chromosomal localization of rDNA arrays in the progenitors, and to trace the evolution of such arrays in the allopolyploid genomes. We also analyzed herbarium specimens of both *C.* × *insueta* and *C.* × *schulzii,* collected shortly after their discovery in the early 1970s, and compared these ancient ribotype patterns with those of the present day. We found evidence for concerted evolution starting to homogenize rDNA repeats at the triploid stage, and that typically, in both hybrids the paternally inherited genes and loci were eliminated.

Materials and Methods

Plant material

Information about the study site in the Urnerboden Alpine valley was obtained from Mandakova *et al.* (2013). Plant material was collected in the field, and from plants sampled in the field and cultivated in a glasshouse. The list of analyzed accessions is given in Supporting Information Table S1. For DNA isolation, leaf material from individual plants was either freshly collected or dried in silica gel. For chromosome preparations, whole young inflorescences were fixed overnight in freshly prepared ethanol : acetic acid (3 : 1), transferred to 70% ethanol and stored at −20°C until required. Apart from the individuals sampled from the present-day populations, we also analyzed herbarium specimens that were sampled in the 1970s–1980s in Urnerboden and deposited in the ZT herbarium (Zürich, Switzerland; Table S1).

PCR amplification, cloning and sequencing of ITS region

The ITS (ITS1-*5.8S*-ITS2) region of rDNA (Fig. 1a) was amplified using the protocol described in Mummenhoff *et al.* (1997). In total, 32 field-sampled individuals were analyzed by sequencing (3)5-20(30) clones per individual, summing 370 cloned sequences (Table S1). Direct sequences of eight individuals (two each of *C. amara*, *C. rivularis*, *C.* × *insueta* and *C.* × *schulzii*) were also obtained. In addition, four old herbarium specimens from the ZT herbarium were sequenced, using 9–16 clones per specimen, totaling 47 cloned sequences (Table S1). All sequences from cloning experiments were deposited in GenBank (accession numbers KF987809-KF988135).

Analysis of ITS cloned sequences

The cloned ITS sequences were aligned in Geneious 6.1.5 (Biomatters Ltd, New Zealand). The Neighbor-Net method, which generates split networks, was applied, based on uncorrected p-distances and using the SplitsTree 4.13.1 software (Huson & Bryant, 2006). Genetic diversity of the ITS sequences was assessed by calculating gene diversity (*H*) and nucleotide diversity (π_n) using the Arlequin 3.5 software (Excoffier *et al.*, 2005).

Chromosome preparations

Mitotic chromosomal spreads were prepared from cells of young, diploid flower buds as described by Mandakova *et al.* (2013). Selected slides were post-fixed in 4% formaldehyde in 2 ×saline sodium citrate (SSC; 20 ×SSC: 3 M sodium chloride, 300 mM trisodium citrate, pH 7.0) for 10 min, washed in 2 ×SSC twice for 5 min, and dehydrated in an ethanol series (70%, 90% and 100%, 2 min each). If appropriate, the slides were treated with 0.1 mg ml^{-1} pepsin (Sigma) in 0.01 M HCl at 37°C for 5 min, and post-fixed and dehydrated as already described.

DNA probes

Arabidopsis thaliana BAC clone T15P10 (AF167571), bearing *35S* rRNA gene repeats, was used for chromosomal *in situ* localization of NORs. An rDNA probe was labeled with Cy3-dUTP, whereas genomic DNAs (gDNA) were labeled with either biotin-dUTP or digoxigenin-dUTP via nick translation, as described by Lysak & Mandakova (2013). For Southern hybridization, the

LOW RESOLUTION FIG

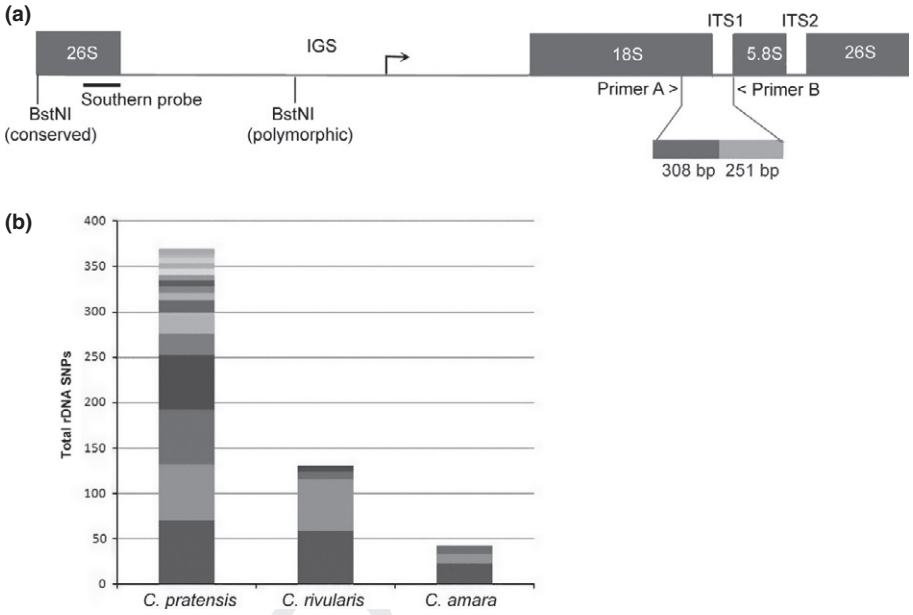


Fig. 1 Scheme of the rDNA unit (a) and intragenomic rDNA polymorphism based on next generation sequencing (NGS) reads (b). In (a) positions of 454 sequencing primers, lengths of the mapped region and restriction sites used for Southern blot hybridization are shown. (b) Graph showing the number and proportion of major (> 5%) SNPs among the NGS reads of *Cardamine* parental species.

26S rDNA probe was a 220 bp-long PCR product containing the 3′ end of the *C. amara 26S* rRNA gene (Fig. 1a). Primers were as follows: 26S_forward: 5′-GAA TTCACCCAAGTGTTGGGA T-3′; 26S_reverse: 5′-AGAGGCGTTCAGTCATAATC-3′.

FISH, GISH and Southern hybridization

In situ hybridization followed the procedures described in Mandakova *et al.* (2013). rDNA FISH was combined with GISH, where 100 ng of labeled *35S* rDNA probe and 400 ng of each labeled gDNA were used. Following hybridization, stringent washing was carried out three times for 5 min in 20% formamide in 2 ×SSC at 42°C. Biotin-labeled probes were detected by avidin–Texas Red (Vector Laboratories, UK), and signals were amplified by biotinylated goat anti-avidin (Vector Laboratories) and avidin–Texas Red. Digoxigenin-labeled probes were detected by mouse anti-digoxigenin (Jackson Immuno-Research) and goat anti-mouse–Alexa Fluor 488 (Molecular Probes, USA) antibodies. Chromosomes were counterstained with 2 μg ml^{-1} DAPI in Vectashield. Fluorescence signals were analyzed and photographed using an Olympus BX-61 epifluorescence microscope and CoolCube CCD camera (MetaSystems). Individual images were merged and processed using Photoshop CS software (Adobe Systems).

Southern blotting followed the protocol described by (Kovarik *et al.*, 2005) using probes labeled with [α-^{32}P]dCTP (Izotop, Hungary) in a random-primed reaction (DekaPrime kit, Fermentas, Lithuania).

RNA isolation and reverse transcription

Total RNA from fresh leaf tissue was isolated using an RNeasy kit (Qiagen) following the protocol supplied by the manufacturer. RNA was measured using a spectrophotometer and quality checked by gel electrophoresis. DNA contamination was removed using DNase (Turbo DNA free; Ambion). The reverse transcription mixture (20 μl) typically contained 1 μg RNA, 2 pmol of random nonamer primers, 10 nmol of each dNTP and 200 U of the enzyme (Superscript II RT; Invitrogen). The RT-CAPS analysis involved amplification of the *18S*-ITS1-5.*8S* region using the 18S_for and 5.8_rev primers described by Matyasek *et al.* (2007). The PCR products were digested with diagnostic restriction enzymes and fragments were separated on a 1.2% agarose gel.

Four-hundred and fifty-four sequencing, clustering and mutation analysis

Emulsion PCR of the *18S*-ITS1-*5.8S* subregion (for primer positions, see Fig. 1a) has been described in Matyasek *et al.* (2012). The reads obtained from amplicon sequencing (Roche 454 GS-FLX; EuroFins MWG, Germany) were sorted into two groups comprising primer A and primer B sequences. The reads were then sorted into clusters containing sequences with 100% identity over the aligned regions. The mapped regions included 251–252 bp of ITS1 and 308 bp of the *18S* gene (3′ end). All alignments were performed using the Smith-Waterman algorithm with default parameters, with an identity threshold E-value of e^{-100} and a maximum of 10 mutations per read. Mutation analysis was carried out using the Smith Waterman algorithm implemented within the MIRA platform (Eurofins MWG). The single nucleotide polymorphisms (SNPs) were detected in the 454SNP/454HCDiff.files retrieved from the GsMapper. On average, *c.* 400 clusters were obtained from each sample. Because a significant proportion of the errors observed on this instrument were small indels, mostly arising from inaccurate homopolymer lengths, reads with indels were grouped with similar reads without indels. The estimated error rate of Roche 454 GS-FLX platform is 10^{-3}–10^{-4} substitutions (Kircher & Kelso, 2010). This means that every 4th to 40th read mapped to a *c.* 250-bp sequence is likely to contain a technical error. Only highly confident clusters were therefore considered for analysis of variants:

singletons and low copy clusters with < 10% reads were disregarded. For phylogenetic analyses, representative sequences from ITS1 clusters were extracted, adjusted to the same lengths (500 bp), aligned, and used in a construction of a neighbor-joining tree (SeaView program, Gouy *et al.*, 2010).

