



OPEN ACCESS

Edited by:

Sonia Garcia,
Botanical Institute of Barcelona
(CSIC), Spain

Reviewed by:

Natalia Borowska-Zuchowska,
University of Silesia in Katowice,
Poland

Josep A. Rossello,

Jardí Botànic UV, Spain

Radka Vozárová,

Academy of Sciences of the Czech
Republic, Czechia

***Correspondence:**

Jelena Mlinarec
jmlinarec@oikon.hr

†ORCID:

Jelena Mlinarec
orcid.org/0000-0002-2627-5374

Ljudevit Luka Boštjančić
orcid.org/0000-0001-8941-9753

Nenad Malenica
orcid.org/0000-0001-8784-0456

Višnja Besendorfer
orcid.org/0000-0001-9706-4921

Specialty section:

This article was submitted to
Plant Systematics and Evolution,
a section of the journal
Frontiers in Plant Science

Received: 30 March 2022

Accepted: 07 June 2022

Published: 06 July 2022

Citation:

Mlinarec J, Boštjančić LL,
Malenica N, Jurković A, Boland T,
Yakovlev SS and Besendorfer V
(2022) Structure and Methylation
of 35S rDNA in Allopolyploids
Anemone multifida ($2n = 4x = 32$,
BBDD) and *Anemone baldensis*
($2n = 6x = 48$, AABDD) and Their
Parental Species Show Evidence
of Nucleolar Dominance.
Front. Plant Sci. 13:908218.
doi: 10.3389/fpls.2022.908218

Structure and Methylation of 35S rDNA in Allopolyploids *Anemone multifida* ($2n = 4x = 32$, BBDD) and *Anemone baldensis* ($2n = 6x = 48$, AABDD) and Their Parental Species Show Evidence of Nucleolar Dominance

Jelena Mlinarec^{1*}, Ljudevit Luka Boštjančić^{2,3†}, Nenad Malenica^{4†}, Adela Jurković⁴, Todd Boland⁵, Sonja Siljak Yakovlev⁶ and Višnja Besendorfer^{4†}

¹ Oikon Ltd.-Institute of Applied Ecology, Zagreb, Croatia, ² LOEWE Centre for Translational Biodiversity Genomics (LOEWE-TBG), Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Frankfurt, Germany, ³ Department of Computer Science, ICube, UMR 7357, CNRS, Centre de Recherche en Biomédecine de Strasbourg, University of Strasbourg, Strasbourg, France, ⁴ Division of Molecular Biology, Department of Biology, University of Zagreb, Horvatovac, Croatia, ⁵ Memorial University of Newfoundland's Botanical Gardens, St. John's, NL, Canada, ⁶ CNRS, AgroParisTech, Ecologie Systématique Evolution, Université Paris-Saclay, Orsay, France

Transcriptional silencing of 35S rDNA loci inherited from one parental species is occurring relatively frequently in allopolyploids. However, molecular mechanisms by which it is selected for transcriptional silencing remain unclear. We applied NGS, silver staining and bisulfite sequencing to study the structure, expression and methylation landscape of 35S rDNA in two allopolyploids of common origin, allotetraploid *Anemone multifida* ($2n = 4x = 32$, genome composition BBDD) and allohexaploid *A. baldensis* ($2n = 6x = 48$, AABDD), and their genome donors, *A. sylvestris* ($2n = 16$, AA), *A. cylindrica* ($2n = 16$, BB) and *A. parviflora* ($2n = 16$, DD). The size of the recovered 35S rDNA units varied from 10,489 bp in *A. cylindrica* to 12,084 bp in *A. sylvestris*. *Anemone* showed an organization typical of most ribosomal 35S rDNA composed of NTS, ETS, rRNA genes, TTS and TIS with structural features of plant IGS sequences and all functional elements needed for rRNA gene activity. The NTS was more variable than the ETS and consisted of SRs which are highly variable among *Anemone*. Five to six CpG-rich islands were found within the ETS. CpG island located adjacent to the transcription initiation site (TIS) was highly variable regarding the sequence size and methylation level and exhibited in most of the species lower levels of methylation than CpG islands located adjacent to the 18S rRNA gene. Our results uncover hypomethylation of *A. sylvestris*- and *A. parviflora*-derived 35S rDNA units in allopolyploids *A. multifida* and *A. baldensis*. Hypomethylation of *A. parviflora*-derived 35S rDNA was more prominent in *A. baldensis* than in *A. multifida*. We showed that *A. baldensis* underwent coupled *A. sylvestris*-derived 35S rDNA array expansion and *A. parviflora*-derived 35S rDNA copy number

decrease that was accompanied by lower methylation level of *A. sylvestris*-derived 35S rDNA units in comparison to *A. parviflora*-derived 35S rDNA units. These observations suggest that in *A. baldensis* nucleolar dominance is directed toward *A. sylvestris*-derived chromosomes. This work broadens our current knowledge of the 35S rDNA organization in *Anemone* and provides evidence of the progenitor-specific 35S rDNA methylation in nucleolar dominance.

Keywords: bisulfite sequencing, CpG islands, genome size, epigenetic modification, IGS, Ranunculaceae, silver staining, polyploidy

INTRODUCTION

In eukaryotes, the 35S rDNA is arranged as highly repeated tandems and located at one or more chromosomal loci (Garcia et al., 2017). The rDNA unit is composed of the coding region for the 18S, 5.8S and 26S rRNA genes, the internal transcribed spacers (ITS1 and ITS2) and the intergenic spacer (IGS). IGS is composed of the 3' external spacer (3'-ETS), the non-transcribed region (NTS), and the 5' external transcribed spacer (5'-ETS). The 35S rDNA in plants is highly variable in copy number and in the nucleotide composition and length of IGS, however, the functional role of this region is relatively conservative (Rogers and Bendich, 1987). Repetitive elements (subrepeats), transcription initiation (TIS) and termination sites (TTS) can be found within IGS, functioning as enhancers, sequences with self-complementarity that could generate a conserved secondary structure, which are involved in the regulation of transcription of the 18-5.8-26S rDNA cistrons. Only 35S rDNA loci with active rRNA transcription and processing can form a nucleolus during the interphase of each cell cycle, and therefore, can be named as nucleolar organizer regions (NORs) (Ritossa and Spiegelman, 1965).

DNA methylation is an epigenetic mark that plays an essential role in regulating gene expression. DNA methylation is known to occur in all cytosine contexts (CpG, CpNpG, CpNpN) (Lawrence et al., 2004). CpG islands are regions rich in CpG, CpNpG, CpNpN di- and tri-nucleotides, no smaller than 200 bp and mostly present in promoters. The DNA methylation of CpG islands regulate gene expression through transcriptional silencing of the corresponding gene (Takai and Jones, 2002; Grummt and Pikaard, 2003; Lim et al., 2019). In 35S rDNA, the IGS is a region involved in transcriptional regulation, therefore, contains many pyrimidine base-pairs as methylation sites (Inácio et al., 2014). The enzymatic addition of a methyl group to DNA is performed by DNA methyltransferase (DNMT) on the 5'-carbon of the pyrimidine ring in cytosine. It is well known that the methylation in the promoter region represses the rRNA gene expression in NORs (Grummt and Pikaard, 2003). Due to the presence of the ribosomal genes in many copies, only a small amount of these

genes is transcriptionally active at any given time. The remaining of rDNA units bears epigenetic modifications, characteristic of heterochromatin, keeping it transcriptionally silent (Vieira et al., 1990; Neves et al., 1995; Costa-Nunes et al., 2010).

Allopolyploidy, interspecific hybridization followed by chromosome doubling, is one of the major driving forces in plant evolution (Liu and Wendel, 2003). The merger of two or more divergent genomes within one nucleus is challenging regarding long-term viability, and the sub-genomes undergo structural changes such as deletions, translocations, transposon activation and meiotic irregularities (Mlinarec et al., 2012a). Besides, subgenomes may undergo the reversible silencing of one of the NOR-bearing homoeologues, phenomenon termed as Nucleolar Dominance (ND) (Leitch and Leitch, 2008; Weiss-Schneeweiss et al., 2013; Wendel et al., 2018). Although our understanding of the role of DNA methylation in establishing and maintaining the ND has increased significantly in the last decades (Chen and Pikaard, 1997; Lawrence et al., 2004; Earley et al., 2006; Costa-Nunes et al., 2010), the mechanisms by which one ancestral rDNA set is selected for silencing remains elusive (Symonova, 2019). It was shown that in *Arabidopsis thaliana* and *Hordeum vulgare*, the 35S rRNA genes were silenced based on the chromosomal position in which they resided (Nicoloff et al., 1979; Schubert and Kunzel, 1990; Mohannath et al., 2016). In the allotetraploid *Arabidopsis suecica*, the small interfering RNA (siRNA)-directed DNA methylation pathway is required to inactivate the 35S rDNA of an *A. thaliana* origin (Preuss et al., 2008).

Anemone sensu stricto (s.s.) (c. 150 species, basic chromosome number $x = 8$), consists of perennial, low-growing herbs of worldwide distribution and with considerable diversity in morphology (Tamura, 1995). Various studies in the genus *Anemone* used the variation in the internal transcribed spacers (ITS1 and ITS2) of the 35S rDNA and the 5S rDNA intergenic spacers (NTS) mainly for systematic and phylogenetic purposes (Meyer et al., 2010; Mlinarec et al., 2016). However, the sequence structure and the methylation landscape of the 35S rDNA has never been studied in *Anemone*.

The genus *Anemone* offers an interesting plant model in studies of 35S rDNA evolution and molecular mechanisms of silencing of ancestral rDNA set in allopolyploids. *Anemone multifida* ($2n = 4x = 32$, BBDD) and *A. baldensis* ($2n = 6x = 48$, AABDD) are allopolyploids of hybrid origin with *A. sylvestris* ($2n = 16$, AA), *A. cylindrica* ($2n = 16$, BB) and *A. parviflora* ($2n = 16$, DD) as their genome donors

Abbreviations: 5-mC, 5-Methylcytosine; bal, *Anemone baldensis*; cyl, *A. cylindrica*; DAPI, 4',6'-diamidino-2-phenylindole; ETS, external transcribed spacer; FISH, fluorescence *in situ* hybridization; IGS, intergenic spacer; ITS, internal transcribed spacer; mul, *A. multifida*; ND, Nucleolar dominance; NOR, Nucleolar organizer region; NTS, non-transcribed region; par, *A. parviflora*; rDNA, ribosomal DNA; SR, subrepeat; syl, *A. sylvestris*; TIS, transcription initiation site; TTS, transcription termination sites.

(Mlinarec et al., 2012a, 2016). Allotetraploid *A. multifida* originated from the cross of the species the most similar to *A. cylindrica* (North America, moderate climate) and *A. parviflora* (North America, arctic climate), while allohexaploid *A. baldensis* (Europe, mountainous climate) originated from the cross of *A. multifida* (North and South America, mountainous climate) and *A. sylvestris* (Northern Europe and Asia, moderate climate). In *A. multifida* and *A. baldensis*, a uniparental/biparental 35S rDNA inheritance from the D subgenome donor (*A. multifida*) and the A and D subgenome donor (*A. baldensis*) has occurred as a result of the complete elimination of the rDNA units from the B subgenome donor (Mlinarec et al., 2012a). *Anemone sylvestris* and *A. cylindrica* have two 35S rDNA loci positioned terminally on short arm of two acrocentric chromosome pairs, while allopolyploids *A. multifida* and *A. baldensis* possess four and eight 35S rDNA signals, respectively. In *A. multifida*, these four 35S rDNA sites are located terminally in the D subgenome, while in *A. baldensis*, four 35S rDNA sites are located both in the A subgenome and the D subgenome (Mlinarec et al., 2012a,b).

Bisulfite-sequencing is an efficient method to determine 5-methylcytosine (5-mC) content at single-base resolution. It is based on sodium-bisulfite modification reaction which deaminates cytosine (but not the 5-mC) to uracil (U). The result is the conversion of C to thymine (T) in DNA during the synthesis of the complementary strand allowing the discrimination of unmethylated cytosine (C) from 5-mC. Presence of C in the sequencing read indicates that the cytosine was methylated, as the 5-mC is protected during sodium bisulfite reaction.

In order to shed more light on the 35S rDNA evolution and molecular mechanisms that lie behind epigenomic regulation in allopolyploid, we applied next generation sequencing (NGS), silver staining and bisulfite sequencing to study the organization and DNA methylation landscape of 35S rDNA in *Anemone* allopolyploids and their progenitor species. The following questions were posed: (i) what is the molecular organization of 35S rDNA in allopolyploids *A. multifida* ($2n = 4x = 32$, BBDD) and *A. baldensis* ($2n = 6x = 48$, AABDD) and their parental species *A. sylvestris* ($2n = 16$, AA), *A. cylindrica* ($2n = 16$, BB) and *A. parviflora* ($2n = 16$, DD)?, (ii) is the methylation landscape uniform or is it changing depending on the region within the ETS analyzed? and (iii) is DNA methylation involved in the preferential suppression of the homoeologous rDNA units?

MATERIALS AND METHODS

Samples and DNA Extraction

Anemone parviflora was obtained from the wild from Canada, Flower's Cove, Newfoundland (voucher number: UNB60590), while *A. sylvestris* (voucher number: ZA40577), *A. cylindrica* (voucher number: ZA40583), *A. multifida* (voucher number: ZA40585) and *A. baldensis* (voucher number: ZA40571) were grown in pots in Zagreb, Botanical Garden of University of Zagreb (Table 1).

Fresh leaves of *Anemone sylvestris*, *A. cylindrica*, *A. multifida* and *A. baldensis* and silica dried *A. parviflora* leaves were

TABLE 1 | List of *Anemone* species studied, their voucher number, origin and collection data.

Species	Voucher number	Origin
<i>A. baldensis</i>	ZA40571	Garden material from Botanical Garden of University of Zagreb, Croatia
<i>A. cylindrica</i>	ZA40583	Seed from Botanical Garden Chemnitz, Germany, Botanical garden University of Zagreb, Croatia
<i>A. parviflora</i>	UNB60590	Flower's Cove, Newfoundland, Canada, collected by T. Boland
<i>A. multifida</i>	ZA40585	Wenatchee Mountains, Washington State, United States, purchased from Kevoek Garden, United Kingdom, Botanical garden University of zagreb
<i>A. sylvestris</i>	ZA40577	Čučerje, Medvednica, Croatia, Botanical Garden of the University of Zagreb

used for DNA extraction. Care was taken to collect young light green leaves from all plants. Plants were in the period of developing flower buds. Total genomic DNA (gDNA) was isolated from ~100 mg of pooled leaves using a commercial plant gDNA isolation kit (DNeasy® Plant Mini Kit - QIAGEN) according to the manufacturer's instructions. The amount and purity of isolated gDNA was determined by using the NanoVue spectrophotometer (GE, Healthcare, United States).

