Molecular Phylogenetics and Evolution 100 (2016) 109–123

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular and cytogenetic evidence for an allotetraploid origin of *Chenopodium quinoa* and *C. berlandieri* (Amaranthaceae)

Bozena Kolano [a,*], Jamie McCann [b], Maja Orzechowska [a], Dorota Siwinska [a], Eva Temsch [b], Hanna Weiss-Schneeweiss [b]

[a] Department of Plant Anatomy and Cytology, University of Silesia, Jagiellonska 28, 40-032 Katowice, Poland
[b] Department of Botany and Biodiversity Research, University of Vienna, Rennweg 14, Vienna, Austria

ARTICLE INFO

Article history:
Received 4 January 2016
Revised 1 April 2016
Accepted 6 April 2016
Available online 7 April 2016

Keywords:
Allopolyploidy
Genome evolution
Fluorescence *in situ* hybridization
Genome size
Phylogeny
rDNA

ABSTRACT

Most of the cultivated chenopods are polyploids, but their origin and evolutionary history are still poorly understood. Phylogenetic analyses of DNA sequences of four plastid regions, nrITS and nuclear 5S rDNA spacer region (NTS) of two tetraploid chenopods ($2n = 4x = 36$), Andean *C. quinoa* and North American *C. berlandieri*, and their diploid relatives allowed inferences of their origin. The phylogenetic analyses confirmed allotetraploid origin of both tetraploids involving diploids of two different genomic groups (genomes A and B) and suggested that these two might share very similar parentage.

The hypotheses on the origin of the two allopolyploid species were further tested using genomic *in situ* hybridization (GISH). Several diploid *Chenopodium* species belonging to the two lineages, genome A and B, suggested by phylogenetic analyses, were tested as putative parental taxa. GISH differentiated two sets of parental chromosomes in both tetraploids and further corroborated their allotetraploid origin. Putative diploid parental taxa have been suggested by GISH for *C. quinoa* and *C. berlandieri*. Genome sizes of the analyzed allotetraploids fit nearly perfectly the expected additive values of the putative parental taxa. Directional and uniparental loss of rDNA loci of the maternal A-subgenome was revealed for both *C. berlandieri* and *C. quinoa*.

1. Introduction

The genus *Chenopodium sensu lato* (*s.l.*) is one of the largest genera in the family Amaranthaceae and comprises mostly annual herbs (Aellen, 1960; Jellen et al., 2011). Recent molecular phylogenetic analysis indicated that *Chenopodium s.l.* is a polyphyletic genus encompassing at least six different evolutionary lineages (*Dysphanieae*, *Anserineae*, *Chenopodiastrum*, *Oxybasia*, *Lipandra* and *Chenopodium sensu stricto*; Fuentes-Bazan et al., 2012a). Most of the economical important *Chenopodium* species belong to *Chenopodium sensu stricto* (*s.s.*) and represent either crops (e.g. diploid *C. pallidicaule* or tetraploid *C. quinoa*) or noxious weeds (e.g. Eurasian hexaploid *C. album* or diploid *C. ficifolium*; Jellen et al., 2011). The most important cultivated *Chenopodium* species is *C. quinoa*, a traditional Andean seed crop that has tremendously grown in popularity in Europe and North America in the past decades. Its grain has an excellent balance of carbohydrates, lipids, and proteins with essential amino acid compositions optimal for human nutrition (Popenoe et al., 1989). Earlier analyses of isozymes and molecular markers revealed that *Chenopodium quinoa* is closely related to another American tetraploid, *Chenopodium berlandieri* (Kistler and Shapiro, 2011; Maughan et al., 2006; Wilson, 1990). *C. berlandieri* includes both cultivated and wild plants. Mesoamerican *C. berlandieri* subsp. *nuttalliae* has two vegetable cultigens, Huauzontle and Quelite, and the seed crop Chia roja whereas wild *C. berlandieri* subsp. *berlandieri* is broadly distributed in North America (Kistler and Shapiro, 2011; Wilson, 1990). Both *C. quinoa* and *C. berlandieri* are tetraploids with $2n = 4x = 36$ and possess symmetrical karyotypes with small, mostly metacentric and submetacentric chromosomes (Bhargava et al., 2006; Palomino et al., 2008). They both have small genome sizes (*C. quinoa*, 1C = 1.49 pg; *C. berlandieri* subsp. *nuttalliae*, 1C = 1.52 pg; Kolano et al., 2012a; Palomino et al., 2008). According to the recent phylogeny-based classification of *Chenopodium s.l.* both species belong to *Chenopodium s.s.* together with other polyploids such as the *C. album* polyploid complex and some other diploid species (Fuentes-Bazan et al., 2012a). At least three evolutionary lineages can be distinguished among the diploids of *Chenopodium s.s.* (Kolano et al., 2015; Štorchová et al., 2014; Walsh et al., 2015; Ward, 2000). The first lineage includes

* Corresponding author.
E-mail address: bozena.kolano@us.edu.pl (B. Kolano).