The 454 read libraries available for *C. amara*, *C. rivularis* and *C. pratensis* (Mandakova *et al.*, 2013) were searched for the *18S*-ITS1 sequences using stand-alone BLAST. Returned sequences were extracted using tools implemented within the CLC Genomics workbench software. SNP and DIP (insertion/deletion) analyses were carried out using the GLC Genomics Workbench with the following parameters/requirements: window length of 11 bp, maximum of 2 gaps, a minimum coverage of 10, variants should occur at a minimum frequency of 0.05, with a maximum of 1000 variants expected.

Results

Diversity of ITS variants obtained from cloned sequences

The alignment of 417 cloned ITS1-*5.8S*-ITS2 sequences (*c.* 621 bp each), comprising 269 polymorphic positions and 249 distinct sequences, indicated high ITS heterogeneity. Identical clones were identified within each taxon, both within and amongst individuals, and also between each pair of the following taxa: *C. rivularis*, *C. pratensis*, *C.* × *insueta* and *C.* × *schulzii,* as well as between *C. amara* and *C.* × *insueta*. The greatest diversity of ITS sequences was recorded in *C. pratensis*, followed by the two allopolyploids, whereas the diploids, *C. amara* and *C. rivularis,* displayed minor variations (see Table 1). Direct sequencing of the selected individuals (Table S1) yielded electropherograms with signal overlaps, reinforcing the intra-individual variation that was observed among the ITS clones. Whereas direct sequences of *C. amara* and *C. rivularis* displayed only polymorphic nucleotide sites (double peaks), those of *C.* × *insueta* and *C.* × *schulzii* were unreadable due to polymorphisms for 1–2 bp-long indels.

Identification of rDNA variants in progenitor species based on NGS

Blast searches of NGS libraries of sequences from progenitor genomes yielded *c.* 600 reads homologous to a *c.* 700-bp fragment of rDNA comprising the *18S* (3′ end) -ITS1 sequence (Fig. 1a). This would roughly correspond to 550, 800 and 1200 gene copies in *C. amara*, *C. rivularis* and *C. pratensis*, respectively, based on read representation and genome size (1C is 217 Mb in *C. amara*, 380 Mb in *C. rivularis* and 754 Mb in *C. pratensis* (Mandakova *et al.*, 2013)). Mapped reads were subjected to mutation analysis. Among the progenitor genomes, there were 37 polymorphic sites, four occurring in the coding region and 33 in the ITS1 region (Table S2). Thirteen sites overlapped with those previously reported (Franzke & Mummenhoff, 1999) whereas there were 24 novel sites. Most polymorphisms occurred in the ITS region, but *C. pratensis* also exhibited additional polymorphisms in the *18S* coding region. *Cardamine amara* differed from *C. rivularis*/*C. pratensis* in 16 positions, whereas *C. rivularis* and *C. pratensis* could be distinguished by frequencies of individual SNPs. There were 16 highly confident (occurring in > 5% reads) polymorphic sites in *C. pratensis*, four in *C. rivularis* and three in *C. amara*. *Cardamine pratensis* also possessed a large number of high (40–60%) frequency SNPs, indicating the presence of abundant gene families in this genome. A quantitative representation of SNPs among reads reflecting rDNA array homogeneity (Fig. 1b) revealed decreasing intragenomic rDNA homogeneity among the species as follows: *C. amara* > *C. rivularis* ≫ *C. pratensis*.

Phylogenetic relationships between progenitor species based on ITS cloned sequences, species-specific ITS variants and the position of ITS clones from allopolyploids

A Neighbor-Net graph revealed three main groups, denoted here as the *C. amara* group, the *C. rivularis* group, and the *C. rivularis*+*C. pratensis* group (marked as A-, R- and P-genome groups, Fig. 2). All *C. amara* clones were placed within the *C. amara* group; however, although most *C. rivularis* clones were within the *C. rivularis* group (71%), some were also present in the *C. rivularis*+*C. pratensis* group (29%). All but one clone of *C. pratensis* were within the *C. rivularis*+*C. pratensis* group. Within the *C. rivularis*+*C. pratensis* group, a subgroup was identified, which comprised most of the *C. pratensis* clones (72%) but none from *C. rivularis*. *Cardamine pratensis* and *C. rivularis* shared some ITS sequences, but *C. rivularis*- and *C. pratensis*-specific ITS classes could also be identified. However, *C. amara* displayed a distinct, species-specific class of ITS repeats.

Clones retrieved from *C.* × *insueta* were mostly placed within the *C. rivularis* group (56%) and the *C. rivularis*+*C. pratensis* group (29%), with very few clones assigned to the *C. pratensis* subgroup (2%) or *C. amara* group (6%). Clones from

Table 1 Diversity of the internal transcribed spacer (ITS) cloned sequences

Taxon (ploidy)	No. of ind.	No. of clones/distinct sequence types	*H* (± SD)	π_n (± SD)	Group assignment
C. amara (2x)	6	58/31	0.905 (0.0303)	0.00403 (0.002447)	A
C. rivularis (2x)	6	38/18	0.899 (0.0436)	0.00555 (0.003221)	R, R+P
C. pratensis (4x)	6	64/54	0.992 (0.0055)	0.01536 (0.007911)	R+P, (R)
C. × *insueta* (3x)	9	135/97	0.943 (0.0159)	0.01531 (0.007825)	A, R, R+P, outl.
C. × *schulzii* (6x)	5	75/50	0.945 (0.0206)	0.01415 (0.007318)	R, R+P, outl.

H, gene diversity; π_n, nucleotide diversity; SD, standard deviation. Group assignment: A, A-genome group; R, R-genome group; R+P, R+P-genome group; outl., outlying (unassigned) positions (see Fig. 2). Sequences of old herbarium specimens are not included here.