Silver Staining and Genomic *in situ* Hybridization

Nuclei and metaphase chromosomes from the root tip cells were used for determination of active Nucleolar Organizing Regions (NORs). Metaphase chromosome spreads were obtained according to Mlinarec et al. (2006). Silver staining was performed according to the method of Hizume et al. (1980). Three individuals were analyzed from each species. In total, from 448 to 536 cells were analyzed per each species (Supplementary Table 1). We could not analyze the nucleoli of *A. parviflora* due to lack of living plants.

Flow Cytometry Analysis

The 2C nuclear DNA content was assessed by flow cytometry according to methods described by Marie and Brown (1993) and Mlinarec et al. (2009). *Triticum aestivum* cv. Chinese spring (2C = 30.9 pg and 43.7% G) was used as an internal standard. A slightly modified Galbraith's nuclei isolation buffer (Galbraith et al., 1983): 45 mM MgCl₂, 30 mM sodium citrate, 60 mM MOPS (4-morpholine propane sulphonate, pH = 7), and 1% (w/v) polyvinylpyrrolidone 10,000, pH 7.2) containing 0.1% (w/v) Triton X-100, supplemented with fresh 5 mM sodium metabisulphite and RNase (2.5 U/ml), was used for nuclei isolation and nuclei were stained with 100 µg/mL propidium iodide (Sigma Chemical Co., St. Louis, United States), the DNA intercalating fluorochrome dye. The nuclei suspension was filtered through 48 µm nylon mesh. To obtain the mean DNA content, six *Anemone multifida* and *A. baldensis* individuals were measured separately and with repetition on a flow cytometer (CyFlow SL3, Partec, Munster, Germany) with a 532 nm laser.

At least 5,000 to 10,000 nuclei were analyzed for each of six samples. The 2C DNA value was calculated using the linear relationship between the fluorescent signals from stained nuclei of the *A. multifida* and *A. baldensis* specimen and the *Triticum aestivum* internal standard.

Illumina NextSeq Sequencing and Clustering of Genomic Reads

A total of 1.25-17 μ g of high molecular weight (> 20 kb) genomic DNA from *A. parviflora*, *A. cylindrica*, *A. sylvestris*, *A. multifida* and *A. baldensis* was all sequenced using the Illumina HiSeq X platform (375×10^6 reads/lane, library insert size 350 bp) with a total 50 Gbp output for all five species. Raw Illumina 150 bp pair-end reads from low coverage DNA-seq experiments were obtained from a commercial service (Macrogen, Netherlands). Raw Illumina reads from DNA-seq experiment are available from SRA-NCBI under the BioProject ID PRJNA830872. Sequencing coverage, genome size, read archive accessions and other parameters are shown in Table 2.

Reconstruction of the Ribosomal DNA Array

Three to five million pair end reads from each genome were subjected to clustering analysis using RepeatExplorer and TAREAN pipeline (Novák et al., 2013; Boštjančić et al., 2021). Clusters containing 35S rDNA sequences were manually identified and confirmed with the BLASTn search against sequences deposited in GeneBank. These partial ribosomal DNA arrays were used as baits for the reconstruction of the full-length ribosomal DNA array loci with GetOrganelle 1.7.3.4 (Jin et al., 2020) and recommended parameters for reconstruction of plant ribosomal DNA arrays: -R 10 -k 35,85,115 -F embplant-nr. To confirm the correct identification of ribosomal DNA array assembly graphs were visualized with Bandage 0.8.1 (Wick et al., 2015). Boundaries of each coding region were determined from the alignment of *Anemone* ribosomal DNA arrays with the 35S rDNA sequence of *Alnus jorullensis* subsp. *jorullensis* (NCBI acc.: MF136530). Putative transcription initiation site (TIS) 5'-TATATTAGGGGGG-3' and putative transcription termination site (TTS) 5'-CCCTCCCC-5' as boundaries of the non-transcribed spacer (NTS) and external transcribed spacer (ETS) were identified based on the similarity those identified based on the comparative analysis of sequences with literature data concerning different plant species (Inácio et al., 2014; Hu et al., 2019). Dot-plot analysis of NTS and ETS region was conducted using LAST with threshold score = 39 ($E = 8.4e-11$) at online Mafft web interface¹ (Kato et al., 2019). Pairwise sequence identity of the NTS and ETS regions of intragenic spacer (IGS) were determined from the pairwise alignment of IGS regions with BioEdit (Hall, 1999). To determine the copy number of ribosomal DNA array in individual *Anemone* genomes, reads that were unmapped to the chloroplast genomes were mapped to the conserved 26S rDNA region of each ribosomal DNA array with bowtie2 v2.2.5 (Langmead and Salzberg, 2012),

¹<https://mafft.cbrc.jp/alignment/server/>

TABLE 2 | Basic statistics of NGS carried out in this study and copy number of the 35S rDNA units in allopolyploids *Anemone multifida* and *Anemone baldensis* and their ancestral species.

Species	1C [pg]	1C [Mb]	Sequencing coverage ^a	GC [%]	Read archive accession	Number of mapped reads to 26S rDNA loci	Coverage 26S rDNA ^b	Total number of reads used for mapping ^c	Genome coverage of reads used for mapping ^d	Number of 35S rDNA copies per 1C ^e	Subgenome	Subgenomic copy number of 35S rDNA per 1C ^h
<i>Anemone parviflora</i>	5.9 ^f	5,770.20	9.87	39	SRR18890536	2,51,920	9,556.02	17,10,09,442	3.85	2,480		
<i>Anemone cylindrica</i>	9.3 ^f	9,095.40	17.37	38	SRR18890535	4,14,074	15,565.30	78,86,75,298	11.27	1,381		
<i>Anemone sylvestris</i>	8.06 ^g	7,882.68	15.95	38	SRR18890538	6,14,468	23,196.44	65,50,48,152	10.8	2,147		
<i>Anemone multifida</i>	15.25	14,914.50	3.56	39	SRR18890539	1,22,614	4,650.04	27,83,53,314	2.43	1,917	AA	2,934
<i>Anemone baldensis</i>	21.01	20,542.89	2.73	38	SRR18890537	2,47,230	9,375.60	29,21,94,250	1.85	5,070	DD	2,136

^aCalculated as follows: Number of Reads * Average Read length (bp) * Genome size (Mb) * 106, with an average read length of 150 bp.

^bThe length of 26S rDNA sequence that was used as a reference is 3402 bp.

^cReads that mapped to the chloroplast genome were excluded from analysis.

^dCalculated as follows: Number of Reads * Average Read length (bp) * Genome size (Mb) * 106.

^eCalculated as follows: Coverage of 26S rDNA * Genome coverage of reads used for mapping.

^fFrom Rochefels et al. (1966).

^gFrom Vessely et al. (2013).

^hCalculated from the coverage of IGS subgenomic variants which is 55.2x (Subgenome A): 40.2 x (Subgenome D) (see Figure 2).

using the flags: `-very-sensitive-local`, `-no-mixed`, `-no-unal`, `-no-discordant`. Coverage information for the 26S rDNA was extracted using samtools coverage 1.11-2 (Li et al., 2009).

Identification of Repetitive Sequences in the Intergenic Spacer

Forward, reverse, palindromic and complement repeats were identified using the online web interface for REPuter² (Kurtz et al., 2001) with the following settings: Hamming distance of 3 and minimum repeat size of 30 bp (Zhang et al., 2019). For the identification of tandem repeats, we utilized the online web interface for the program Tandem Repeats Finder³ (Benson, 1999), the minimal repeat size was 10 bp and reported similarity percentage groups were 90-99 and 100% (Li et al., 2020).

Phylogenetic Analysis

Phylogenetic relationships and inheritance patterns were reconstructed using the NTS and ETS regions of reconstructed ribosomal DNA arrays. Sequence alignment was conducted with MAFFT v7.481 with selected E-INS-i method (Katoh et al., 2005). Maximum likelihood (ML) analysis of each dataset was conducted in IQ-TREE (Nguyen et al., 2015) with 10,000 bootstrap replicates and automatic selection of gene model TN + F + G4 for ETS tree and TPM3u + F + I + G4 for NTS tree. Phylogenetic trees were visualized in FigTree 1.4.4⁴.

Bisulfite Conversion

Bisulfite conversion was done using the EZ DNA Methylation-Lightning Kit (Zymo Research, Irvine, California, United States) according to the manufacturer's instructions. Approximately 300 ng of genomic DNA was used as input. The converted DNA was used immediately for PCR amplification or stored at -20°C for later use.

CpG Islands Detection

To identify the presence of CpG islands within IGS of the 35S rDNA sequence, the NewCPGreport tool⁵ was employed. This program estimated putative "CpG islands" using the observed versus expected GC numbers as well as their grouping patterns in the sequence. As a default, the CpG island should contain $>50\%$ CpG and Obs/Exp ratio >0.6 .

Primer Design, PCR Amplification and Cloning of Bisulfite Converted DNA

Primers for gDNA of bisulfite converted DNA templates were designed with Primer3⁶ and MethPrimer 2.0⁷, respectively. Regions within the CpG1, CpG2, and CpG3 islands of *A. parviflora*, *A. sylvestris*, *A. cylindrica* (Supplementary Table 2 and Supplementary Figure 1) were selected. As a template for

primer design, we used 35S rDNA sequence from *A. parviflora*, *A. sylvestris* and *A. cylindrica* obtained from RepeatExplorer. Specific primer pairs have been used for amplification of genomic DNA and bisulfite converted DNA. All PCRs were performed using GoTaq[®] Green Master Mix (Promega, Madison, WI, United States): 1X GoTaq[®] Green Master Mix, 10 pmol of each primer (Macrogen, Amsterdam, Netherlands) and 1 μl of template DNA (16 ng), in a 25 μl final reaction volume. The PCR program consisted of 35 cycles, each with 1 min denaturation at 95°C , 10 sec annealing from 54 to 58°C , 1 min extension at 72°C , and a final extension of 20 min. Annealing temperatures specific for each primer pair are presented in Supplementary Table 2. Amplicons were extracted and purified using ReliaPrep[™] DNA Clean-Up and Concentration System and cloned into pGEM-T Easy vector according to the manufacturer's instruction (Promega, Madison, WI, United States). Positive clones were PCR-screened using vector SP6 and T7 primers. Bisulfite-converted DNA (13 to 45 clones per sample) and genomic DNA (1-3 clones per sample) were Sanger sequenced by Macrogen (Amsterdam, Netherlands).

DNA Methylation Analysis

The genomic and bisulfite converted sequences (clones) were overlapped by ClustalX sequence alignment tool 1.83 and aligned. To identify the methylated cytosines, the ClustalX alignment file was uploaded to CyMATE software platform at <http://www.cymate.org> (Hetzl et al., 2007; Müllner and Hetzl, 2008). The CG, CHG and CHH methylation contexts were analyzed. Statistical analyses were performed using Wilcoxon signed-rank test (function `wilcox.exact`, from `exactRankTests` (version 0.8-34) package) implemented in the R. *P*-values were corrected for multiple testing using the Benjamini-Hochberg adjustment, with $\alpha = 0.05$.

RESULTS

The Number of Active NORs Revealed by Silver Staining

To determine the number of transcriptionally active 35S rDNA sites, we applied silver staining. The number of nucleoli in the individual cell and their frequencies were presented in Supplementary Table 1. The number of observed nucleoli ranged from one to four in diploids *A. cylindrica* and *A. sylvestris*, as well in allotetraploid *A. multifida* and allohexaploid *A. baldensis*. The most frequent number of nucleoli was two in *A. sylvestris*, *A. cylindrica* and *A. multifida* and three in *A. baldensis*. The maximum number of four nucleoli present in *A. sylvestris*, *A. cylindrica* and *A. multifida* corresponded to the four 35S rDNA sites, i.e., NORs, thus confirming that all rDNA sites can be active within the single cell (Figures 1A-C). In *A. baldensis*, the maximum number of four nucleoli, as well as the existence of four silver stained signals (e.g., NORs) in metaphase chromosomes indicate that in *A. baldensis* four out of eight 35S rDNA sites are active (Figures 1D,E).

Molecular organization of the 35S rDNA of allopolyploids *A. multifida* and *A. baldensis* of common origin and their

²<https://bibiserv.cebitec.uni-bielefeld.de/reputer/>

³<https://tandem.bu.edu/trf/trf.html>

⁴<https://github.com/rambaut/figtree/>

⁵https://www.ebi.ac.uk/Tools/seqstats/emboss_newcpgreport/

⁶<https://bioinfo.ut.ee/primer3-0.4.0/>

⁷www.urogene.org/methprimer2/tester-invitation.html

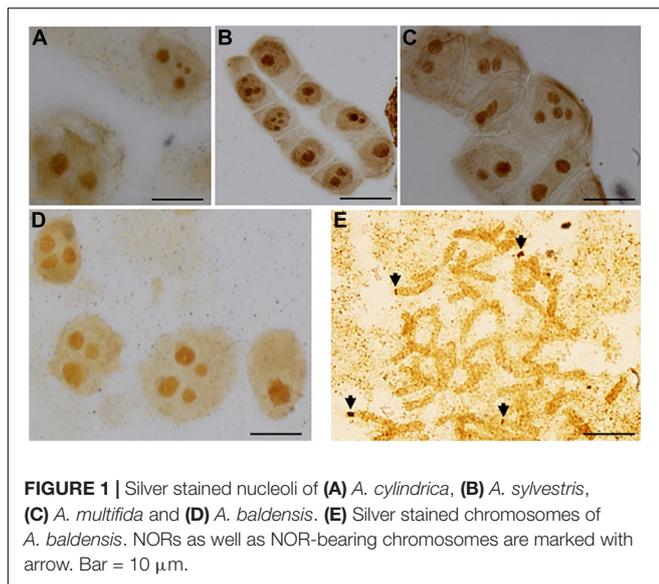


FIGURE 1 | Silver stained nucleoli of (A) *A. cylindrica*, (B) *A. sylvestris*, (C) *A. multifida* and (D) *A. baldensis*. (E) Silver stained chromosomes of *A. baldensis*. NORs as well as NOR-bearing chromosomes are marked with arrow. Bar = 10 μ m.

progenitor species *A. cylindrica*, *A. sylvestris* and *A. parviflora*: a comparative analysis.

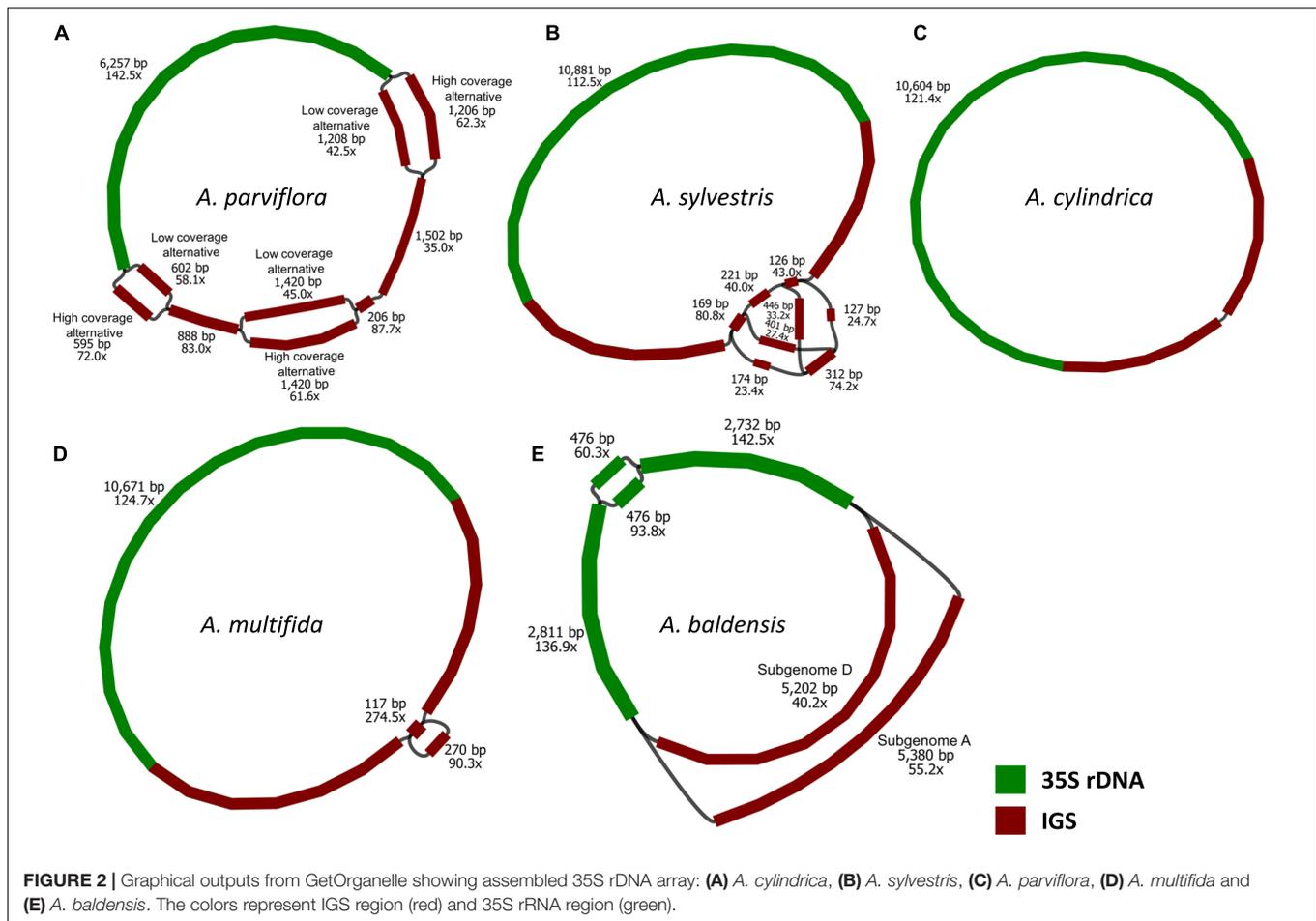
The genome size results showed that the average 2C DNA value in pg was 30.05 ± 0.46 (credibility of value (CV%) = 1.51) for *A. multifida* and 42.01 ± 1.25 (CV% = 2.98) for *A. baldensis*. Genome size values for parental species are taken from the literature: *A. cylindrica* (2C = 18.6 and 21.4 pg, Rothfels et al., 1966 and Bai et al., 2012, respectively), *A. parviflora* (2C = 11.8 pg, Rothfels et al., 1966) and *A. sylvestris* (2C = 17.02 and 16.12 pg, Barow and Meister, 2003; Veselý et al., 2013, respectively). Using high-throughput genomic data, we assembled 35S rDNA units in all five *Anemone* species. The 35S rDNA assembly graphs which contained the complete 35S rDNA unit were constructed using GetOrganelle 1.7.3.4 (Jin et al., 2020; Figure 2). The total length of the 35S rDNA unit varied from 10,489 bp in *A. cylindrica*, 10,715 bp in *A. multifida*, 10,761–10,939 bp in *A. baldensis*, 11,269–11,278 bp in *A. parviflora* and 12,084 bp in *A. sylvestris* (Supplementary Table 3). The total length of the recovered 18S–5.8S–26S regions (including ITS1 and ITS2 sequences) varied from 5763 bp in *A. multifida*, 5764 bp in *A. sylvestris*, *A. parviflora* and *A. baldensis* to 5770 bp in *A. cylindrica* (Supplementary Table 3). The homogeneity of the 18S, 5.8S and 26S rDNA rDNA units was high with little variation in genic sequences and both ITS regions. The length variation was accounted for by differences in length of ITS1 of *A. sylvestris* and *A. baldensis* (176 bp), vs. *A. parviflora*, *A. multifida* and *A. cylindrica* (175 bp) which contained a single nucleotide deletion after nucleotide position site 55. The 18S, 5.8S and 26S rDNA sequences of *A. sylvestris*, *A. parviflora*, *A. cylindrica*, *A. multifida* and *A. baldensis* showed from 93.7 to 97% nucleotide identity with their counterparts from *Alnus jorullensis* subsp. *jorullensis* (NCBI acc.: MF136530).

The genome proportion that each 35S rDNA cluster comprised was calculated from the number of mapped reads to 26S rDNA sequences compared with the total number of reads used for mapping. Reads that mapped to the chloroplast genome

were excluded from the analysis. From the genome proportion, the copy numbers of 35S rDNA sequences were estimated using the formula: coverage of 26S rDNA/genome coverage of reads used for mapping. Copy number per 1C ranged from 1381 in *A. cylindrica*, 1917 in *A. multifida*, 2147 in *A. sylvestris*, 2480 in *A. parviflora* to 5070 copies in *A. baldensis* distributed over two loci (Table 2).

The size of IGS sequences generated by GetOrganelle varied from 4702 bp in *A. cylindrica*, 4935 bp in *A. multifida*, 4980–5158 in *A. baldensis*, 5487–5492 bp in *A. parviflora* and to 6302 bp in *A. sylvestris*. Sequence similarity of IGS ranged from 46.5 to 78.8%, being the lowest between *A. sylvestris* and *A. parviflora* (46.5%) and the highest between *A. cylindrica* and *A. baldensis* (78.8%). The elementary structure of the *Anemone* IGS sequence generated by GetOrganelle consisted of six distinct regions: NTS, ETS, TTS, TIS, SR, and CpG islands (Figure 3). With the help of the alignment of our sequence with the transcription initiation regions (TIS) of other species, including Poaceae (Chang et al., 2010; Huang et al., 2017; Hu et al., 2019), Cucurbitaceae (Zentgraf et al., 1990), Fabaceae (Kato et al., 1990), Brassicaceae (Gruendler et al., 1991), Solanaceae (Borisjuk et al., 1997), and Punicaceae (Parvaresh and Talebi, 2014), we found the putative TIS of *Anemone* may be TATATTAGGGG (positions 3500–3512 in *A. sylvestris*, 2402–2414 in *A. parviflora*, 2163–2175 in *A. multifida*, 1901–1913 in *A. cylindrica*, 2208–2220 in *A. baldensis*) where A after TATATT represent start position (+ 1) (Supplementary Table 3 and Supplementary Figure 2). We also found that in *Anemone* sp. the putative TTS is CCCTCCCC (Yang et al., 2015; Hu et al., 2019). The IGS of *Anemone* contains an abundant and highly variable number of methylation sites that are irregularly distributed along the whole sequence. Interestingly, the majority of CCGG motifs are placed upstream 5' end of the NTS. In total, we counted 16 (*A. sylvestris*, *A. baldensis*-HC variant), 17 (*A. cylindrica*), 36 (*A. multifida*, *A. baldensis*-LC variant), and 44 (*A. parviflora*-HC variant, *A. parviflora*-LC variant) CCGG motifs. Furthermore, we counted from 169 (*A. cylindrica*) to 244 (*A. sylvestris*) CHG sites and from 632 (*A. cylindrica*) to 933 (*A. sylvestris*) CHH methylation sites in the IGS of five *Anemone* species investigated in this study. CHG and CHH sites are more frequent in the NTS than in the ETS region.

The region between the putative TTS and TIS represents the NTS (1899 bp in *A. cylindrica*, 2161 bp in *A. multifida*, 2206–2353 in *A. baldensis*, 2398–2400 bp in *A. parviflora*, 3498 bp in *A. sylvestris*), whereas the remaining IGS region is the ETS (2760 bp in *A. multifida*, 2760–2791 bp in *A. baldensis*, 2789 bp in *A. cylindrica*, 2790 bp in *A. sylvestris*, 3073–3080 bp in *A. parviflora*) (Supplementary Table 3). The NTS appeared as significantly more variable between investigated *Anemone* species in comparison to the ETS (Table 3). NTS showed 29.3–96.7% similarity among five *Anemone* species investigated in this study, being the lowest between *A. sylvestris* and *A. baldensis*-LC variant (29.3%) and the highest between *A. multifida* and *A. baldensis*-LC variant (96.7%), while the ETS showed 64–100% similarity, being the lowest between *A. cylindrica* and *A. baldensis*-LC variant (64%) and the highest between *A. sylvestris* 1, 2, and 3 variants (100%) (Table 3).



We identified two variants of IGS in *A. baldensis*: *A. baldensis*-HC (5159 bp) and *A. baldensis*-LC (4981 bp) (HC stands for High Copy and LC stands for Low Copy) (Figures 2, 3). The IGS variants differed in size, nucleotide sequence, abundance and origin. *Anemone baldensis*-HC and *A. baldensis*-LC variants exhibited 32% similarity in the NTS and 64.6% similarity in the ETS region. Furthermore, the *Anemone baldensis*-HC variant has higher coverage in comparison to the *A. baldensis*-LC variant along the whole IGS region (coverage proportion = 55.2x (*A. baldensis*-HC variant): 40.2x (*A. baldensis*-LC variant) (Figure 2E). The last, *Anemone baldensis*-HC variant is the most similar to the IGS of *A. sylvestris*, while the *A. baldensis*-LC variant is the most similar to the IGS of *A. multifida* (Figure 4 and Table 3). In *A. parviflora*, two IGS variants, *A. parviflora*-HC and *A. parviflora*-LC, have been observed, exhibiting high and low coverage, respectively. The variants have 94.9% of similarity in the region of NTS and 98.03% similarity in the region of ETS (Table 3). In *A. sylvestris*, three IGS variants, *A. sylvestris* 1, *A. sylvestris* 2 and *A. sylvestris* 3, have been observed. Interestingly, all three variants differ in the nucleotide sequence of the NTS, ranging from 90.8 to 93.5% of similarity, while the size is conserved (Table 3).

The analysis shows that among five *Anemone* sp. investigated in this study *A. sylvestris* possesses the highest number of

subrepeats (SRs), both tandem and forward. In *A. sylvestris*, the NTS is highly enriched in SRs along the whole length (Figure 3 and Supplementary Figure 3). Similarly, the NTS of *A. baldensis* HC variant is also enriched with the SRs along its length, however, slight reduction in the number of SRs has been observed in the allopolyploid (Figure 3 and Supplementary Figure 3). The NTS of *A. baldensis*-LC variant, *A. multifida*, *A. parviflora* and *A. cylindrica* show significantly lower number of SRs in comparison to the NTS of *A. sylvestris* and *A. baldensis*-HC variant. *A. baldensis*-LC variant, *A. multifida* and *A. parviflora* exhibit similar distribution of SRs, located in two regions of the NTS. *A. cylindrica* possesses one region in the central part of the NTS enriched with the SRs (Figure 3 and Supplementary Figure 3). Contrary to the NTS, the ETS showed a more conserved number and organization of SRs among five *Anemone* species investigated in this study. In all five *Anemone* species SRs are located in two regions in the central part of the ETS (Figure 3 and Supplementary Figure 4). All five *Anemone* species investigated in this study showed similar number of repeats in the ETS region (Supplementary Figure 5). On contrary, all three *A. sylvestris* variants (1, 2, and 3) showed significantly higher number of repeats in the NTS region in comparison with the NTS region of *A. baldensis*, *A. cylindrica*, *A. multifida* and *A. parviflora* (Supplementary Figure 5).