COLOR

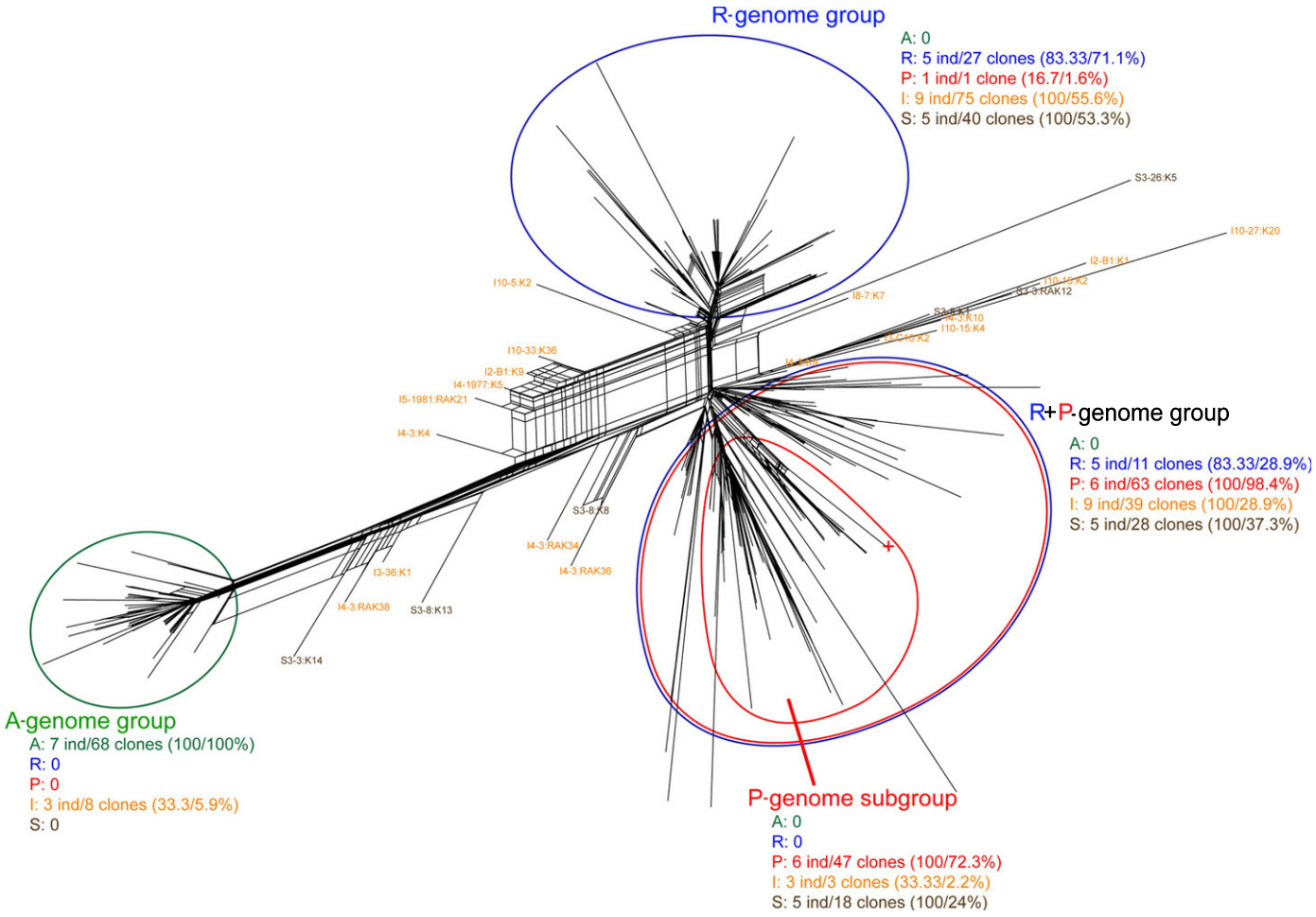


Fig. 2 Neighbor-Net split graph based on internal transcribed spacer (ITS) cloned sequences (36 individuals, 417 sequences, see Supporting Information Table S1 for the accession list). Three main groups are highlighted, specifying the number of individuals and clones included (the proportion of the total number of individuals and clones analyzed per taxon is in brackets): A, *Cardamine amara*; R, *C. rivularis*; P, *C. pratensis*; I, *C.* × *insueta*; S, *C.* × *schulzii*. Sequences from the herbarium specimens are included in the graph, but not in the statistics (see text).

C. × *schulzii* were placed within the *C. rivularis* (53%) and *C. rivularis*+*C. pratensis* groups (37%; also abundant within *C. pratensis*, 24%), but none were placed in *C. amara*. Numerous sequences of *C.* × *insueta*, and *C.* × *schulzii* also appeared in outlying positions, suggesting a signal conflict. These sequences displayed a certain level of recombination among those from the three main groups.

Internally transcribed spacers variation recorded among the analyzed herbarium specimens corresponded to that of the field-sampled individuals. Thus, all clones of ZT35744 (*C. amara*) were placed within *C. amara*; two clones each of ZT35751 and ZT35752 (both *C.* × *insueta*) were present in the *C. amara* group whereas the remaining 24 clones were placed in *C. rivularis* and *C. rivularis*+*C. pratensis*; and all clones of ZT35755 (*C.* × *schulzii*) were positioned within the *C. rivularis* and *C. rivularis*+*C. pratensis* groups.

Analysis of rDNA variants by NGS in allopolyploids

Because the parental species contributed to divergent rDNA family sequences, we carried out mutation analysis of PCR amplicons. The samples included present-day plants of *C.* × *insueta*, *C.* × *schulzii*, and a herbarium specimen of *C.* × *schulzii*, collected in 1974 (for accessions see Table S1). Reads were mapped to *C. amara* and *C. rivularis*/*C. pratensis* consensus sequences. The numbers of mapped reads are given in Table S3. The reads were then grouped into clusters with 100% sequence identity (Table S4). Table S5 shows major clusters comprising > 95% of highly confident reads (> 10 reads per cluster). The major clusters in each allopolyploid were formed by reads derived from *C. rivularis*/*C. pratensis*,whereas there were only a few clusters derived from *C. amara*. To determine the phylogenetic relationship between the dominant rDNA families, we extracted representative *18S*-ITS1 sequences from the top (high-frequency) clusters (Table S5). In the neighbor-joining tree (Fig. 3), the consensus sequences clustered into three well-separated groups. Group A contained sequences homologous to the major *C. amara* ribotypes, accounting for 10% and 8% of reads in *C.* × *insueta* and *C.* × *schulzii*, respectively (Table S5). Group P comprised four *C.* × *schulzii* clusters grouped with the most abundant *C. pratensis* ribotype, accounting for *c.* 35% of *C.* × *schulzii* reads. Group R comprised four clusters from

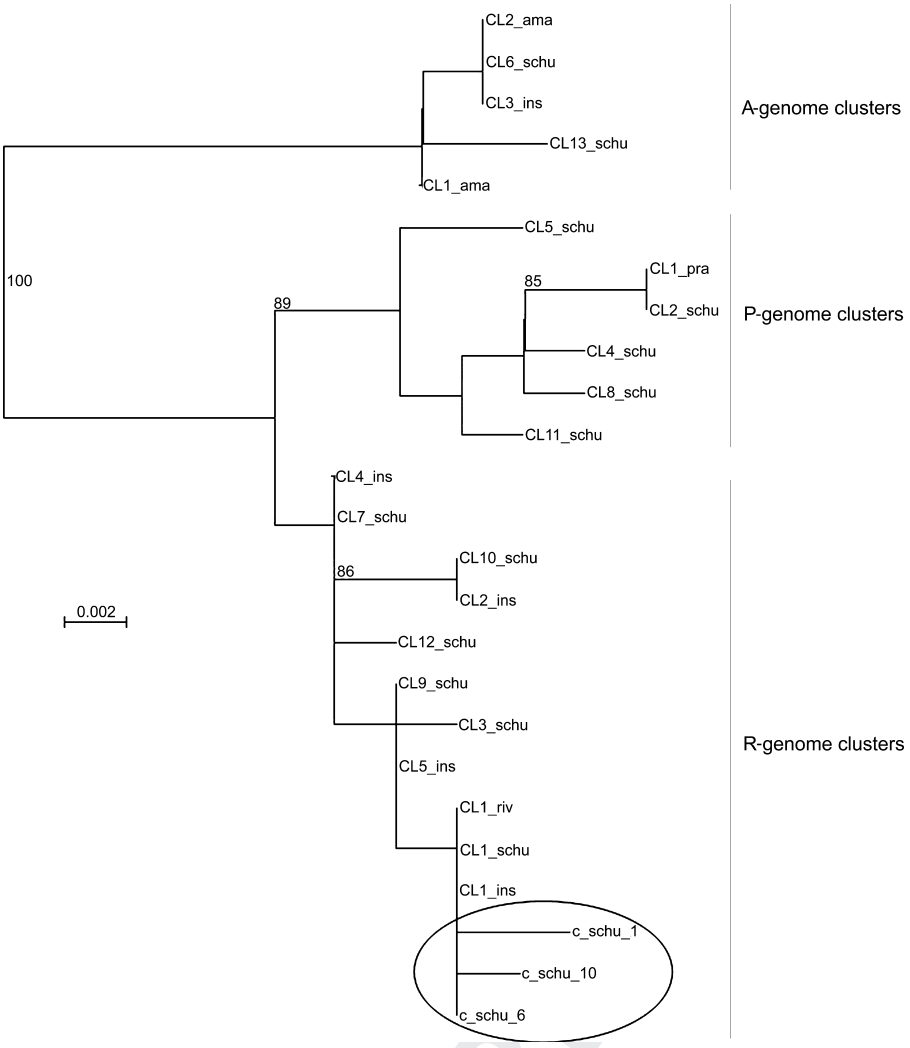


Fig. 3 Phylogenetic relationships between the most abundant rDNA variants in the parental species and allopolyploids. The neighbor-joining tree was constructed using sequences from the top clusters, representing > 95% of total rDNA reads. Branch support is indicated at each node and was calculated using 1000 bootstrap replicates. The read abundance decreased in the direction from CL1 to CLn. The CL1_ama, CL1_pra and CL1_riv represent the most abundant variants in parental species. Circled sequences represent clones amplified from cDNA of *Cardamine* × *schulzii*.

C. × *insueta* and three clusters from *C.* × *schulzii,* representing 58% and 90% of *C.* × *schulzii* and *C.* × *insueta* reads, respectively.

Most parental SNPs were found in hybrids and their frequency corresponded with cluster representation. For example, *Cardamine* × *schulzii* reads had 59.8% of A and 40.2% of C nucleotides at position +122 (ITS1) whereas this site was almost monomorphic in *C. rivularis* (A) and *C. amara* (C); *C. pratensis* had 76.6% of C and 23.4% of A at this site (Table S2). Similar midparental values of base calls were observed in hybrids at positions +125 (ITS1) and −102 (18S coding region), that is, for nucleotides distinguishing the parental species.