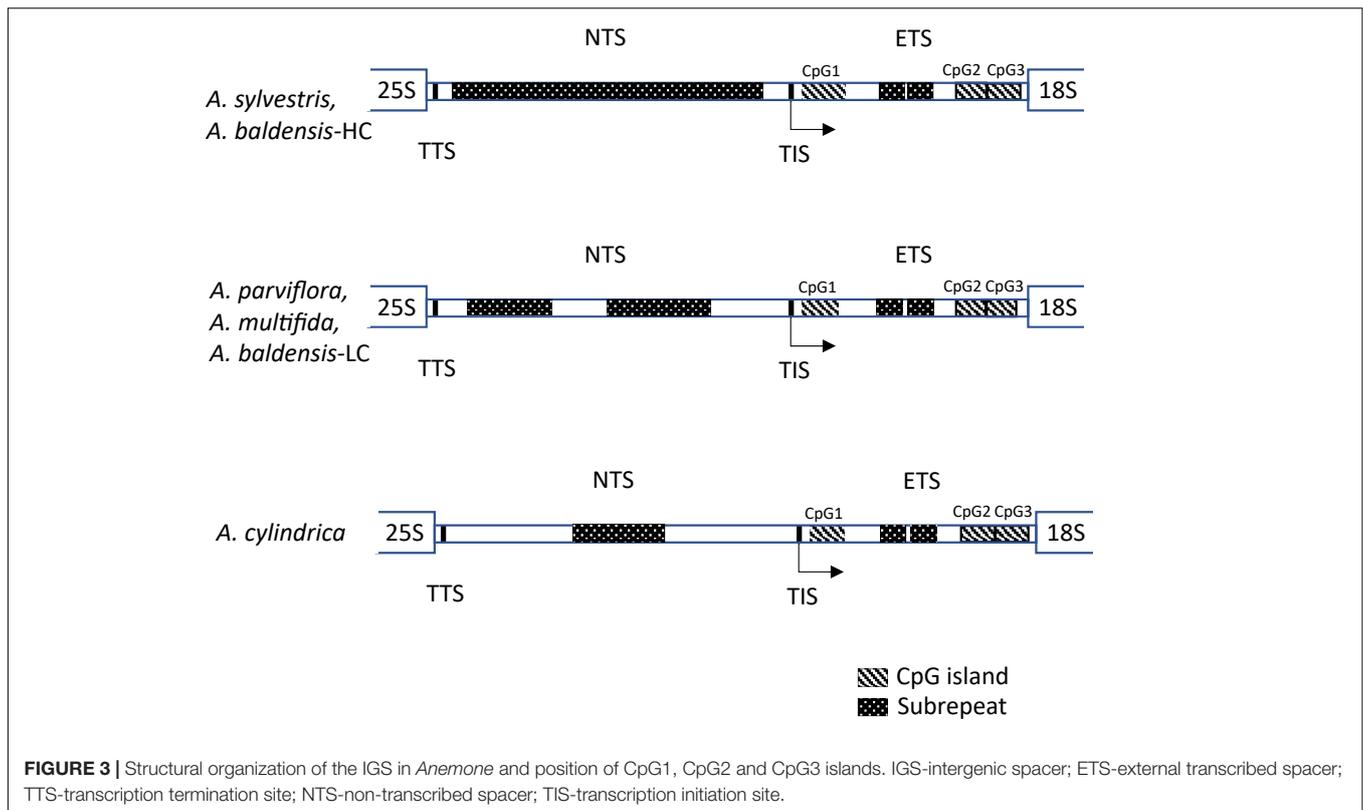


TABLE 3 | Pairwise identity matrix for two Intragenic spaces regions (IGS), Non-transcribed spacer (NTS) above the diagonal (gray background) and external transcribed spacer (ETS) below the diagonal (blue background) of five *Anemone* species and their IGS variants identified in this study.

Pairwise identity (%)	<i>A. baldensis-HC</i>	<i>A. baldensis-LC</i>	<i>A. cylindrica</i>	<i>A. multifida</i>	<i>A. parviflora-HC</i>	<i>A. parviflora-LC</i>	<i>A. sylvestris 1</i>	<i>A. sylvestris 2</i>	<i>A. sylvestris 3</i>
<i>A. baldensis-HC</i>	-	32	50.8	32.4	32.2	32.5	67	67	67
<i>A. baldensis-LC</i>	64.6	-	46.4	96.7	54.9	56.2	29.3	29.3	29.3
<i>A. cylindrica</i>	92.5	64	-	46.5	44.9	45.7	48.2	48.2	48.2
<i>A. multifida</i>	64.7	99.7	64.1	-	53.2	54.4	29.6	29.6	29.6
<i>A. parviflora-HC</i>	79.6	78.8	78.5	78.9	-	94.9	30.4	32.7	32.7
<i>A. parviflora-LC</i>	79.4	78.7	78.4	78.8	98.03	-	30.6	32.3	32.3
<i>A. sylvestris 1</i>	99.71	64.8	92.6	64.9	79.8	79.6	-	91.2	90.8
<i>A. sylvestris 2</i>	99.71	64.8	92.6	64.9	79.8	79.6	100	-	93.5
<i>A. sylvestris 3</i>	99.71	64.8	92.6	64.9	79.8	79.6	100	100	-

The application of the CpG Island Report program was conducted using the following parameters: Obs/Exp ratio >0.60, C% + %G >50.00 and minimum length >200. Thus, the CpG-rich regions in the ETS identified in this study satisfied almost all of the criteria proposed for vertebrate CpG islands (Gardiner-Garden and Frommer, 1987): they were larger than 200 bases, had higher G + C contents than the surrounding DNA regions and had O/E CpG values close to or slightly higher than those statistically expected. Five CpG islands were observed in the ETS of *A. cylindrica*, *A. baldensis-HC* variant and *A. sylvestris*, while six CpG islands were observed in the ETS of *A. multifida*, *A. parviflora-HC* variant and *A. baldensis-LC* variant (Supplementary Table 4).

DNA Methylation Analysis

To determine the level of cytosine methylation within the 35S rDNA, we used Sanger sequenced bisulfite-treated genomic DNA from three diploid and two polyploid *Anemone* taxa (Table 1). We were particularly focused on the CpG islands within the ETS as those regions appeared to be involved in transcription regulation (Deaton and Bird, 2011). In each species, we amplified three CpG islands, which corresponded to the CpG1', CpG4', and CpG5' islands in *A. sylvestris*, *A. baldensis-HC* and CpG1', CpG5', and CpG6' islands in *A. cylindrica*, *A. parviflora-HC*, *A. multifida* and *A. baldensis-LC*, determined by Emboss Newcpgreport tool (Supplementary Table 4). Analyzed regions within the ETS are named as CpG1, CpG2, and CpG3 islands and are shown in

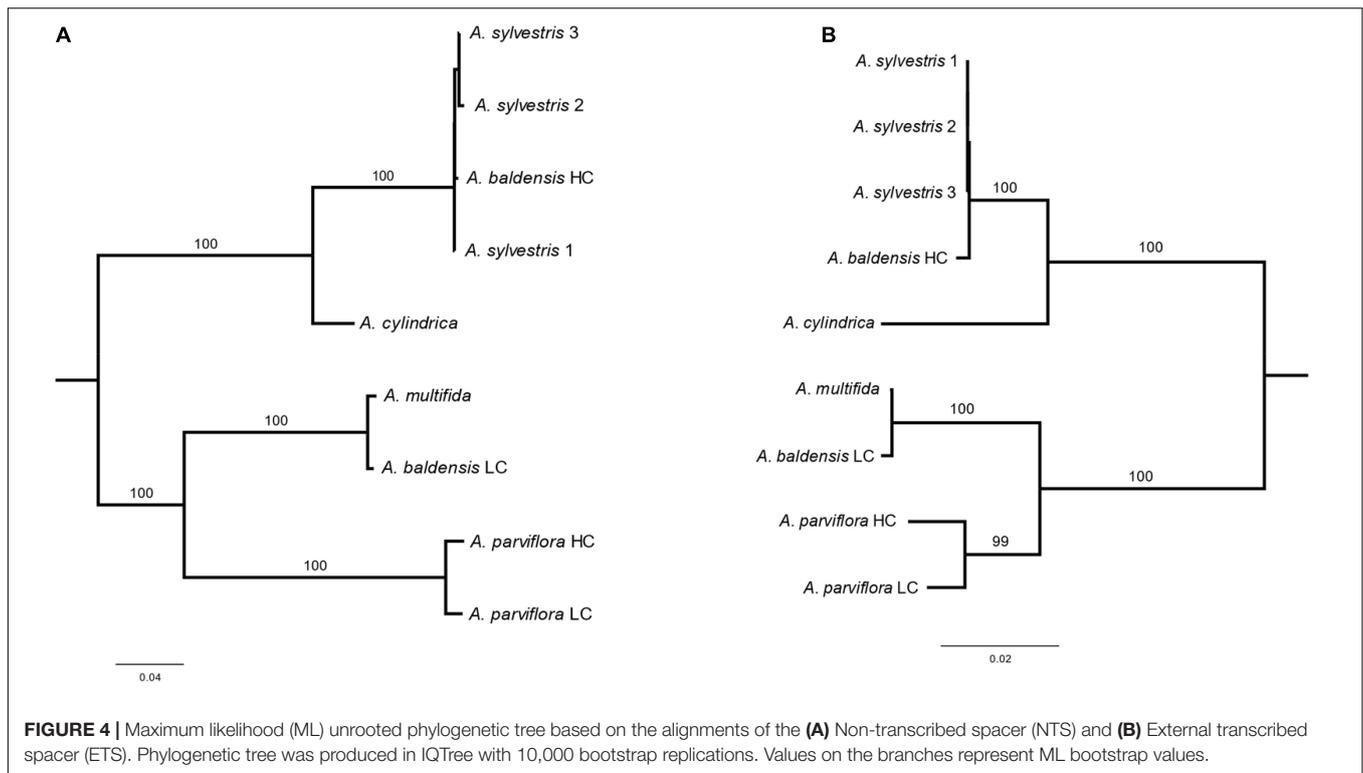


Figure 3 and **Supplementary Figure 1** (marked in bold). In *A. parviflora*, we sequenced and analyzed each of 14, 15, and 17 clones of CpG1, CpG2, and CpG3. In *A. cylindrica*, we sequenced and analyzed each of 13 clones of CpG1 and CpG3. In *A. sylvestris* we sequenced and analyzed 15 clones of CpG1, 15 clones of CpG2 and 16 clones of CpG3 (**Table 4**). In allotetraploid *A. multifida* we sequenced and analyzed 18 clones of CpG1, 15 clones of CpG2 and 37 clones of CpG3. All clones of *A. multifida* belonged to the D subgenome, since the 35S rDNA from the B subgenome is lost in the process of diploidization (Mlinarec et al., 2012a). In *A. baldensis*, we sequenced and analyzed 9 clones of CpG1, 11 clones of CpG2 and 34 clones of CpG3 of the A subgenome, and 22 clones of CpG1, 15 clones of CpG2 and 11 clones of CpG3 of the D subgenome (**Table 4**). Size of the analyzed sequences ranged from 211 to 637 bp (**Table 4**). The results of bisulfite sequencing are presented as diagrams at a single clone resolution (**Supplementary Figure 6**) and summarized in **Table 4**, **Supplementary Table 5**, and **Figure 5**. **Table 4** contains average methylation frequencies, while **Supplementary Table 5** contains methylation frequencies for each position within the CpG island.

Total methylation frequency of CpG1 island ranged from 19.19 to 95.43%, being the lowest in the A subgenome of *A. baldensis*, and the highest in *A. parviflora*. Total methylation frequency of CpG2 island ranged from 37.81 to 71.61%, being the lowest in the A subgenome of *A. baldensis*, and the highest in *A. parviflora*. Total methylation frequency of CpG3 island ranged from 42.69 to 87.40%, being the lowest in the *A. sylvestris*, and the highest in *A. parviflora*. In general, CpG1 island showed lower methylation frequencies (19.19–59.36%), followed by the CpG2

island (37.81–60.83%), while CpG3 island showed the highest methylation frequency (42.69–87.40%). The only exception was the CpG1 island of *A. parviflora* and *A. multifida* showing considerably high methylation frequency, 95.43 and 85.21%, respectively (**Table 4**).

In *Anemone*, we found considerable variation between clones originating from the same individual (**Table 4**). For example, in *A. multifida*, three clones contained 54.17% of total methylation, while 25 clones exhibited 84.38–100% of methylation regarding CpG1 island.

In *Anemone*, the highest methylation was observed for ETS cytosine at symmetrical CG motifs (average methylation = ~88%), followed by CHG motifs (average methylation ~55.88%), while considerably low (average methylation of 23.31%) methylation was found at non-symmetrical (CHH) sites (**Table 4**).

Discrimination Between the 35S rDNA Originating From the A and D Homoeologues in *Anemone baldensis*

In *Anemone*, two CpG1 size variants, CpG1-637 and CpG1-277, were revealed by cloning and sequencing of the CpG1 region, which were 637 bp and 277 bp long, respectively. Longer variant CpG1-637 dominated in *A. sylvestris* and its homologs in *A. baldensis* (the A subgenome), while shorter variant CpG1-277 dominated in *A. cylindrica*, *A. parviflora* and its homologs in *A. baldensis* (the D subgenome) and *A. multifida* (the D subgenome). Detailed analysis showed that variant CpG1-637 occurred from duplication of CpG1-277

TABLE 4 | Methylation levels in the ETS of the 35S rDNA in studied species determined by bisulfite sequencing.

Species	Subgenome	CpG region	CpG1 variant	Size (bp)	Number of clones	methyalted (%)	non-methyalted (%)	CGN	CHG	CHH
<i>A. parviflora</i>		CpG1		230	14	95.43	4.57	92.31	97.67	96.14
		CpG2		346	15	71.67	25.95	98.82	93.29	60.29
		CpG3		379	17	87.40	12.60	99.31	94.12	86.43
<i>A. cylindrica</i>		CpG1		211	13	59.36	40.64	94.23	77.88	25.54
		CpG3		225	13	47.69	52.31	94.44	51.75	5.49
<i>A. sylvestris</i>		CpG1	short	277	2	28.68	71.32	65.00	35.00	5.80
		CpG1	long	637	13	33.25	66.75	72.27	44.14	10.97
		CpG2		280	15	41.55	58.45	93.65	42.00	12.63
		CpG3		267	16	42.69	57.31	97.48	47.60	4.18
<i>A. multifida</i>	D	CpG1	short	222	38	91.09	8.91	98.94	97.26	86.68
	D	CpG2		346	15	60.83	39.17	97.1	82	41.82
	D	CpG3		379	37	87.08	12.92	99.55	97.63	77.16
<i>A. baldensis</i>	A	CpG1	long	637	9	19.19	80.81	50.00	15.08	4.33
	A	CpG2		278	11	37.81	62.19	94.09	29.29	11.59
	A	CpG3		267	34	43.09	56.91	96.91	47.51	5.29
	D	CpG1	short	323	22	29.68	70.32	63.69	20.13	6.06
	D	CpG2		346	15	51.29	48.71	98.26	80.67	26.30
	D	CpG3		266	11	52.38	47.62	98.35	59.74	11.11

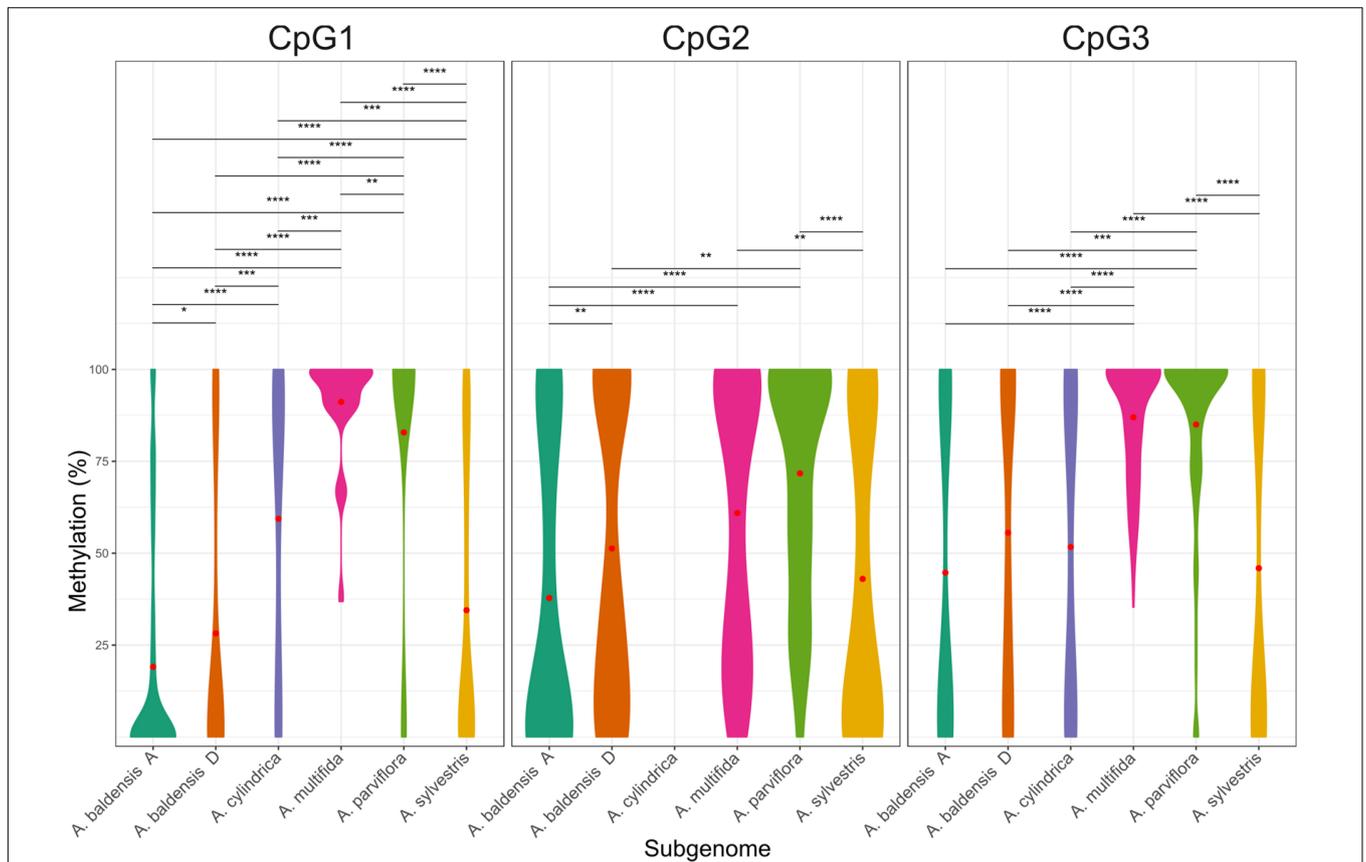


FIGURE 5 | Violin plots representing the overall cytosine methylation level in regions CpG1, CpG2, and CpG3 within the 5'ETS of 35S rDNA established by bisulfite sequencing of Sanger-sequenced clones. Wider regions of the violin plot represent methylation levels that occur more frequently. The values on the graph represent percentage of the methylated clones per each site for each type of modification. Different colors represent different species. Red dots represent mean methylation level. Statistical significance is represented as: * for $\alpha \leq 0.05$, ** for $\alpha \leq 0.01$, *** for $\alpha \leq 0.001$ and **** $\alpha \leq 0.0001$.

(Supplementary Figure 7). Thus, in *A. baldensis*, shorter CpG1 variant (CpG1-277) originated from the *A. parviflora*-like subgenome, while longer CpG1 variant (CpG1-637) originated from the *A. sylvestris*-like subgenome. A primer pair designed to amplify the CpG1 region in *A. sylvestris* enabled discrimination of CpG1 regions from the A and D subgenomes of *A. baldensis* according to the length of the CpG1.

To study the comparative methylation of the CpG2 region between the A and D subgenome of *A. baldensis* we took advantage of the CpG2 divergence. Specific primers were designed to amplify the CpG2 regions of A and D homoeologous in *A. baldensis*. Primers synthesized according to the *A. parviflora* 35S rDNA sequence were used to amplify CpG2 region from the D subgenome, while primers synthesized according to the 35S rDNA of *A. sylvestris* were used to amplify CpG2 region from the A subgenome of *A. baldensis* (Supplementary Table 2).

The primers that were constructed according to the 35S rDNA sequence from *A. sylvestris* amplified CpG3 regions from both A and D subgenomes of *A. baldensis*, however, those were still discriminated according to the ten SNPs within the amplified sequence (Supplementary Table 2). The CpG3 island that was similar to those of *A. parviflora* originated from the D subgenome, while the CpG3 island similar to *A. sylvestris* originated from the A subgenome.

Comparative 35S rDNA Methylation Analysis Between *Anemone parviflora* and *Anemone sylvestris* and Their Homologs in Allopolyploids *Anemone multifida* and *Anemone baldensis*

Results of this study clearly show a decrease in the methylation frequency (= hypomethylation) of the 35S rDNA between parental species and their allopolyploids, e.g., both allopolyploids showed a decrease in methylation level in the 35S rDNA loci. Hypomethylation was more prominent in allohexaploid *A. baldensis* than in allotetraploid *A. multifida*. In particular, average methylation frequency of the CpG1 and CpG2 regions in *A. parviflora* (95.43 and 71.67%, respectively) was higher in comparison to average methylation frequency in its homologs in *A. multifida* (91.09 and 60.83%, respectively) and *A. baldensis* (29.68 and 51.29%, respectively) (Table 4 and Figure 5). Average methylation frequency of CpG3 region was similar between *A. parviflora* (87.4%) and in its homologs in *A. multifida* (87.08%), and higher when compared with its homologs in *A. baldensis* (52.38%). Furthermore, average methylation frequency of CpG1 and CpG2 regions was higher in *A. sylvestris* (33.25 and 41.55%, respectively) in comparison to its homologs in *A. baldensis* (19.19% and 37.81%, respectively), while the average methylation frequency of CpG3 region of *A. sylvestris* (42.69%) was similar to its homologs in *A. baldensis* (43.09%) (Table 4 and Figure 5). Statistical analyses showed that the differences in methylation level of 35S rDNA between *A. parviflora* and its homologs in *A. baldensis* are statistically significant, while the differences between *A. parviflora* and its homologs in *A. multifida*, as well as differences between

A. sylvestris and its homologs in *A. baldensis* are not significant (Supplementary Table 6).

In allohexaploid *A. baldensis*, the D subgenome is more methylated than the A subgenome considering all three CpG islands, and also considering all three CGN, CHG and CHH classes (Table 4 and Figure 6). Statistical analyses showed that the differences in methylation level between the A and D subgenomes of *A. baldensis* are statistically significant considering CpG1 and CpG2 islands, while differences are not significant considering CpG3 island (Supplementary Table 7).

DISCUSSION

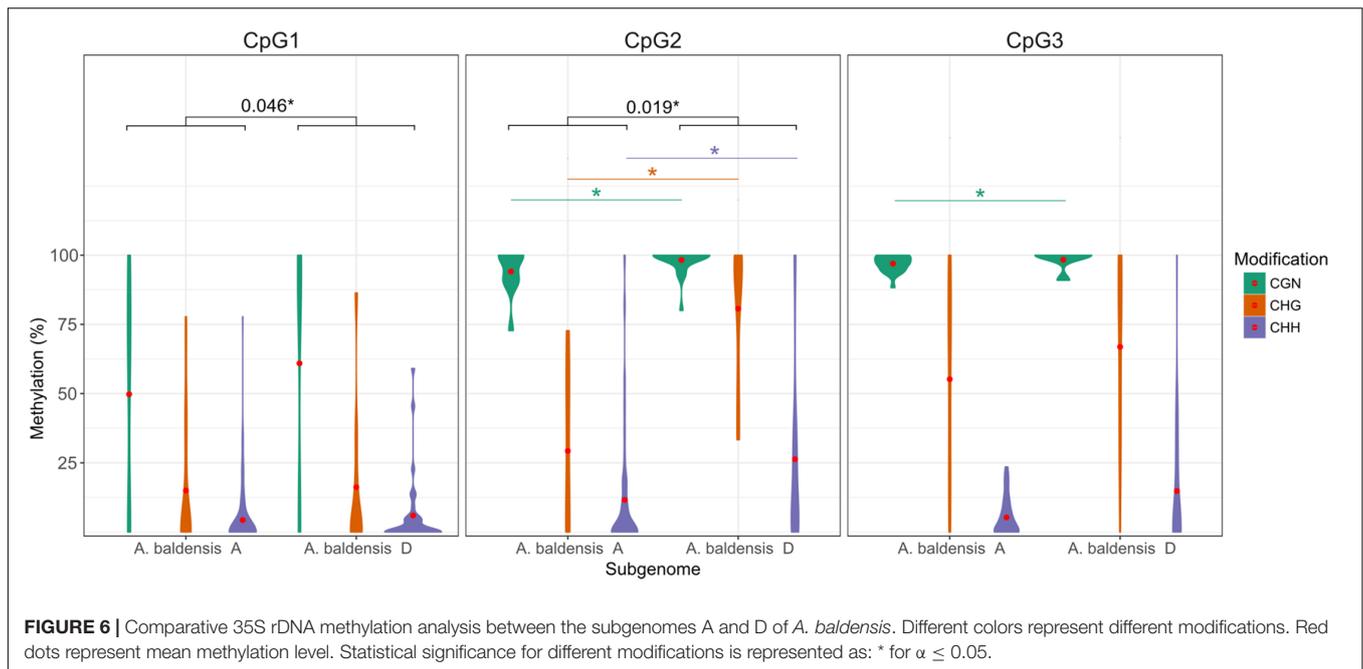
The sequence organization of the complete 35S rDNA array and the methylation status of three CpG-rich regions within the ETS was presented and compared for the first time in *A. multifida* ($2n = 4x = 32$, BBDD) and *A. baldensis* ($2n = 6x = 48$, AABDD) and their diploid parental species *A. sylvestris* ($2n = 2x = 16$, AA), *A. cylindrica* ($2n = 2x = 16$, BB) and *A. parviflora* ($2n = 2x = 16$, DD).

Polyploids *Anemone multifida* and *Anemone baldensis* and Their Parental Species *Anemone sylvestris*, *Anemone cylindrica* and *Anemone parviflora* Share the 35S rDNA Organization

The total length of the recovered 35S rDNA units varied from 10,489 bp in *A. cylindrica* to 12,084 bp in *A. sylvestris*. Comparative analysis of 35S rDNA between *Anemone* species showed that the size differences of the 35S rDNA was caused by the variability in sequence length of the IGS. In *Anemone*, the NTS (1899–3498 bp) showed high variability in sequence length, in comparison to significantly less variable size of ETS (2760–3080 bp). Similar was observed in other plant genera such as *Stipa* sp. (*Poaceae*) in which the NTS was much more variable in sequence length (1453–2347 bp), than the ETS (677–684 bp) (Krawczyk et al., 2017).

Results of this study reveal great differences in the amount and position of SRs within the NTS between *Anemone*, while the ETS is less variable in the amount and position of the SRs. In *A. sylvestris* almost the complete NTS consists of SRs, while in *A. cylindrica* only a small central region of the NTS is composed of the SRs. Furthermore, in allohexaploid *A. baldensis*, reduction of SRs in *A. sylvestris*-derived 35S rDNA units has been observed. Thus, results of this study clearly show the fast-evolving nature of the NTS within the IGS and that the SRs are very likely responsible for generating such great variability. It is known that the SRs serve as recombination hotspots generating the high multiplicity of ribosomal genes (Rogers and Bendich, 1987). In *Fagaceae* and *Asteraceae*, variability in the length of IGS sequences was mainly due to different copy numbers of SRs in the NTS and ETS region, respectively (Linder et al., 2000; Inácio et al., 2014).

The IGS in *Anemone* (4702–6302 bp) fit well within the range of IGS of other plant species such as *Fagus sylvatica*



(1715–1858 bp), *Quercus suber* (1980–2242 bp), *Stipa* sp. (2193–3098 bp), *Erianthus arundinaceus* (2955–3299 bp) and *Trillium* sp. (7400–12200 bp) (Yakura et al., 1983; Inácio et al., 2014; Krawczyk et al., 2017; Hu et al., 2019). Also, the ETS of *Anemone* (2760–3080 bp) fit well among known ETS lengths (680–3170 kb) (Linder et al., 2000; Vander Stappen et al., 2003; Inácio et al., 2014; Krawczyk et al., 2017; Hu et al., 2019). In *Helianthus*, the lengths of the ETS sequences varied from approximately 1600 bp (*H. annuus*, *H. argophyllus*, *H. praecox* ssp. *hirtus*) to 2100 bp (*H. atrorubens*) (Linder et al., 2000), while in *Stipa* sp. it ranged from 677 to 684 bp (Krawczyk et al., 2017).

Both regions, the NTS and the ETS, showed a strong phylogenetic signal (bootstrap values above 99 in both the NTS and the ETS trees) among five *Anemone* species investigated in this study (Figure 4). Furthermore, both the NTS and the ETS as molecular markers enabled to reveal the origin of allopolyploids *A. multifida* and *A. baldensis*. In agreement with the previous studies of Mlinarec et al. (2012a), *A. baldensis* high copy IGS variant (*A. baldensis*-HC variant) was the most similar to *A. sylvestris*, donor of the A subgenome, while *A. baldensis* low IGS copy variant (*A. baldensis*-LC variant) was the most similar to *A. multifida*, as well as to *A. parviflora*, donor of the D subgenome. Thus, the results of this study confirm that *A. sylvestris* (AA) and *A. multifida* (DD) are ancestral species of *A. baldensis* (AABBDD). Nonetheless, owing to its higher number of sub-repeats and due to the overall variability of the NTS region, the NTS region might not be a suitable marker for the phylogenetic inference in other distantly related species. Furthermore, in *A. baldensis*, *A. sylvestris*-derived 35S rDNA units are presented in considerably higher copy number than the *A. parviflora*-derived 35S rDNA units (Table 2). This result is in line with previous results of Southern blot and FISH experiments showing the presence of significantly weaker 35S

rDNA signals in the *A. parviflora*-like bands (Southern blot) and chromosomes (FISH) in comparison to the *A. sylvestris*-like ones (Mlinarec et al., 2012a). Baldwin and Markos (1998) calculated an approximately 1.3- to 2.4-fold higher rate of sequence evolution by nucleotide substitution in the ETS region studied than in ITS1 + ITS2 in *Calycadenia* (Compositae). Due to the extremely high rate of evolution within the IGS region, it managed to resolve the complex intraspecific relationships in Asteraceae, Poaceae and Compositae (Baldwin and Markos, 1998; Linder et al., 2000; Krawczyk et al., 2017).