We analyzed individual variants in parental species and derived allopolyploids (Fig. 4; Tables S4, S5). The highly abundant ribotypes occurring in *C. amara*, *C. rivularis* and *C. pratensis* genomes were denoted as A, R and P, respectively. Fig. 4(a) shows representation of A-genome variants in *C. amara* and two derived allopolyploids. The major *C. amara* A1 variant was reduced in hybrids forming the seventh and eleventh cluster in *C.* × *insueta* and *C.* × *schulzii,* respectively. By contrast, the minor A2 variant of *C. amara* remained unchanged in copies and represented > 70% of the A-genome rDNA in both *C.* × *insueta* and *C.* × *schulzii*. This variant was slightly under-represented (*P* value < 0.01, Fig. 4a; Table S6) in the present-day *C.* × *schulzii* compared to the herbarium specimen. The A5 variant was a clear recombinant sequence containing 5′ ITS1 from *C. rivularis* and the 3′ ITS1 from *C. amara*. The A6 variant harbored a 14-bp deletion in the coding region as well as numerous mutations, and it probably originated from a nonfunctional pseudogene. Both minor A5 and A6 variants were only found in hybrids. Among the *C. rivularis*/*C. pratensis* variants, R1 was the most abundant type in both *C.* × *insueta* and *C.* × *schulzii* accounting for > 46% and 35% of their rDNAs, respectively (Fig. 4b). The R1 variant was under-represented (*P* value < 0.001, Fig. 4b; Table S6) in the herbarium sample compared to current plants of *C.* × *schulzii*. However, the R5 variant that was abundant in *C.* × *insueta,* was found at greatly reduced frequencies in both *C.* × *schulzii* samples. The *C. pratensis* P1 and P2 variants were almost equally represented in *C.* × *schulzii* (each accounting for *c.* 13% of rDNA) whereas they did not occur in *C.* × *insueta,* even among low copy clusters.

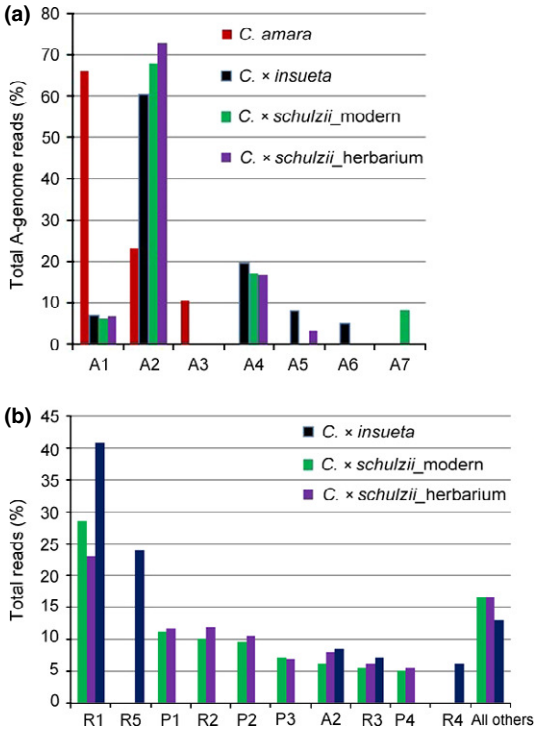


Fig. 4 Representation of rDNA variants in the parental and allopolyploid genomes. (a) A- genome variants in parental *Cardamine amara* and allopolyploid *C.* × *insueta*/*C.* × *schulzii*. (b) R-and P- genome variants in *C.* × *schulzii* and *C.* × *insueta*.

Population level analysis of rDNA patterns

We used Southern blot hybridization to investigate restriction site polymorphisms within rDNA units. Genomic DNAs of *C. amara* (three individuals, Table S1), *C. rivularis* (3), *C. pratensis* (4), *C.* × *insueta* (5) and *C.* × *schulzii* (6) were digested with *Bst*NI and hybridized on blots to the *26S* rDNA probe (Fig. 5). *Bst*NI had one conserved recognition site in the *26S* coding region and a second, more variable site in the IGS region (Fig. 1a). In *C. amara*, two or three bands were visualized after hybridization. The bottom band was always more intense than the doublet located in the upper part of the gel. Digests of *Cardamine rivularis* DNA contained a single band (R) which had slightly greater electrophoretic mobility than the major *C. amara* band. In *C. pratensis* digests, the probe hybridized with one or two bands (P) depending on the individual analyzed. The bottom P band was present in all accessions, whereas the upper band was more variable and only occurred in some accessions. In *C.* × *insueta* digests, the probe hybridized to two fragments, the minor upper fragment having the same mobility as that from *C. amara*, and the lower fragment corresponding to bands from *C. rivularis* and *C. pratensis*. Digests of *C.* × *schulzii* contained three bands that were inherited from all three parents. Two out of the three *C. amara* restriction fragments were not present in any accessions of *C.* × *insueta* and *C.* × *schulzii*. Hybridization intensity was estimated by measuring amounts of radioactivity in individual bands. In *C.* × *insueta*, the ratio between R/A bands was *c.* 10 : 1 (Table 2). In *C.* × *schulzii*, the ratio between R+P/A was nearly 20 : 1. The *C.* × *schulzii* herbarium specimen could not be analyzed because its DNA was too degraded to permit enzymatic restriction analysis.

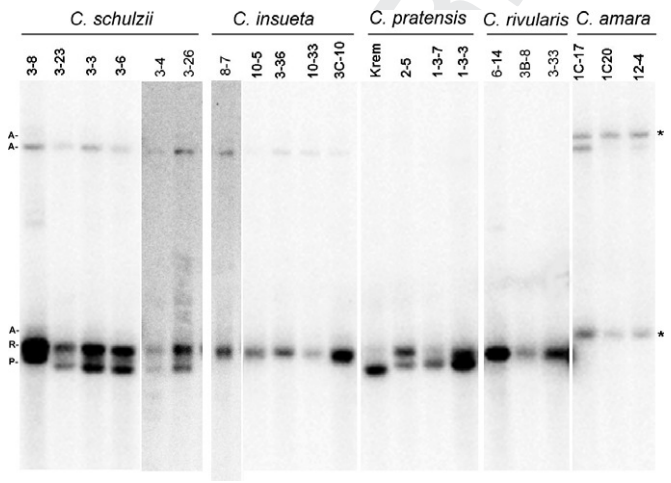


Fig. 5 Southern blot hybridization analysis of rDNA from *Cardamine* parental species and their hybrids. Species-specific bands are indicated. A, *C. amara*; R, *C. rivularis*; P, *C. pratensis*. *C. amara* fragments lost in hybrids are indicated by an asterisk.

Expression analysis of rDNA

In order to determine the level of rDNA expression in polyploids we used cleaved amplified polymorphic site analysis (CAPS). The *18S*-ITS1 region was amplified using the *18S* forward and *5.8S* reverse primers (RT-CAPS), yielding 700-bp products from both genomic and cDNA (Fig. 6a). The A- and R/P-genome variants could be discriminated based on *Taq*I restriction site polymorphisms. Consistent with the restriction map, *C. amara* had a specific, slowly migrating band in the 500 bp region, and a rapidly migrating 200-bp band (Fig. 6b). Both *C. rivularis*/*C. pratensis* exhibited a doublet of bands in the 200–300 bp region, the upper band being specific for both species. Control PCRs carried out on genomic DNA from hybrids showed the presence of bands inherited from the parents (left panel, Fig. 6b). In cDNA samples, only bands derived from *C. rivularis*/*C. pratensis* were visualized, whereas the upper *C. amara* bands were absent (right panel). *Sau*96I recognizes a T/C polymorphism in the *18S* coding region of *C. pratensis* and *C.* × *schulzii* rDNA (Fig. 6a). Consistent with the quantitative NGS analysis (Table S5) most of the product amplified from *C. pratensis* genomic DNA was digested by the enzyme (left panel, Fig. 6c) whereas products from the remaining species were undigested or only slightly digested (*C.* × *schulzii*). Products amplified from *C. pratensis* cDNA (right panel) were undigested, indicating that variants harboring a T>C mutation at −102 (Table S2) were not expressed although they were abundant in the genome. ITS clones amplified from *C.* × *schulzii* cDNA grouped with *C. rivularis*, but none of them grouped with *C. amara* or *C. pratensis* (Fig. 3).

New Phytologist

Table 2 Proportions of homeologous rRNA genes in *Cardamine* × *schulzii* and *C.* × *insueta* as determined by independent methods

		P+R-genome units (%)				A-genome units (%)			
	Genome composition	Expected[1]	Cloning	NGS	Southern blot	Expected	Cloning	NGS	Southern blot
C. × *insueta*	RAA	66	94[2]	86	94	24	6[2]	14	6
C. × *schulzii* modern	PPPPRA	85	100[3]	92	95	15	0[3]	8	5
C. × *schulzii* herbarium	PPPPRA	85	100	89	nd	15	0	11	nd

[1]Expected percentages were calculated as if there was an additivity in homeologous gene copies.
The gene copies were calculated from gene proportion of 454 reads in parental species.
[2]About 7% of cloned sequences were recombinants between the P+R and A repeat types (*in vitro* or natural recombination).
[3]About 4% of cloned sequences were recombinants between the P+R and A repeat types (*in vitro* or natural recombination).

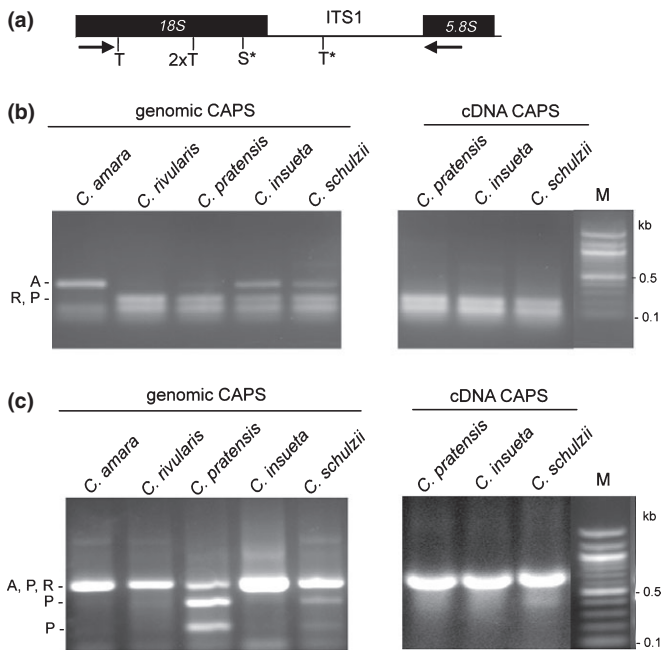


Fig. 6 Dominant expression of *Cardamine rivularis* rDNA units in allopolyploids. (a) Scheme of the region used for CAPS assay. T, *Taq*I sites; S, *Sau*96I site. Polymorphic sites are indicated by asterisks. (b, c) Gels show genomic CAPS (left) and cDNA-CAPS (right) profiles from the parental and hybrid genomes. The PCR products were digested with *Taq*I (b) and *Sau*96I (c). A, *C. amara*; R, *C. rivularis*; P, *C. pratensis*. In (b), note the absence of the *C. amara*-specific band in cDNA-CAPS of *C.* × *insueta*/*C.* × *schulzii* samples. Note the absence of digestion of amplification products with *Sau*96I in the cDNA CAPS panel (c). M, size markers; the strong bands are 500- and 1000-bp long, respectively (from the bottom). For accessions, see Supporting Information Table S1.

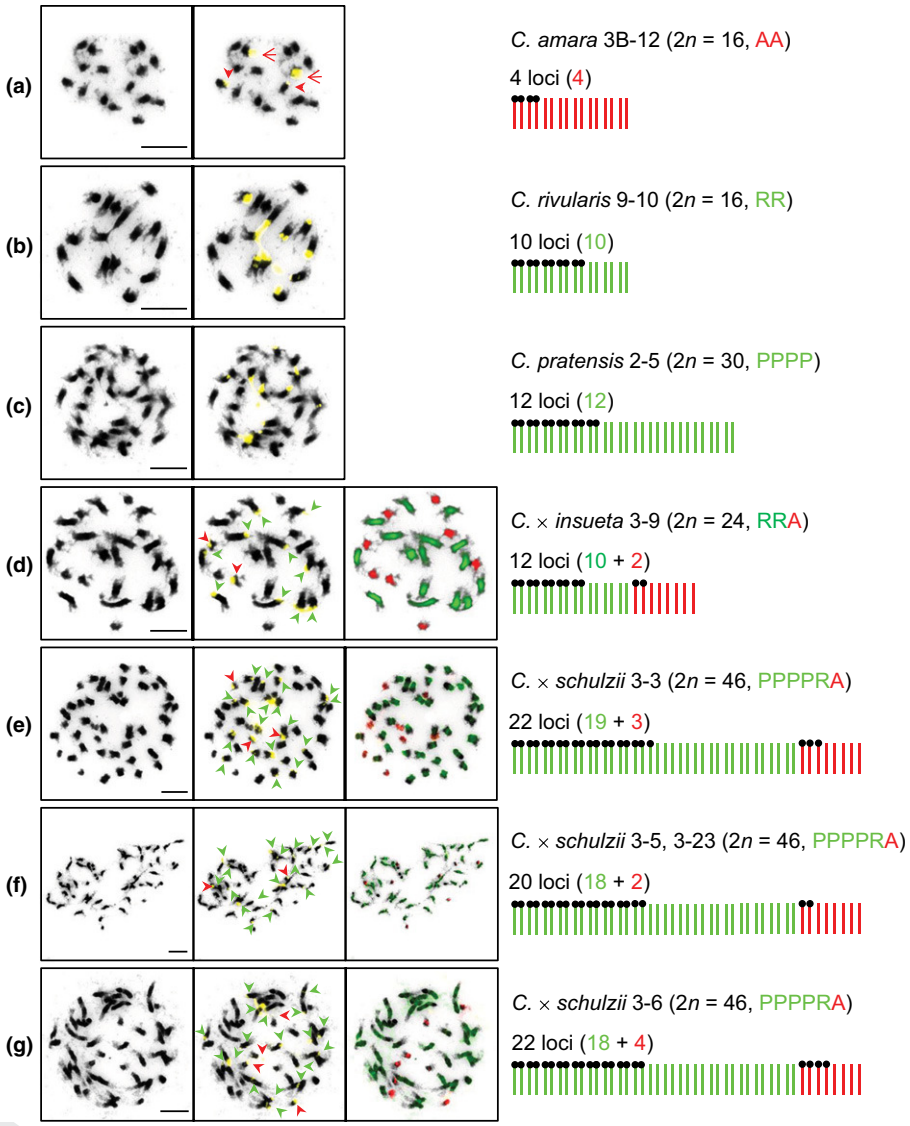
Chromosome localization of rDNA

We used FISH to determine the number of rDNA-bearing chromosomes in 10 individuals of *C. amara*, 16 individuals of *C. rivularis*, two individuals of *C. pratensis*, 27 individuals of *C.* × *insueta* and in four individuals of *C.* × *schulzii* (see Table S1 for accessions). In the parental species, the *35S* rDNA probe hybridized to 4 chromosomes in *C. amara*, and to 10 and 12 chromosomes in *C. rivularis* and *C. pratensis*, respectively (Fig. 7a–c). In the chromosomes of all three species, rRNA gene repeats occurred mostly at terminal positions, being heteromorphic, with larger and smaller (minor) loci. In order to characterize the NOR-bearing chromosomes in hybrids, we used a *35S* rDNA probe combined with parental gDNAs that identified A and R/P-genome chromosomes. We have shown previously (Mandakova *et al.*, 2013) that *C. rivularis* gDNA also labels *C. pratensis* chromosomes when used as a FISH probe. In *C.* × *insueta*, the rDNA probe hybridized to 12 sites (Fig. 7d). After GISH, 2 and 10 sites were located on A- and R-genome chromosomes, respectively. The additivity (10+2) of the *35S* rDNA loci was confirmed in the *C.* × *insueta* RRA-genome. However, the signals on A chromosomes were weak and major sites were absent compared to the parents. There were no differences between the number of rDNA loci in *C.* × *insueta* individuals.

In *C.* × *schulzii*, the expected number of rDNA loci, based on the PPPPRA genomic constitution, was 19, that is, 12 donated from *C. pratensis* and 7 from *C.* × *insueta* (5 R-genome and 2 A-genome loci). However, none of the accessions exhibited an additive number of loci. Instead, the number of loci varied from 20 to 22 among individuals (Fig. 7e–g). GISH indicated that all individuals had at least one extra locus on the R/P chromosomes, whereas two accessions (3-3 and 3-6) had 1 or 2 extra loci on A-genome chromosomes. Compared with parental species (Fig. 7a–c), the *35S* probe hybridized relatively weakly to the *C.* × *schulzii* chromosomes. Hence, many of its loci could be marked as 10 'minor'.

Discussion

Allopolyploids of recent ancestry represent unique systems for examining interactions between divergent rDNA repeats and the early effects of sequence homogenization. *Cardamine* allopolyploids formed < 150 yr ago, as documented recently by cytogenetic and molecular evidence (Mandakova *et al.*, 2013; Zozomová-Lihová *et al.*, 2014). In these allopolyploids, which are characterized by long generation times (long-lived perennials), efficient vegetative propagation, irregular meiosis and suppressed sexual reproduction (Urbanska-Worytkiewicz, 1977; Mandakova *et al.*, 2013), retention of divergent parental rDNA and low rDNA homogeneity were expected due to reduced recombination frequency. However, our results indicate a strong bias in their parental rDNA copy numbers, differential transmission of variants, expression dominance and increased NOR mobility.



COLOR

Fig. 7 Chromosome *in situ* localization of *35S* rDNA in *Cardamine* species. (a–c), *35S* rDNA probe (yellow signals) hybridized to mitotic chromosomes of *C. amara*, *C. rivularis* and *C. pratensis*. (d–g), *35S* rDNA probe (yellow signals) localized on mitotic chromosomes of *C.* × *insueta* and four accessions of *C.* × *schulzii*. The genomic origin of rDNA loci is indicated by red (A-genome) and green (R/P-genome) arrowheads based on GISH with gDNA of *C. amara* (red fluorescence) and *C. rivularis* (green fluorescence). Note that two large rDNA loci (arrows) in the A-genome were lost in both hybrids. Ideograms show the number and chromosomal position of localized rDNA sites (solid black circles). Chromosomes were counterstained by DAPI and inverted in Adobe Photoshop. Bars, 10 μm.