In *Anemone*, the ETS is discretely GC-rich (52.38–53.64%). The CpG islands in the ETS account for the GC richness of this region. The placement of CpG islands within the ETS in *Anemone* is similar to cork oak and beech which exhibited 54 and 62% GC, respectively (Inácio et al., 2014). In *Arabidopsis thaliana* CpG islands were found in the SR region that is discretely rich in GC (53%; Gruendler et al., 1991).

While the length of the IGS varied in size between the *Anemone* species, its units present a specific order of conserved motifs in closely related species. *Anemone*, investigated in this study showed an organization typical of most ribosome IGS composed of NTS, ETS, TTS, and TIS, with structural features of plant IGS sequences and all functional elements needed for rRNA gene activity (Hu et al., 2019). Potential transcriptional enhancers and promoters for the RNA polymerase I machinery and one putative transcription initiation site were detected in all five species based on comparisons with the TIS motif of other species (Kato et al., 1990; Zentgraf et al., 1990; Gruendler et al., 1991; Borisjuk et al., 1997; Chang et al., 2010; Parvaresh and Talebi, 2014; Huang et al., 2017; Hu et al., 2019). TIS region detected in investigated *Anemone* has a TATATT sequence upstream the initiating A, like in the majority of the plants studied so far, including the Gymnosperms (Linder et al., 2000; Hu et al., 2019).

Moreover, in *Anemone* the TIS sequence, TATATTAGGGG, is the most similar to those of *Saccharum officinarum* and *Saccharum robustum* from the Poaceae family (Hu et al., 2019). The pyrimidine-rich motif CCCTCCCC, serving as proximal terminator at the 5' end of 26S rDNA, is also highly similar to the TTS at the 5' end of IGS in other plants (Yang et al., 2015; Hu et al., 2019).

Cloning and sequencing of the CpG1 region in *A. sylvestris* revealed two size variants, CpG1-637 and CpG1-277, downstream of the TIS region. The longer variant originated as duplication from shorter variant and it probably arose during divergence of *A. sylvestris*, as it is absent in *A. parviflora* and *A. cylindrica*. The presence of two size variants allowed discrimination of CpG1 islands according to the parental origin in *A. baldensis*. Such intraspecies length variability in the IGS of 35S rRNA genes is shown in other plant species, e.g., *Atropa*, *Aegilops*, *Nicotiana*, *Prunus* and *Brachypodium* (Volkov et al., 1993, 2017; Borisjuk et al., 1997; Shcherban et al., 2008; Borowska-Zuchowska et al., 2020).

The Methylation of the 35S rDNA in *Anemone*

In seed plants and most animals, the multicopy genes encoding ribosomal RNA (rDNA) typically exhibit heterochromatic features and high levels of DNA methylation (Matyášek et al., 2019). *Anemone* species analyzed in this study show wide range of methylation in 35S rDNA loci, determined by bisulfite sequencing, from relatively low levels (19.19%) to high levels (95.43%) in the CpG1 islands of *A. baldensis* and *A. parviflora*, respectively. Methylation frequencies observed in *Anemone* in this study fit well within the methylation frequencies observed in other seed plants. *Brassica rapa* and *B. oleracea* showed 30 and 17%, respectively, of the overall methylation level of ETS revealed by bisulfite sequencing (Shcherban et al., 2008). *Fagus sylvatica*, *Quercus faginea*, *Q. rubra* and *Castanea mollissima* showed high levels of methylation ($\geq 80\%$), while in *Q. suber*, *Q. pyrenaica* and *C. sativa* the levels of methylation were lower, ranging from 30 up to 60% regarding the *Bam*HI restriction site within 26S rDNA (Inácio et al., 2014). Concerning the 18S rRNA gene, only *Q. suber*, and *Q. faginea* showed moderate levels of methylation: 30 and 15%, respectively, while *F. sylvatica*, *Q. rubra*, *C. sativa* and *C. mollissima* showed low level ($< 3\%$) of methylation regarding *Bam*HI restriction site (Inácio et al., 2014). Furthermore, *Cycas revoluta* showed moderate levels of methylation (approximately 33% of total methylation) in 35S rDNA as revealed by bisulfite sequencing of 23 Sanger-sequenced clones (Wang et al., 2016). *Solanum lycopersicum* and *Theobroma cacao* showed moderate levels of 35S rDNA methylation, 33 and 24.2% of total methylation, respectively, determined from high-throughput bisulfite sequencing (Matyášek et al., 2019). Contrary to plants, bryophyte showed little or no methylation in both 35S and 5S loci, determined from high-throughput bisulfite sequencing (Matyášek et al., 2019). Total methylation levels of 35S rDNA, including both ITSs and non-repetitive parts of IGS subregions, ranged from negligible (1.7% in *Physcomitrella patens*, 2.4% in *Polytrichum formosum*, 6.9%

in *Marchantia polymorpha*) to moderate (16.6% in *Dicranum scoparium*) (Matyášek et al., 2019). In bryophytes, low levels of CG and CHG methylation in 35S rDNA have been confirmed by two assays, using methylation-sensitive restriction enzymes and by High-Resolution Cytosine Methylation Analysis of rDNA by Whole Genome Bisulfite Sequencing (Matyášek et al., 2019).

Considerably high variations in 35S rDNA methylation frequency between species studied so far could be an artifact caused by different methods used and also by the fact that in most studies only a relatively small number of copies within an array was analyzed. Comparing data from different studies that used different methods can lead to misinterpretation of the results. Therefore, in addition to focusing on epigenomics of single cell type and increasing data sets, methylation data should be confirmed by more than one assay, which would provide a more balanced, comprehensive and critical view of the research issue (Kumar et al., 2018).

Of special interest within the 35S rDNA array is the IGS of rRNA genes, where the promoter region (TIS) and other regulatory elements are located. CpG islands within IGS with higher average GC content are associated with regulation of rRNA gene expression regarding the cytosine methylation (Berger, 2007). Furthermore, the methylation property of the CpG islands within the ETS may have major implications on 35S rRNA gene regulation. In *Anemone*, segments of different methylation frequencies may have derived from arrays that have different impact on the transcription regulation of 35S rRNA genes. The proximity of CpG1 island to the TIS region and its lower level of methylation in comparison to distantly located CpG2 and CpG3 island, revealed in this study, suggests its putative role in regulation of gene transcription mediated by DNA methylation. It is estimated that about 60% of promoter sequences are located close to CpG islands in different organisms and, contrary to expectation, these are not highly methylated areas even though they contain a relatively high GC amount (Antequera, 2003; Deaton and Bird, 2011). In bryophytes *Polytrichum formosum*, *Physcomitrella patens*, *Marchantia polymorpha*, *Dicranum scoparium*, as well as angiosperms *Solanum lycopersicum* and *Theobroma cacao*, the methylcytosine density was relatively uniform across the rDNA units with no apparent differences between coding and non-coding regions, established by whole genome bisulfite sequencing (Matyášek et al., 2019).

Results of this study showed that *A. parviflora* and *A. multifida* exhibit significantly higher methylation levels of 35S rDNA (95.43 and 91.09%, respectively) in comparison with *A. cylindrica*, *A. sylvestris* and *A. baldensis* (19.19–56.36%) in the same region of ETS (CpG1 island). One possible explanation is that differences in methylation level could be related to the age of the leaves and/or individual. However, we consider this hypothesis less likely as gDNAs of all plants used in this study were isolated from the young leaves and from individuals of similar age. Recent investigations of Potabattula et al. (2020) consistently showed that methylation of the rDNA transcription unit including upstream control element (UCE), core promoter, 18S rDNA, and 28S rDNA in human sperm significantly increased with donor's age, revealed by using bisulfite pyrosequencing.

In *Anemone* all 35S rDNA loci are located distantly at the ends of chromosomes (Mlinarec et al., 2012a,b). Therefore, it is not likely that methylation in *Anemone* is related to the location on chromosomes. In *Arabidopsis*, 5S rDNA units located proximally to centromeres were more methylated than those located distantly (Vaillant et al., 2008).

In *Anemone* we observed considerable variations in methylation of CpG1 island between clones originating from the same individual. Astonishing epigenetic variation in methylation patterns was observed in other plants such as *Brassica napus* in which virtually none of the clones had identical methylation profile (Książczyk et al., 2011). In *B. oleracea* the number of methylated Cs ranged between 6–35%, while it ranged between 19–40% in *B. rapa* (Książczyk et al., 2011). Variations could be due to the differences in transcriptional status of the rRNA genes within the array. It is considered that within the tandemly arranged units, highly methylated genes are heterochromatic and inactive, while genes with low or no methylation levels are active (Suzuki and Bird, 2008). In *A. multifida*, moderately methylated CpG1 clones with 54.17% of methylation possibly originate from the active part of 35S rDNA arrays, while completely methylated CpG1 clones (100% of methylation) probably originate from inactive part of 35S rDNA arrays. In mammals, only about one hundred of 45S rRNA genes are estimated to be transcribed at any one time (Grummt and Pikaard, 2003). In *Danio rerio* (zebrafish) only twelve 5S rDNA copies appeared to be active in oocytes, while a large number of genes in another locus are silenced and activated only at later developmental stages (Locati et al., 2016). Alternatively, in *A. multifida*, completely methylated 35S rDNA fraction may derive from rDNA pseudogenes. So far, we have no evidence that 35S rDNA pseudogenes constitute a significant proportion of 35S rDNA arrays in *Anemone*. Previous work showed that *A. apennina* had significant deletions in both spacer and 5S rDNA coding region, however, pseudogenic character of 5S rDNA remained to be confirmed (Mlinarec et al., 2016). On the contrary, in *Cycas revoluta* pseudogenes showed less methylation (25%) compared with functional genes (53%), which is explained by a reduced number of symmetrical methylated sites in pseudogenized copies (Wang et al., 2016).

In all investigated *Anemone* species in this study, the CG and CHG sites appeared consistently to be more frequently methylated than the non-symmetrical CHH sites. Higher methylation frequency of the CG and CHG sites in comparison to the non-symmetrical CHH sites is typical for plant rDNA (Garcia et al., 2012). In *Cycas revoluta*, 35S rDNA cytosines at symmetrical CG and CHG motifs were shown to be highly methylated (>50% of cytosines), while low (c. 7%) methylation was found at non-symmetrical (CHH) sites (Wang et al., 2016). In *Artemisia*, *Tagetes*, *Helichrysum*, *Elachanthemum* and *Helianthus* the methylation density at different motifs had the identical tendency descending in order: CG > CHG > CHH (Garcia et al., 2012). In *Solanum lycopersicum* and *Theobroma cacao*, 35S rDNA cytosines at CG motifs exhibited 60.1 and 73.5% methylation, respectively, while the methylation at CHH sites was 15.6 and 4%, respectively (Matyášek et al., 2019).

Variation in rDNA Methylation Between Allopolyploids *Anemone multifida* and *Anemone baldensis* and Their Parental Species *Anemone sylvestris* and *Anemone parviflora*

Uniparental elimination of rDNA is a regular process of DNA deletion induced by allopolyploidy (Kovarik et al., 2008; Shcherban et al., 2008). In our previous work we showed that allopolyploids *A. multifida* and *A. baldensis* were subdued by the complete elimination of the *A. cylindrica*-like 35S rDNA units (Mlinarec et al., 2012a). There are many examples of other allopolyploid species undergoing uniparental inheritance such as *Aegilops*, *Brachypodium*, *Brassica* and *Nicotiana* (Kovarik et al., 2008; Shcherban et al., 2008; Książczyk et al., 2011; Borowska-Zuchowska and Hasterok, 2017). Results of this study showed a copy number reduction of the *A. parviflora*-derived 35S rDNA units in *A. multifida* and *A. baldensis*: from 2480 copies per haploid genome in *A. parviflora* to 1916 and 2136 copies in the *A. parviflora*-derived subgenome of *A. multifida* and *A. baldensis* (the D subgenome), respectively (Table 2). Contrary, a marked amplification of *A. sylvestris*-derived 35S rDNA units has been observed in *A. baldensis*: from 2147 copies per haploid genome in *A. sylvestris* to 2933 copies in the *A. sylvestris*-derived subgenome of *A. baldensis* (the A subgenome) (Table 2). Contraction of copy number of some rDNA families inherited from one parent was observed in *Aegilops sharonensis* \times *Ae. umbellulata* (Shcherban et al., 2008).

The expression of 35S rRNA genes in individual 35S rDNA loci can be assessed by silver staining. We demonstrated in this study that in *A. multifida* all two *A. parviflora*-derived 35S rDNA loci are active. On the contrary, presence of the highest number of four nucleoli in *A. baldensis* suggest that in this species two 35S rDNA loci may be suppressed. Since nucleolar fusion frequently occurs during the cell cycle (mononucleation), only the highest number of nucleoli estimated in a large sample size should be taken as evidence of the number of active rDNA loci (Ochatt and Seguí-Simarro, 2021, Rosselló et al., 2022). In this study the number of nucleoli was estimated in relatively large samples size of 536 cells in four individuals (Supplementary Table 1). This is further confirmed by performing silver staining on the metaphase chromosomes of *A. baldensis* revealing the presence of four silver stained signals, e.g., four NORs located on two pairs of acrocentric chromosomes (Figure 1E). Higher methylation frequency in the *A. parviflora*-derived 35S rDNA units in comparison to the *A. sylvestris*-derived 35S rDNA units might indicate suppression of *A. parviflora*-derived NORs in allohexaploid *A. baldensis*. In allotetraploid *Aegilops sharonensis* \times *Ae. umbellulata* differences in the expression level between individual NOR loci were not strictly correlated with gene dosage in a particular locus (Shcherban et al., 2008).

To test the 35S rDNA methylation changes in response to allopolyploidy, we performed bisulfite sequencing of CpG islands within the ETS. Considerably higher level of methylation in the 35S rDNA units of both *A. sylvestris* and *A. parviflora* in comparison to its homologs in *A. baldensis* (the A and D

subgenomes) suggests that the repressed 35S rDNA copies may be reactivated in a polyploid. Thus, it seems that the *A. parviflora*-derived 35S rDNA loss and *A. sylvestris*-derived 35S rDNA expansion were accompanied by hypomethylation of the ETS within both the *A. parviflora*-derived and *A. sylvestris*-derived 35S rDNA units. Hypomethylation of the promoter region within rDNA units from one subgenome was observed in synthetic allotetraploid *Aegilops sharonensis* \times *umbellulata*, revealed by digesting DNA with isoschizomers *Hpa*II and *Msp*I, which have different sensitivities to cytosine methylation (Shcherban et al., 2008). Contrary to *A. baldensis* and *Aegilops sharonensis* \times *umbellulata*, in allopolyploid *Brassica napus*, C-genome NORs became hypermethylated, while the methylation status of A-genome NORs remained unchanged (Książczyk et al., 2011). Further, in the hybrid grass *Brachypodium hybridum*, the methylation level of *B. distachyon* homologs remained unchanged, while those of *B. stacei* became hypermethylated (Borowska-Zuchowska et al., 2020). In bryophytes, the rDNAs were unmethylated both in gametophyte and sporophyte tissues suggesting that, within species, the ploidy level and developmental stage do not influence methylation (its absence) at this locus (Matyášek et al., 2019).

Interestingly, we showed that the hypomethylation of *A. parviflora*-derived 35S rDNA was more prominent in allohexaploid *A. baldensis* than in allotetraploid *A. multifida* which could be related with the hybrid vigor. Namely, *A. baldensis* shows greater hybrid vigor in comparison to *A. multifida*. It is shown than in *Arabidopsis*, decrease in DNA methylation 1 is required to produce a full level of hybrid vigor (Kawanabe et al., 2016).

Considerably higher levels of DNA methylation of the *A. parviflora*-like 35S rDNA units in comparison to the *A. sylvestris*-like 35S rDNA units considering CpG1 and CpG2 islands within ETS, indicate the involvement of DNA methylation in establishing and maintaining the Nucleolar Dominance (ND) in allohexaploid *A. baldensis*. The preferential methylation of the *A. parviflora*-like 35S rDNA loci accompanied by the copy number decrease indicates that in *A. baldensis* the ND is toward the *A. sylvestris*-like 35S rDNA loci. Considerable number of studies on plants that exhibit ND demonstrate significant differences in the DNA methylation levels between the transcriptionally active and repressed rDNA units (Vieira et al., 1990; Neves et al., 1995; Chen and Pikaard, 1997; Houchins et al., 1997; Komarova et al., 2004; Lawrence et al., 2004; Costa-Nunes et al., 2010; Dobešová et al., 2015; Borowska-Zuchowska et al., 2020). In hybrid grass *Brachypodium hybridum*, the actively transcribed D-genome 35S rDNA units exhibited a low DNA methylation level, while the *B. stacei*-like rRNA genes were characterized by a high level of DNA methylation (Borowska-Zuchowska and Hasterok, 2017).

In *Anemone* polyploids, the ND was observed in leaves, however, other tissues might show different pattern. A tissue-specific expression pattern of the rDNA homeologs has been observed in many plant hybrids and allopolyploids (Volkov et al., 2007; Borowska-Zuchowska et al., 2021). For instance, the ND was stable in leaves but not in roots of *B. hybridum* (Borowska-Zuchowska et al., 2021). Furthermore, although a fully established ND was observed in leaves, a trace expression of

the *A. thaliana*-derived rRNA genes was detected in the root-tip cells of the mature *Arabidopsis suecica* plants (Pontes et al., 2007).

Our preliminary plastid phylogenomics of *Anemone* reveal that in polyploid *A. baldensis*, *A. sylvestris*-derived subgenome is maternal indicating that in *A. baldensis* the ND is toward the maternal (sub)genome. Maternal control of 35S rDNA expression and epigenetic silencing of paternal 35S rRNA was observed in *Xenopus* hybrids (Michalak et al., 2015). However, there are also examples in the literature that the ND is independent of maternal or paternal effects (Sleutels et al., 2000); or that the ND in plant allopolyploids is not a maternal effect (Chen and Pikaard, 1997).

CONCLUSION

In this study, we provide in depth comparative analysis of the 35S rDNA methylation among five *Anemone* species. The organization of 35S rDNA is presented for the first time in any member of the Ranunculaceae family so far. A heterogenous methylation pattern across the ETS is shown. Our research uncovers contrasting changes in 35S rDNA copy number in allopolyploid *Anemone baldensis* accompanied with diverse methylation frequencies within 35S rDNA arrays between different subgenomes within allopolyploid. These observations suggest that in *A. baldensis* the nucleolar dominance (ND) is toward *A. sylvestris*-derived homologs resulting in considerable disruption at both genomic and epigenomic levels. Taken together, this work enhances our current knowledge of the 35S rDNA organization in *Anemone* and provides evidence for progenitor-specific 35S rDNA methylation in the ND.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the repository SRA-NCBI under the BioProject ID PRJNA830872 (<http://www.ncbi.nlm.nih.gov/bioproject/830872>), accession numbers SAMN27739250, SAMN27739251, SAMN27739252, SAMN27739253, SAMN27739254.

AUTHOR CONTRIBUTIONS

VB, JM, and NM conceived the study and designed the experimental part of the study. TB carried out by the field work. JM, NM, and AJ carried out the DNA extraction, bisulfite sequencing and cloning. VB carried out the silver staining. LB performed the bioinformatic and statistical analyses. JM and NM designed the primers. SY performed the flow cytometry. JM carried out the DNA methylation analysis, wrote the manuscript, and interpreted the results. All authors read, edited, enhanced the original version of the manuscript, and approved its final version.

FUNDING

This work was supported by grant IP-2016-06-8432 from the Croatian Science Foundation and the institutional project financed by the University of Zagreb.

ACKNOWLEDGMENTS

We would like to thank Oikon Ltd. for providing us with support in conducting this research.

REFERENCES

- Antequera, F. (2003). Structure, function and evolution of CpG island promoters. *Cell. Mol. Life Sci.* 60, 1647–1658. doi: 10.1007/s00018-003-3088-6
- Bai, C., Alverson, S. W., Follansbee, A., and Waller, M. D. (2012). New reports of nuclear DNA content for 407 vascular plant taxa from the United States. *Ann. Bot.* 110, 1623–1629. doi: 10.1093/aob/mcs222
- Baldwin, B. G., and Markos, S. (1998). Phylogenetic utility of the external transcribed spacer (ETS) of 18S-26S rDNA: congruence of ETS and ITS trees of *Calycadenia* (Compositae). *Mol. Phylogenet. Evol.* 10, 449–463. doi: 10.1006/mpev.1998.0545
- Barow, M., and Meister, A. (2003). Endopolyploidy in seed plants is differently correlated to systematics, organ, life strategy and genome size. *Plant Cell Environ.* 26, 571–584. doi: 10.1046/j.1365-3040.2003.00988.x
- Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acid Res.* 27, 573–580. doi: 10.1093/nar/27.2.573
- Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature* 447, 407–412. doi: 10.1038/nature05915
- Borisjuk, N. V., Davidjuk, Y. M., Kostishin, S. S., Miroshnichenko, G. P., Velasco, R., Hemleben, V., et al. (1997). Structural analysis of rDNA in the genus *Nicotiana*. *Plant Mol. Biol.* 35, 655–660.
- Borowska-Zuchowska, N., and Hasterok, R. (2017). Epigenetics of the preferential silencing of *Brachypodium stacei*-originated 35S rDNA loci in the allotetraploid grass *Brachypodium hybridum*. *Sci. Rep.* 7:5260. doi: 10.1038/s41598-017-05413-x
- Borowska-Zuchowska, N., Kovarik, A., Robaszekiewicz, E., Tuna, M., Savaş Tuna, G., Gordon, S., et al. (2020). The fate of 35S rRNA genes in the allotetraploid grass *Brachypodium hybridum*. *Plant J.* 103, 1810–1825. doi: 10.1111/tpj.14869
- Borowska-Zuchowska, N., Robaszekiewicz, E., Mykhailuk, S., Wartini, J., Pinski, A., Kovarik, A., et al. (2021). To be or not to be expressed: the first evidence of a nucleolar dominance tissue-specificity in *Brachypodium hybridum*. *Front. Plant Sci.* 12:768347. doi: 10.3389/fpls.2021.768347
- Boštančić, L. L., Bonassin, L., Anušić, L., Lovrenčić, L., Besendorfer, V., Maguire, I., et al. (2021). The *Pontastacus leptodactylus* (Astacidae) repeatome provides insight into genome evolution and reveals remarkable diversity of satellite DNA. *Front. Genet.* 11:611745. doi: 10.3389/fgene.2020.611745
- Chang, K. D., Fang, S. A., Chang, F. C., and Chung, M. C. (2010). Chromosomal conservation and sequence diversity of ribosomal RNA genes of two distant *Oryza* species. *Genomics* 96, 181–190. doi: 10.1016/j.ygeno.2010.05.005
- Chen, Z. J., and Pikaard, C. S. (1997). Transcriptional analysis of nucleolar dominance in polyploid plants: biased expression/silencing of progenitor rRNA genes is developmentally regulated in *Brassica*. *Proc. Natl. Acad. Sci. U.S.A.* 94, 3442–3447. doi: 10.1073/pnas.94.7.3442
- Costa-Nunes, P., Pontes, O., Preuss, S. B., and Pikaard, C. S. (2010). Extra views on RNA-dependent DNA methylation and MBD6-dependent heterochromatin formation in nucleolar dominance. *Nucleus* 1, 254–259. doi: 10.4161/nucl.1.3.11741
- Deaton, A. M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes* 25, 1010–1022. doi: 10.1101/gad.2037511
- Dobešová, E., Malinská, H., Matyášek, R., Leitch, A. R., Soltis, D. E., Soltis, P. S., et al. (2015). Silenced rRNA genes are activated and substitute for partially eliminated active homeologs in the recently formed allotetraploid, *Tragopogon mirus* (Asteraceae). *Heredity* 114, 356–365. doi: 10.1038/hdy.2014.111
- Earley, K., Lawrence, R. J., Pontes, O., Reuther, R., Enciso, A. J., Silva, M., et al. (2006). Erasure of histone acetylation by *Arabidopsis* HDA6 mediates large-scale gene silencing in nucleolar dominance. *Gene Dev.* 20, 1283–1293. doi: 10.1101/gad.1417706
- Galbraith, D. W., Harkins, K. R., Maddox, J. M., Ayres, N. M., Sharma, D. P., and Firoozabady, E. (1983). Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* 220, 1049–1051. doi: 10.1126/science.220.4601.1049
- García, S., Crhák Khaitová, L., and Kovařík, A. (2012). Expression of 5S rRNA genes linked to 35S rDNA in plants, their epigenetic modification and regulatory element divergence. *BMC Plant Biol.* 12:95. doi: 10.1186/1471-2229-12-95
- García, S., Kovařík, A., Leitch, A., and Garnatje, T. (2017). Cytogenetic features of rRNA genes across land plants: analysis of the plant rDNA database. *Plant J.* 89, 1020–1030. doi: 10.1111/tpj.13442
- Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. *J. Mol. Biol.* 196, 261–282.
- Gruendler, P., Unfried, I., Pascher, K., and Schweizer, D. (1991). rDNA intergenic region from *Arabidopsis thaliana*. Structural analysis, intraspecific variation and functional implications. *J. Mol. Biol.* 221, 1209–1222. doi: 10.1016/0022-2836(91)90929-z
- Grummt, I., and Pikaard, C. S. (2003). Epigenetic silencing of RNA polymerase I transcription. *Nat. Rev. Mol. Cell Biol.* 4, 641–649.
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98.
- Hetzl, J., Foerster, A. M., Raidl, G., and Mittelsten Scheid, O. (2007). CyMATE: a new tool for methylation analysis of plant genomic DNA after bisulfite sequencing. *Plant J.* 51, 526–536. doi: 10.1111/j.1365-313X.2007.03152.x
- Hizume, M., Sato, S., and Tanaka, A. (1980). A highly reproducible method of nucleolar organizing region staining in plants. *Stain Technol.* 55, 87–90. doi: 10.1007/BF00317037
- Houchins, K., O'Dell, M., Flavell, R. B., and Gustafson, J. P. (1997). Cytosine methylation and nucleolar dominance in cereal hybrids. *Mol. Gen. Genet.* 255, 294–301. doi: 10.1007/s004380050500
- Hu, X., Yu, F., Huang, Y., Sun, L., Li, X., Yang, S., et al. (2019). Characterization analysis of the 35S rDNA intergenic spacers in *Erianthus arundinaceus*. *Gene* 694, 63–70. doi: 10.1016/j.gene.2019.01.026
- Huang, Y., Yu, F., Li, X., Luo, L., Wu, J., Yang, Y., et al. (2017). Comparative genetic analysis of the 45S rDNA intergenic spacers from three *Saccharum* species. *PLoS One* 12:e0183447. doi: 10.1371/journal.pone.0183447
- Inácio, V., Rocheta, M., and Morais-Cecílio, L. (2014). Molecular organization of the 25S–18S rDNA IGS of *Fagus sylvatica* and *Quercus suber*: a comparative analysis. *PLoS One* 9:e98678. doi: 10.1371/journal.pone.0098678
- Jin, J. J., Yu, W. B., Yang, J. B., Song, Y., dePamphilis, C. W., Yi, T.-S., et al. (2020). GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* 21:241. doi: 10.1186/s13059-020-02154-5
- Kato, A., Nakajima, T., Yamashita, J., Yakura, K., and Tanifuji, S. (1990). The structure of the large spacer region of the rDNA in *Vicia faba* and *Pisum sativum*. *Plant Mol. Biol.* 14, 983–993. doi: 10.1007/BF00019395
- Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* 33, 511–518. doi: 10.1093/nar/gki198
- Katoh, K., Rozewicki, J., and Yamada, D. K. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Kawanabe, T., Ishikura, S., Miyaji, N., Sasaki, T., Wu, L. M., Itabashi, E., et al. (2016). Role of DNA methylation in hybrid vigor in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 113, E6704–E6711. doi: 10.1073/pnas.1613372113
- Komarova, N. Y., Grabe, T., Huigen, D. J., Hemleben, V., and Volkov, R. A. (2004). Organization, differential expression and methylation of rDNA in artificial *Solanum* allopolyploids. *Plant Mol. Biol.* 56, 439–463. doi: 10.1007/s11103-004-4678-x
- Kovarik, A., Dadejova, M., Lim, Y. K., Chase, M. W., Clarkson, J. J., Knapp, S., et al. (2008). Evolution of rDNA in *Nicotiana allopolyploids*: a potential link between rDNA homogenization and epigenetics. *Ann. Bot.* 101, 815–823. doi: 10.1093/aob/mcn019
- Krawczyk, K., Nobis, M., Nowak, A. A., Szczecińska, M., and Sawicki, J. (2017). Phylogenetic implications of nuclear rRNA IGS variation in *Stipa* L. (Poaceae). *Sci. Rep.* 7:11506. doi: 10.1038/s41598-017-11804-x

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.908218/full#supplementary-material>