Phylogenetic relationships between parental species

The recently inferred phylogeny of *Cardamine* provided only poor resolution, most likely due to rapid speciation (Carlsen *et al.*, 2009). However, *C. amara* and *C. rivularis*/*C. pratensis* were shown to belong to two related clades, still with unresolved positions. The phylogenetic relationships (Figs 2, 3) between rDNA variants revealed clear distinctions between *C. amara* and the closely related species *C. rivularis* and *C. pratensis*. The latter two have been assigned to a single polyploid complex (Franzke & Hurka, 2000). The genomes are indeed closely related, as proven by cytogenetic and molecular data (Franzke & Hurka, 2000; Mandakova *et al.*, 2013; Zozomová-Lihová *et al.*, 2014). Here we show that both *C. rivularis* and *C. pratensis* share > 90% polymorphisms in the *18S*-ITS1, confirming that the species separated relatively recently, perhaps postglacially (Franzke & Hurka, 2000). However, deep sequencing revealed quantitative differences in representation of rDNA clusters, and cloning disclosed common as well as distinct rDNA variants amplified in the *C. rivularis* and *C. pratensis* genomes.

Both NGS and cloning revealed decreasing homogeneity of rDNA arrays in the taxon sequence: *C. amara* > *C. rivularis* ≫ *C. pratensis*. rDNA heterogeneity should be viewed in the context of ancient hybridization events and inherited ancestral polymorphisms. For *C. pratensis*, which exhibits multiple ploidy levels throughout Europe, a complex evolutionary history has been envisaged, affected by recurrent reticulations and multiple formations of polyploids (Marhold *et al.*, 2004). Variation in Southern blot profiles also suggests multiple origins and/or differential homogenization of rDNA families. Interestingly, the level of *18S* gene polymorphism in *C. pratensis* is higher than in most other species analyzed by NGS (Matyasek *et al.*, 2012; Straub *et al.*, 2012) and most variants are not expressed (Fig. 6c). Pseudogene abundance coincides with a large number of loci in this species. rDNA heterogeneity in *C. rivularis* is more difficult to explain because the species is considered to be a true diploid ($2n = 16$). A

possible explanation is the presence of multiple rDNA loci generating array diversity (Matyasek *et al.*, 2012). Indeed, although closely related *C. rivularis* and *C. pratensis* differ by 14 chromosomes, both species have a comparable number of *35S* rDNA loci – 10–12 (Fig. 7). Pleistocene histories of alpine species frequently indicate the presence of multiple periglacial refugia and colonization routes (Schönswetter *et al.*, 2005). It is feasible, therefore, that postglacial colonization from multiple refugia, accompanied by genetic admixture between divergent populations, could have caused increased ITS heterogeneity in *C. rivularis*.

Loss of paternal rDNA repeats in the triploid *C.* × *insueta*

It is estimated that *c.* 400 genes from *C. amara* rDNA (72%) were lost in the course of *C.* × *insueta* evolution. Based on the analysis of ITS clones, underrepresentation of *C. amara* rDNA repeats was already evident in *C.* × *insueta* herbarium specimens collected in 1977 and 1981, indicating that repeat loss occurred relatively early in the evolution of the hybrid. There were no major differences in the R/A homeolog gene ratios among the subpopulations, although shifts in variant frequency cannot be excluded because *C.* × *insueta* may have arisen multiple times (Zozomová-Lihová *et al.*, 2014). Quantitative variant analysis of NGS reads revealed that not all rDNA families were equally affected by deletions. Intriguingly, the major, and transcriptionally active, *C. amara* family was almost completely lost in hybrids, whereas only the minor family, apparently already silenced in the parent, was retained. Reduction of the number of rDNA repeats and loci is a common phenomenon in allopolyploids, and is usually explained by recombination events involving homologous or homeologous interactions in meiosis (Nieto Feliner & Rossello, 2012). However, *C. amara* chromosomes have no pairing partners in *C.* × *insueta* (Mandakova *et al.*, 2013). Homogenization of rDNA repeats in triploids is not unprecedented in the literature. For example, it has been shown that in triploid lizards, which reproduce by parthenogenesis, one parental rDNA array completely replaces that of the other partner (Hillis *et al.*, 1991). In plants, there are also examples of triploid hybrids showing deviation from locus number additivity (Lepen & Puizina, 2011). Thus, despite decreased fertility and irregular meiotic cycles, rDNA loci may not be genetically stable in triploid hybrids.

Evidence for intergenomic homogenization of rDNA in the *C.* × *schulzii* allohexaploid

The integrative approach of molecular and cytogenetic analyses indicates that intergenomic concerted evolution is ongoing in rDNA of the trigenomic allohexaploid, *C.* × *schulzii*. Several lines of evidence favor this scenario:

Changes in the A-genome One of the most prominent features was the increased number of A-genome NORs in some accessions of *C.* × *schulzii*. Because there was no such variation in the parental *C.* × *insueta,* the increase arose after the hybridization event and during diversification of *C.* × *schulzii*. Although novel loci in chromosomes that were originally free of rDNA have been reported in synthetic hybrids (Weiss & Maluszynska, 2000; Skalicka *et al.*, 2003; Pontes *et al.*, 2004), reductions of loci are more frequent in wild species (Kotseruba *et al.*, 2003; Lim *et al.*, 2007; Liu & Davis, 2011; Kolano *et al.*, 2012; Mlinarec *et al.*, 2012). One explanation is that large clusters are dispersed, initially to minor loci that may eventually transpose, expand or become lost in stabilized allopolyploids. Transposon activity is known to be influenced by polyploidy (Parisod *et al.*, 2010) and has been proposed to mobilize rDNA in plants (Raskina *et al.*, 2004).

Changes in the R/P-genome Cytogenetic and molecular markers indicate that *C. pratensis* has contributed more DNA to the *C.* × *schulzii* genome (PPPPRA) than has *C.* × *insueta*. If so, then the expected P/R gene ratio would be 2.8/1 in *C.* × *schulzii,* considering differences in rDNA copy number of the hybridizing species. However, the observed ratio was 0.68/1. Among the Sanger ITS clones, there were also substantially fewer *C.* × *schulzii* clones in the P-genome group than in the R-genome group (Fig. 2). These observations are consistent with the prevalent genes being inherited from the *C.* × *insueta* maternal donor. The ratios between individual ribotypes (Fig. 4b) were not exactly the same for *C.* × *schulzii* and *C.* × *insueta*, possibly reflecting a multiple origin of *C.* × *insueta* (Zozomová-Lihová *et al.*, 2014). Despite near homogenization towards maternal rDNA (*C. rivularis*), in both hybrids (as already observed in specimens from the 1970s), the P/R rDNA heterogeneity was substantial, reflecting high rDNA diversity inherited from these progenitors. Finally, many nonfunctional pseudogenes were transmitted from *C. pratensis* to *C.* × *schulzii,* and so the activity or inactivity of loci could potentially be involved in maintaining heterogeneity of arrays (Kovarik *et al.*, 2008). The *C.* × *schulzii* collected in 2012 has a higher content of the dominant R1 variant than a herbarium specimen of an individual collected almost 40 yr ago suggesting increased rDNA homogeneity in modern accessions. Interestingly, most populations of modern *Tragopogon* allotetraploids also have skewed rRNA gene ratios compared to herbarium specimens collected close to the species origin (80 yr) (Kovarik *et al.*, 2005). Thus, the process of rDNA homogenization in evolutionary young allopolyploid systems is still incomplete and ongoing.

Some disproportion in the abundance of *C. amara* ribotypes among the allopolyploid genomes that was evident in the NGS data (8–14%) and Sanger ITS clones (0–6%) (Table 2), points to the high NGS efficiency in detecting rare repeats, in contrast to conventional cloning. Diversity values calculated from NGS reads were comparable or even lower than those calculated from Sanger sequencing of clones (Table S7) indicating the accuracy of 454 sequencing. Therefore, even medium abundance clusters (2–9 reads/cluster) that were not considered in this study could reflect true genomic polymorphisms. PCR-mediated recombination is not a rare event, and is usually difficult to discern from that occurring naturally (Cronn *et al.*, 2002). The presence of a recombinant sequence in the NGS data (A5 variant) indicates that intergenomic recombination may indeed take place in allopolyploid genomes, but certainly at a lower frequency than suggested by ITS cloning.

Maternal/paternal effects influencing the direction of rDNA homogenization in recently formed allopolyploids

Different directions, degrees and patterns of concerted evolution among multiple accessions of the same allopolyploid have been attributed to the stochasticity of this process (Wendel *et al.*, 1995), whereas unidirectional or skewed homogenization indicates that it may be driven by selection (Nieto Feliner & Rossello, 2012). In the literature there is evidence for both scenarios. For example, three *Melampodium* polyploids homogenized the same parental rDNA repeats, regardless of whether they were donated by the maternal or paternal parents (Weiss-Schneeweiss *et al.*, 2012). By contrast, bidirectional concerted evolution, in which arrays become homogenized to alternative progenitor diploids in different allopolyploid derivatives, is evident among cotton, rice and soybean polyploids (Wendel *et al.*, 1995; Rauscher *et al.*, 2004; Bao *et al.*, 2010). In *Cardamine*, homogenization patterns (towards maternal alleles) were uniform across multiple accessions, in spite of multiple polyploid formations inferred for *C.* × *insueta* (Zozomová-Lihová *et al.*, 2014). Intriguingly, most populations of the recently formed *Tragopogon* allotetraploids show a reduction in paternal genes of *T. dubius* origin (Kovarik *et al.*, 2005). Similarly, one population of allohexaploid *Senecio cambrensis* that originated in the British Isles retained only the maternal rDNA type donated by *S. vulgaris* (Lowe & Abbott, 1996). Thus, it seems that it is the maternal repeat that dominates rDNA in all three recently-documented allopolyploid systems. In tobacco, the paternal genome underwent extensive eliminations and molecular rearrangements (Skalicka *et al.*, 2005), whereas the paternal rDNA repeat, paradoxically, dominates its rDNA (Volkov *et al.*, 1999; Kovarik *et al.*, 2004), arguing against the influence of a maternal cytoplasm. However, tobacco is a much older (> 10 000 yr) allopolyploid than *Cardamine*, and hence there might have been additional changes not necessarily related to genomic shock accompanying early generations of allopolyploid species.

Conclusions

Stabilization of parental genomes in the common nucleus has been a topic of considerable debate (Leitch & Leitch, 2008; Jones & Hegarty, 2009), and our results show that in *Cardamine*, two rounds of interspecific hybridization that occurred within < 150 yr resulted in a reduction in many rDNA variants. Paternal genes (*C. amara*) from the first hybridization event are no longer detectable in *C.* × *schulzii* by conventional cloning and sequencing. However, next generation sequencing of ITS amplicons clearly detected the presence of low copy *C. amara*-derived genes in *C.* × *schulzii,* and therefore has helped to resolve the complex evolutionary history of this recently originated allopolyploid. Our results show that in contrast to paternal derived rRNA genes, maternal derived rRNA genes have increased in number in the allopolyploid and appear to have begun to colonize chromosomes derived from the paternal parent.

Acknowledgements

This work was by supported by the European Social Fund (project CZ.1.07/2.3.00/30.0037 and CZ.1.07/2.3.00/20.0189) and Czech Science Foundation (excellence cluster P501/12/G090), and by the Research and Development Support Agency (project no. APVV-0139-12 to J.Z-L.). The authors are grateful to Ulrike Coja for technical assistance with cloning analysis, the curator of the ZT (Zürich) herbarium for allowing us to study and analyze herbarium specimens from Urnerboden, Karol Marhold for assistance in the field sampling and sharing biological knowledge, and the editor and reviewers for their constructive comments. We thank Lucille Schmieding and John D. Brooker for text editing.

References

Abbott RJ, Lowe AJ. 2004. Origins, establishment and evolution of new polyploid species: *Senecio cambrensis* and *S. eboracensis* in the British Isles. *Biological Journal of the Linnean Society* **82**: 467–474.

Ainouche ML, Baumel A, Salmon A. 2004. *Spartina anglica* Schreb.: a natural model system for analysing early evolutionary changes that affect allopolyploid genomes. *Biological Journal of the Linnean Society* **82**: 475–484.

Alvarez I, Wendel JW. 2003. Ribosomal ITS sequences and plant phlyogenetic inference. *Molecular Phylogenetics and Evolution* **29**: 417–434.

Bao Y, Wendel JF, Ge S. 2010. Multiple patterns of rDNA evolution following polyploidy in *Oryza*. *Molecular Phylogenetics and Evolution* **55**: 136–142.

Birchler JA. 2012. Genetic consequences of polyploidy in plants. In: Soltis PE, Soltis DE, eds. *Polyploidy and genome evolution.* Berlin/Heidelberg, Germany: Springer, 21–32.

Brown DD, Wensink PC, Jordan E. 1972. Comparison of ribosomal DNAs of *Xenopus laevis* and *Xenopus mulleri* – evolution of tandem genes. *Journal of Molecular Biology* **63**: 57–73.

Carlsen T, Bleeker W, Hurka H, Elven R, Brochmann C. 2009. Biogeography and phylogeny of *Cardamine* (Brassicaceae). *Annals of the Missouri Botanical Garden* **96**: 215–236.

Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* **104**: 482–489.

Dover GA. 1982. Molecular drive: a cohesive mode of species evolution. *Nature* **299**: 111–117.

Doyle JJ, Flagel LE, Paterson AH, Rapp RA, Soltis DE, Soltis PS, Wendel JF. 2008. Evolutionary genetics of genome merger and doubling in plants. *Annual Review of Genetics* **42**: 443–461.

Eickbush TH, Eickbush DG. 2007. Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**: 477–485.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**: 47–50.

Franzke A, Hurka H. 2000. Molecular systematics and biogeography of the *Cardamine pratensis* complex (Brassicaceae). *Plant Systematics and Evolution* **224**: 213–234.

Franzke A, Mummenhoff K. 1999. Recent hybrid speciation in *Cardamine* (Brassicaceae) – conversion of nuclear ribosomal ITS sequences in statu nascendi. *Theoretical and Applied Genetics* **98**: 831–834.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**: 221–224.

Hemleben V, Ganal M, Gersnter J, Schiebel K, Torres RA. 1988. Organization and length heterogeneity of plant ribosomal RNA genes. In: Kahl G, ed. *The architecture of Eukaryotic Gene.* Weinheim, Germany: VHC, 371–384.

Hillis DM, Moritz C, Porter CA, Baker RJ. 1991. Evidence for biased gene conversion in concerted evolution of ribosomal DNA. *Science* **251**: 308–310.

Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**: 254–267.

Jones RN, Hegarty M. 2009. Order out of chaos in the hybrid plant nucleus. *Cytogenetic and Genome Research* **126**: 376–389.

Kircher M, Kelso J. 2010. High-throughput DNA sequencing – concepts and limitations. *BioEssays* **32**: 524–536.

Kolano B, Tomczak H, Molewska R, Jellen EN, Maluszynska J. 2012. Distribution of 5S and 35S rRNA gene sites in 34 *Chenopodium* species (Amaranthaceae). *Botanical Journal of the Linnean Society* **170**: 220–231.

Kotseruba V, Gernand D, Meister A, Houben A. 2003. Uniparental loss of ribosomal DNA in the allotetraploid grass *Zingeria trichopoda* (2n = 8). *Genome* **46**: 156–163.

Kovarik A, Lim KY, Dadejová M, Matyasek R, Chase M, Knapp S, Clarkson J, Leitch AR. 2008. Evolution of rDNA in *Nicotiana* allopolyploids: a potential link between rDNA homogenization and epigenetics. *Annals of Botany* **101**: 815–823.

Kovarik A, Matyasek R, Lim KY, Skalicka K, Koukalova B, Knapp S, Chase M, Leitch AR. 2004. Concerted evolution of 18-5.8-26S rDNA repeats in *Nicotiana* allotetraploids. *Biological Journal of the Linnean Society* **82**: 615–625.

Kovarik A, Pires JC, Leitch AR, Lim KY, Sherwood AM, Matyasek R, Rocca J, Soltis DE, Soltis PS. 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. *Genetics* **169**: 931–944.

Ksiazczyk T, Kovarik A, Eber F, Huteau V, Khaitova L, Tesarikova Z, Coriton O, Chevre AM. 2011. Immediate unidirectional epigenetic reprogramming of NORs occurs independently of rDNA rearrangements in synthetic and natural forms of a polyploid species *Brassica napus*. *Chromosoma* **120**: 557–571.

Leitch AR, Leitch IJ. 2008. Genomic plasticity and the diversity of polyploid plants. *Science* **320**: 481–483.

Lepen I, Puizina J. 2011. FISH mapping of 18S-5.8S-26S rRNA genes and fluorochrome banding in the triploid viviparous onion *Allium* x *Cornutum clementi* ex Visiani, 1842. *Acta Biologica Cracoviensia. Series: Botanica* **53**: 111–116.

Lim KY, Matyasek R, Kovarik A, Leitch A. 2007. Parental origin and genome evolution in the allopolyploid *Iris versicolor*. *Annals of Botany* **100**: 219–224.

Liu B, Davis TM. 2011. Conservation and loss of ribosomal RNA gene sites in diploid and polyploid *Fragaria* (Rosaceae). *BMC Plant Biology* **11**: 157.

Lowe AJ, Abbott RJ. 1996. Origins of the new allopolyploid species *Senecio cambrensis* (asteraceae) and its relationship to the Canary Islands endemic *Senecio teneriffae*. *American Journal of Botany* **83**: 1365–1372.

Lysak MA, Mandakova T. 2013. Analysis of plant meiotic chromosomes by chromosome painting. *Methods in Molecular Biology* **990**: 13–24.

Malinska H, Tate JA, Matyasek R, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 2010. Similar patterns of rDNA evolution in synthetic and recently formed natural populations of *Tragopogon* (Asteraceae) allotetraploids. *BMC Evolutionary Biology* **10**: 291.

Mandakova T, Kovarik A, Zozomova-Lihova J, Shimizu-Inatsugi R, Shimizu KK, Mummenhoff K, Marhold K, Lysak MA. 2013. The more the merrier: recent hybridization and polyploidy in *Cardamine*. *Plant Cell* **25**: 3280–3389.

Marhold K, Lihova J, Perny M, Bleeker W. 2004. Comparative ITS and AFLP analysis of diploid *Cardamine* (Brassicaceae) taxa from closely related polyploid complexes. *Annals of Botany* **93**: 507–520.