- Książczyk, T., Kovarik, A., Eber, F., Huteau, V., Khaïtova, L., Tesarikova, Z., et al. (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. doi: 10.1007/s00412-011-0331-z
- Kumar, S., Chinnusamy, V., and Mohapatra, T. (2018). Epigenetics of modified DNA bases: 5-methylcytosine and beyond. *Front. Genet.* 9:640. doi: 10.3389/fgene.2018.00640
- Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., and Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res.* 29, 4633–4642. doi: 10.1093/nar/29.22.4633
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lawrence, R. J., Earley, K., Pontes, O., Silva, M., Chen, Z. J., Neves, N., et al. (2004). A concerted DNA methylation/histone methylation switch regulates rRNA gene dosage control and nucleolar dominance. *Mol. Cell* 13, 599–609. doi: 10.1016/s1097-2765(04)00064-4
- Leitch, A. R., and Leitch, I. J. (2008). Genomic plasticity and the diversity of polyploid plants. *Science* 320, 481–483. doi: 10.1126/science.1153585
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). 1000 genome project data processing subgroup, the sequence alignment/map (SAM) format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Li, Q. J., Su, N., Zhang, L., Tong, R. C., Zhang, X. H., Wang, J. R., et al. (2020). Chloroplast genomes elucidate diversity, phylogeny, and taxonomy of *Pulsatilla* (Ranunculaceae). *Sci. Rep.* 10:19781. doi: 10.1038/s41598-020-76699-7
- Lim, W. J., Kim, K. H., Kim, J. Y., Jeong, S., and Kim, N. (2019). Identification of DNA-methylated CpG islands associated with gene silencing in the adult body tissues of the Ogye chicken using RNA-Seq and reduced representation bisulfite sequencing. *Front. Genet.* 10:346. doi: 10.3389/fgene.2019.00346
- Linder, C. R., Goertzen, L. R., Heuvel, B. V., Francisco-Ortega, J., and Jansen, R. K. (2000). The complete external transcribed spacer of 18S-26S rDNA: amplification and phylogenetic utility at low taxonomic levels in Asteraceae and closely allied families. *Mol. Phylogenet. Evol.* 14, 285–303. doi: 10.1006/mpev.1999.0706
- Liu, B., and Wendel, J. F. (2003). Epigenetic phenomena and the evolution of plant allopolyploids. *Mol. Phylogenet. Evol.* 29, 365–379. doi: 10.1016/s1055-7903(03)00213-6
- Locati, M. D., Pagano, J. F., Ensink, W. A., van Olst, M., van Leeuwen, S., Nehrdich, U., et al. (2016). Linking maternal and somatic 5S rRNA types with different sequence specific non-LTR retrotransposons. *RNA* 4, 446–456. doi: 10.1261/rna.059642.116
- Marie, D., and Brown, S. C. (1993). A cytometric exercise in plant DNA histograms, with 2C values of 70 species. *Biol. Cell* 78, 41–51. doi: 10.1016/0248-4900(93)90113-s
- Matyášek, R., Krumpalová, A., Lunerová, J., Mikulášková, E., Rosselló, J. A., and Kovařík, A. (2019). Unique epigenetic features of ribosomal RNA genes (rDNA) in early diverging plants (Bryophytes). *Front. Plant Sci.* 10:1066. doi: 10.3389/fpls.2019.01066
- Meyer, K. M., Hoot, S. B., and Arroyo, M. T. K. (2010). Phylogenetic affinities of South American *Anemone* (Ranunculaceae) including the endemic segregate genera *Barnouidia* and *Oreithales*. *Int. J. Plant Sci.* 171, 323–331.
- Michalak, K., Maciak, S., Kim, Y. B., Santopietro, G., Oh, J. H., Kang, L., et al. (2015). Nucleolar dominance and maternal control of 45S rDNA expression. *Proc. Biol. Sci.* 282:20152201. doi: 10.1098/rspb.2015.2201
- Mlinarec, J., Chester, M., Siljak-Yakovlev, S., Papes, D., Leitch, A. R., and Besendorfer, V. (2009). Molecular structure and chromosome distribution of three repetitive DNA families in *Anemone hortensis* L. (Ranunculaceae). *Chromosome Res.* 17, 331–346. doi: 10.1007/s10577-009-9025-2
- Mlinarec, J., Franjević, D., Bočkor, L., and Besendorfer, V. (2016). Diverse evolutionary pathways shaped 5S rDNA of species of tribe Anemoneae (Ranunculaceae) and reveal phylogenetic signal. *Bot. J. Linn. Soc.* 182, 80–99. doi: 10.1111/boj.12452
- Mlinarec, J., Papeš, D., and Besendorfer, V. (2006). Ribosomal, telomeric and heterochromatin sequence localization in the karyotype of *Anemone hortensis* L. *Bot. J. Linn. Soc.* 150, 177–186. doi: 10.1111/j.1095-8339.2006.00467.x
- Mlinarec, J., Šatović, Z., Malenica, N., Ivančić-Baće, I., and Besendorfer, V. (2012a). The evolution of the tetraploid *Anemone multifida* (2n = 32) and hexaploid *A. baldensis* (2n = 48) (Ranunculaceae) was accompanied by similar genomic changes: evidence for their common origin. *Ann. Bot.* 110, 703–712. doi: 10.1093/aob/mcs128
- Mlinarec, J., Šatović, Z., Mihelj, D., Malenica, N., and Besendorfer, V. (2012b). Cytogenetic and phylogenetic studies of diploid and polyploid members of tribe Anemoneae (Ranunculaceae). *Plant Biol.* 14, 525–536. doi: 10.1111/j.1438-8677.2011.00519.x
- Mohannath, G., Pontvianne, F., and Pikaard, C. S. (2016). Selective nucleolus organizer inactivation in *Arabidopsis* is a chromosome position-effect phenomenon. *Proc. Natl. Acad. Sci. U.S.A.* 113, 13426–13431. doi: 10.1073/pnas.1608140113
- Müllner, C., and Hetzl, J. (2008). CyMATEads: reliable analysis of cytosine methylation in plant and animal DNA using bisulfite sequence data. *Schriftenr. Inf.* 26, 43–52.
- Neves, N., Heslop-Harrison, J. S., and Viegas, W. (1995). rRNA gene activity and control of expression mediated by methylation and imprinting during embryo development in wheat x rye hybrids. *Theor. Appl. Genet.* 91, 529–533. doi: 10.1007/BF00222984
- Nguyen, L.-T., Schmidt, A. H., von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300
- Nicoloff, H., Anastassovakristeva, M., Rieger, R., and Kunzel, G. (1979). Nucleolar dominance as observed in barley translocation lines with specifically reconstructed SAT chromosomes. *Theor. Appl. Genet.* 55, 247–251. doi: 10.1007/BF00265358
- Novák, P., Neumann, P., Pech, J., Steinhaisl, J., and Macas, J. (2013). RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* 29, 792–793. doi: 10.1093/bioinformatics/btt054
- Ochatt, S. J., and Seguí-Simarro, J. M. (2021). Analysis of ploidy in haploids and doubled haploids. *Methods Mol. Biol.* 2287, 105–125. doi: 10.1007/978-1-0716-1315-3_4
- Parvaresh, M., and Talebi, M. (2014). Molecular characterization of ribosomal DNA intergenic spacer (IGS) region in pomegranate (*Punica granatum* L.). *Plant Syst. Evol.* 300, 899–908. doi: 10.1007/s00606-013-0928-1
- Pontes, O., Lawrence, R. J., Silva, M., Preuss, S., Costa-Nunes, P., Earley, K., et al. (2007). Postembryonic establishment of megabase-scale gene silencing in nucleolar dominance. *PLoS One* 2:e1157. doi: 10.1371/journal.pone.0001157
- Potabattula, R., Zacchini, F., Ptak, G. E., Dittrich, M., Müller, T., El Hajj, N., et al. (2020). Increasing methylation of sperm rDNA and other repetitive elements in the aging male mammalian germline. *Aging Cell* 19:e13181. doi: 10.1111/accel.13181
- Preuss, S. B., Costa-Nunes, P., Tucker, S., Pontes, O., Lawrence, R. J., Mosher, R., et al. (2008). Multimegabase silencing in nucleolar dominance involves siRNA-directed DNA methylation and specific methylcytosine-binding proteins. *Mol. Cell* 32, 673–684. doi: 10.1016/j.molcel.2008.11.009
- Ritossa, F. M., and Spiegelman, S. (1965). Localization of DNA complementary to ribosomal RNA in the nucleolus organizer region of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 53, 737–745. doi: 10.1073/pnas.53.4.737
- Rogers, S. O., and Bendich, A. J. (1987). Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Mol. Biol.* 9, 509–520. doi: 10.1007/BF00015882
- Rosselló, J. A., Maravilla, A. J., and Rosato, M. (2022). The nuclear 35S rDNA world in plant systematics and evolution: a primer of cautions and common misconceptions in cytogenetic studies. *Front. Plant Sci.* 13:788911. doi: 10.3389/fpls.2022.788911
- Rothfels, K., Sexsmith, E., Heimburg, E., and Krause, M. O. (1966). Chromosome size and DNA content of species of *Anemone* L and related genera (Ranunculaceae). *Chromosoma* 20, 54–74.
- Schubert, I., and Kunzel, G. (1990). Position-dependent NOR activity in barley. *Chromosoma* 99, 352–359.
- Shcherban, A. B., Badaeva, E. D., Amosova, A. V., Adonina, I. G., and Salina, E. A. (2008). Genetic and epigenetic changes of rDNA in a synthetic allotetraploid, *Aegilops sharonensis* x *Ae. umbellulata*. *Genome* 51, 261–271. doi: 10.1139/G08-006
- Sleutels, F., Barlow, D. P., and Lyle, R. (2000). The uniqueness of the imprinting mechanism. *Curr. Opin. Genet. Dev.* 10, 229–233. doi: 10.1016/s0959-437x(00)00062-9

- Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9, 465–476. doi: 10.1038/nrg2341
- Symonova, R. (2019). Integrative genomics-Importance of the oldest repetitive fraction of the eukaryote genome. *Genes* 10:345. doi: 10.3390/genes10050345
- Takai, D., and Jones, P. A. (2002). Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. U.S.A.* 99, 3740–3745. doi: 10.1073/pnas.052410099
- Tamura, M. (1995). “Angiospermae: Ordnung Ranunculales, Fam. Ranunculaceae,” in *Die Natürlichen Pflanzenfamilien*, 2 Edn. Vol. 17, eds A. Engler, K. Prantl, and P. Hiepkö (Berlin: Duncker et Humboldt), 1–555.
- Vaillant, I., Tutois, S., Jasencakova, Z., Douet, J., Schubert, I., and Tourmente, S. (2008). Hypomethylation and hypermethylation of the tandem repetitive 5S rRNA genes in *Arabidopsis*. *Plant J.* 54, 299–309. doi: 10.1111/j.1365-3113.2008.03413.x
- Vander Stappen, J., Marant, S., and Volckaert, G. (2003). Molecular characterization and phylogenetic utility of the rDNA external transcribed spacer region in *Stylosanthes* (Fabaceae). *Theor. Appl. Genet.* 107, 291–298. doi: 10.1007/s00122-003-1245-9
- Vesely, P., Bures, P., and Šmarda, P. (2013). Nutrient reserves may allow for genome size increase: evidence from comparison of geophytes and their sister non-geophytic relatives. *Ann. Bot.* 112, 1193–1200. doi: 10.1093/aob/mct185
- Vieira, R., Mellosampayo, T., and Viegas, W. S. (1990). 1R chromosome nucleolus organizer region activation by 5-azacytidine in wheat x rye hybrids. *Genome* 33, 707–712.
- Volkov, R. A., Komarova, N. Y., and Hemleben, V. (2007). Ribosomal DNA in plant hybrids: inheritance, rearrangement, expression. *Syst. Biodivers.* 5, 261–276. doi: 10.1017/S1477200007002447
- Volkov, R. A., Kostishin, S. S., and Panchuk, I. I. (1993). rDNA organization in species from subfamily Prunoideae. *Mol. Biol.* 27, 1356–1367.
- Volkov, R. A., Panchuk, I. I., Borisjuk, N. V., Hosiawa-Baranska, M., Maluszynska, J., and Hemleben, V. (2017). Evolutionary dynamics of 45S and 5S ribosomal DNA in ancient allohexaploid *Atropa belladonna*. *BMC Plant Biol.* 17:21. doi: 10.1186/s12870-017-0978-6
- Wang, W., Ma, L., Becher, H., Garcia, S., Kovarikova, A., Leitch, I. J., et al. (2016). Astonishing 35S rDNA diversity in the gymnosperm species *Cycas revoluta* Thunb. *Chromosoma* 125, 683–699. doi: 10.1007/s00412-015-0556-3
- Weiss-Schneeweiss, H., Emadzade, K., Jang, T. S., and Schneeweiss, G. M. (2013). Evolutionary consequences, constraints and potential of polyploidy in plants. *Cytogenet. Genome Res.* 140, 137–150. doi: 10.1159/000351727
- Wendel, J. F., Lisch, D., Hu, G., and Mason, A. S. (2018). The long and short of doubling down: polyploidy, epigenetics, and the temporal dynamics of genome fractionation. *Curr. Opin. Genet. Dev.* 49, 1–7. doi: 10.1016/j.gde.2018.01.004
- Wick, R. R., Schultz, M. B., Zobel, J., and Holt, K. E. (2015). Bandage - Interactive visualization of de novo genome assemblies. *Bioinformatics* 31, 3350–3352. doi: 10.1093/bioinformatics/btv383
- Yakura, K., Kato, A., and Tanifuji, S. (1983). Structural organization of ribosomal DNA in four *Trillium* species and *Paris verticillata*. *Plant Cell Physiol.* 24, 1231–1240.
- Yang, K., Robin, A. H. K., Yi, G. E., Lee, J., Chung, M. Y., Yang, T. J., et al. (2015). Diversity and inheritance of intergenic spacer sequences of 45S ribosomal DNA among accessions of *Brassica oleracea* L. var. *capitata*. *Int. J. Mol. Sci.* 16, 28783–28799. doi: 10.3390/ijms161226125
- Zentgraf, U., Ganai, M., and Hemleben, V. (1990). Length heterogeneity of the rRNA precursor in cucumber (*Cucumis sativus*). *Plant Mol. Biol.* 15, 465–474. doi: 10.1007/BF00019163
- Zhang, T., Xing, Y., Xu, L., Bao, G., Zhan, Z., Yang, Y., et al. (2019). Comparative analysis of the complete chloroplast genome sequences of six species of *Pulsatilla* Miller, Ranunculaceae. *Chin. Med.* 14:53. doi: 10.1186/s13020-019-0274-5

Conflict of Interest: JM was employed by Oikon Ltd.-Institute of Applied Ecology.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Mlinarec, Boštjančić, Malenica, Jurković, Boland, Yakovlev and Besendorfer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.