Matyasek R, Renny-Byfield S, Fulnecek J, Macas J, Grandbastien MA, Nichols R, Leitch A, Kovarik A. 2012. Next generation sequencing analysis reveals a relationship between rDNA unit diversity and locus number in *Nicotiana* diploids. *BMC Genomics* **13**: 722.

Matyasek R, Tate JA, Lim YK, Srubarova H, Koh J, Leitch AR, Soltis DE, Soltis PS, Kovarik A. 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.

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

Mlinarec J, Satovic Z, Malenica N, Ivancic-Bace I, Besendorfer V. 2012. Evolution of the tetraploid *Anemone multifida* (2n = 32) and hexaploid *A. baldensis* (2n = 48) (Ranunculaceae) was accompanied by rDNA loci loss and intergenomic translocation: evidence for their common genome origin. *Annals of Botany* **110**: 703–712.

Mummenhoff K, Franzke A, Koch M. 1997. Molecular phylogenetics of *Thlaspi* sl (Brassicaceae) based on chloroplast DNA restriction site variation and sequences of the internal transcribed spacers of nuclear ribosomal DNA. *Canadian Journal of Botany-Revue Canadienne De Botanique* **75**: 469–482. **12**

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. *Molecular Phylogenetics and Evolution* **44**: 911–919.

Nieto Feliner G, Rossello JA. 2012. Concerted evolution of multigene families and homeologous recombination. In: Wendel JF, ed. *Plant genome diversity*. Wien, Germany: Springer, 171–194.

Parisod C, Alix K, Just J, Petit M, Sarilar V, Mhiri C, Ainouche M, Chalhoub B, Grandbastien MA. 2010. Impact of transposable elements on the organization and function of allopolyploid genomes. *New Phytologist* **186**: 37–45.

Pontes O, Neves N, Silva M, Lewis MS, Madlung A, Comai L, Viegas W, Pikaard CS. 2004. Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proceedings of the National Academy of Sciences, USA* **101**: 18 240–18 245.

Raskina O, Belyayev A, Nevo E. 2004. Quantum speciation in *Aegilops*: molecular cytogenetic evidence from rDNA cluster variability in natural populations. *Proceedings of the National Academy of Sciences, USA* **101**: 14 818–14 823.

Rauscher JT, Doyle JJ, Brown AH. 2004. Multiple origins and nrDNA internal transcribed spacer homeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. *Genetics* **166**: 987–998.

Schönswetter P, Stehlik I, Holderegger R, Tribsch A. 2005. Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology* **14**: 3547–3555.

Skalicka K, Lim KY, Matyasek R, Koukalova B, Leitch AR, Kovarik A. 2003. Rapid evolution of parental rDNA in a synthetic tobacco allotetraploid line. *American Journal of Botany* **90**: 988–996.

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

Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate JA, Mavrodiev E. 2004. Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. *Biological Journal of the Linnean Society* **82**: 485–501.

Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* **60**: 561–588.

Straub SC, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. 2012. Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany* **99**: 349–364.

Urbanska KM, Hurka H, Landolt E, Neuffer B, Mummenhoff K. 1997. Hybridization and evolution in *Cardamine* (Brassicaceae) at Urnerboden, central Switzerland: Biosystematic and molecular evidence. *Plant Systematics and Evolution* **204**: 233–256.

Urbanska KM, Landolt E. 1999. Patterns and processes of man-influenced hybridisation in *Cardamine* L. In: van Raamsdonk LWD, den Nijs HCM, eds. *Plant evolution in man-made habitats*. Amsterdam, the Netherlands: Hugo de Vries Laboratory, 29–47.

Urbanska-Worytkiewicz KM. 1977. Reproduction in natural triploid hybrids (2n = 24) between *Cardamine rivularis* Schur and *C. amara* L. *Berichte des Geobotanischen Institutes der ETH, Stiftung Rübel* **44**: 42–85. **13**

Vallejo-Marin M, Lye GC. 2013. Hybridisation and genetic diversity in introduced *Mimulus* (Phrymaceae). *Heredity* **110**: 111–122.

Veitia RA, Bottani S, Birchler JA. 2013. Gene dosage effects: nonlinearities, genetic interactions, and dosage compensation. *Trends in Genetics* **29**: 385–393.

Volkov RA, Borisjuk NV, Panchuk II, Schweizer D, Hemleben V. 1999. Elimination and rearrangement of parental rDNA in the allotetraploid *Nicotiana tabacum*. *Molecular Biology and Evolution* **16**: 311–320.

Weiss H, Maluszynska J. 2000. Chromosomal rearrangement in autotetraploid plants of *Arabidopsis thaliana*. *Hereditas* **133**: 255–261.

Weiss-Schneeweiss H, Bloch C, Turner B, Villasenor JL, Stuessy TF, Schneeweiss GM. 2012. The promiscuous and the chaste: frequent allopolyploid speciation and its genomic consequences in american daisies (*Melampodium* sect. *Melampodium*; Asteraceae). *Evolution* **6**: 211–228.

Weiss-Schneeweiss H, Emadzade K, Jang T-S, Schneeweiss GM. 2013. Evolutionary consequences, constraints and potential of polyploidy in plants. *Cytogenetic and Genome Research* **140**: 137–150.

Wendel JF, Schnabel A, Seelanan T. 1995. Bidirectional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proceedings of the National Academy of Sciences, USA* **92**: 280–284.

Xiong ZY, Gaeta RT, Pires JC. 2011. Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proceedings of the National Academy of Sciences, USA* **108**: 7908–7913.

Zimmerli S. 1986. Einfluss der Bewirtschaftung auf die Entstehung und Struktur der *Cardamine* Populationen auf dem Urnerboden. *Veröffentlichungen des Geobotanischen Institutes der ETH, Stiftung Rübel* **87**: 141–154.

Zozomová-Lihová J, Krak K, Mandáková T, Shimizu KK, Španiel S, Vít P, Lysák MA. 2014. Multiple hybridization events in *Cardamine* (Brassicaceae) during the last 150 years: revisiting a textbook example of neoallopolyploidy. *Annals of Botany* **113**: 817–830.

Supporting Information

Additional supporting information may be found in the online version of this article.

Table S1 List of the studied *Cardamine* accessions

Table S2 Distribution of polymorphic sites along the 18S coding and ITS1 regions based on NGS reads

Table S3 Summary of *18S*-ITS1 amplicon sequencing analysis

Table S4 Mutation analysis of *18S*-ITS1 PCR amplicons

Table S5 Merged clusters of NGS reads (*18S*-ITS1 amplicons) from the R/P/A subgenomes obtained from *C.* × *insueta* and *C.* × *schulzii*

Table S6 Statistical evaluation of variant contents in present-day and herbarium samples of *C.* × *schulzii*

Table S7 Comparison of nucleotide diversity values obtained for the same accessions by different sequencing methods

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

Author Query Form

Journal: NPH

Article: 12873/2014-17172

Dear Author,
During the copy-editing of your paper, the following queries arose. Please respond to these by marking up your proofs with the necessary changes/additions. Please write your answers on the query sheet if there is insufficient space on the page proofs. Please write clearly and follow the conventions shown on the attached corrections sheet. If returning the proof by fax do not write too close to the paper's edge. Please remember that illegible mark-ups may delay publication.

Many thanks for your assistance.

Query reference	Query	Remarks
1	**AUTHOR: Please check that authors and their affiliations are correct.**	
2	**AUTHOR: Please check all the level headings.**	
3	**AUTHOR: Please provide city name for Biomatters Ltd.**	
4	**AUTHOR: Please provide city name for Vector Laboratories.**	
5	**AUTHOR: Please give address information for Jackson Immuno-Research: town, state (if applicable), and country.**	
6	**AUTHOR: Please provide city and state name for Molecular Probes.**	
7	**AUTHOR: Please give address information for MetaSystems and Adobe Systems: town, state (if applicable), and country.**	
8	**AUTHOR: Please provide city name for Eurofins MWG.**	
9	**AUTHOR: Please check inserted text to explain A, R and P 'A, *C. amara*; R, *C. rivularis*; P, *C. pratensis*' is correct in figure 5.**	
10	**AUTHOR: Please check inserted text to explain A, R and P 'A, *C. amara*; R, *C. rivularis*; P, *C. pratensis*' is correct in figure 6.**	
11	**AUTHOR: Please check the page range is correct for reference Ainouche *et al.* (2004).**	
12	**AUTHOR: Please check the journal title is correct for reference Mummenhoff *et al.* (1997).**	
13	**AUTHOR: Please check the journal title is correct for reference Urbanska-Worytkiewicz (1977).**	
14	**AUTHOR: Figure 1 is of insufficient resolution; please supply a new file for this figure. The specifications for artwork are available at http://authorservices.wiley.com/bauthor/illustration.asp**	
15	**AUTHOR: Figure 4 is of insufficient resolution; please supply a new file for this figure. The specifications for artwork are available at http://authorservices.wiley.com/bauthor/illustration.asp**	
16	**AUTHOR: Figure 5 is of poor quality. Please check required artwork specifications at http://authorservices.wiley.com/bauthor/illustration.asp**	