http://dx.doi.org/10.1016/j.ympev.2016.04.009
1055-7903/

American diploids (e.g. *C. neomexicanum*, *C. incanum*) which possess genome A. The second lineage consists of Old World diploids (*C. ficifolium* and relative diploids) which represent genome B. The third evolutionary lineage consists solely of *C. vulvaria* (Fuentes-Bazan et al., 2012b; Kolano et al., 2015; Štorchová et al., 2014; Walsh et al., 2015). *C. berlandieri* and *C. quinoa* are both allotetraploids (Ward, 2000), the view supported by recent phylogenetic studies based on low copy DNA sequences (Štorchová et al., 2014; Walsh et al., 2015). An earlier study based on 5S rDNA NTS sequences suggested that *C. berlandieri* and *C. quinoa* share at least one common ancestor (Kolano et al., 2011; Maughan et al., 2006). Various diploid species have been suggested as parental genomes of these two allotetraploid species in several previous studies. Morphological and ribozyme data suggested the North American diploids *C. watsonii* or *C. neomexicanum* as potential progenitors of the two cultivated chenopods (Walters, 1987). Recent phylogenetic studies, on the other hand, suggested species related to Eurasian diploids (*C. suecicum* or *C. ficifolium*) and species related to North American *C. standleyanum* or *C. incanum* as diploid progenitors for *C. quinoa* (Štorchová et al., 2014; Walsh et al., 2015). One of the B genome species was also proposed to be involved in hybridization which gave raise to hexaploid species *C. album* and *C. giganteum* (members of *C. album* complex). The second parent of these polyploids was suggested to be Eurasian tetraploid *C. strictum* with genome constitution CD (Walsh et al., 2015; Mandak et al., 2012).

Origin of allopolyploids can also be elucidated using genomic *in situ* hybridization (GISH), which allows physical mapping of putative parental taxa in chromosomes of the polyploid (Markova and Vyskot, 2009). GISH has been successfully used to unravel the genomic origins of a number of polyploid taxa (e.g. tetraploid *Arachis hypogaea*, triploid *Allium* × *cornutum*; Fredotovic et al., 2014; Seijo et al., 2007). However, until now it has not been used in the analysis of the origin and genome composition of tetraploid *Chenopodium* species.

Allopolyploidy resulting from interspecific hybridization accompanied by whole genome duplication may induce a number of genetic and genomic changes that can occur rapidly within a few generations (Ma and Gustafson, 2005; Leitch and Leitch, 2008). One genomic character that is often affected by polyploidy is genome size, which can experience increase, decrease or remain the same as the expected additive value of progenitor taxa, depending on the study system (Leitch and Bennett, 2004; Leitch et al., 2008).

Another large-scale alteration common to allopolyploid genomes affects the nuclear ribosomal RNA genes encoding 18S, 5.8S and 25S rRNAs (35S rRNA), and to a lesser extent, 5S rRNA genes. Each unit of 35S rDNA contains conserved genic regions and the variable internal transcribed (ITS) and intergenic spacer (IGS) regions. In addition to a loss of rDNA loci, concerted evolution acting on tandemly repeated 35S rDNA units has been proposed to occur via recombination-based mechanisms (Matyasek et al., 2012), often accompanied by changes in unit copy numbers (Renny-Byfield et al., 2011). The number and localization of rDNA loci have recently been reported for several *Chenopodium s.l.* species and revealed low number of rDNA loci per chromosome set (Kolano et al., 2012b, 2015). *C. quinoa* possesses one 35S rDNA locus and two loci (one interstitial and one subterminal) of 5S rDNA per haploid genome. One or two loci of 35S rDNA and two or three loci of 5S rDNA, all in subterminal chromosomal positions, were observed in *C. berlandieri*, depending on the accession (Maughan et al., 2006). Comparative analysis of rDNA loci number in polyploid and related diploid chenopods suggested that rDNA underwent reorganization during polyploid genome evolution (Kolano et al., 2012b).

The present study addresses the origin of the tetraploids *C. quinoa* and *C. berlandieri* and attempts to identify their diploid ancestors using phylogenetic analyses of nuclear ribosomal internal transcribed spacer (nrITS), 5S rDNA nontranscribed spacer (NTS) and plastid (cpDNA) DNA sequences and to test the inferred parentage by mapping of parental genomes in the polyploids using genomic *in situ* hybridization (GISH). The aim of this study is also to analyze the patterns and dynamics of the evolution of both types of ribosomal loci (5S and 35S rDNA) in polyploids, as well as to address the evolution of genome size in polyploids in comparison to their putative parental taxa.

2. Materials and methods

2.1. Plant material

Thirty-two accessions representing 17 diploid species and 11 accessions of two tetraploid taxa, C. *quinoa* and *C. berlandieri*, all in the *Chenopodium s.s.*, grown from seeds and cultivated in a greenhouse facility of the University of Silesia were analyzed (Table 1). The final sample comprised 15 of approximately 20–25 *Chenopodium s.s.* species found in the Americas (two species, *C. pallidicaule* and *C. petiolare* from South America, rest from North America; Walsh et al., 2015) and three Eurasian diploid *Chenopodium s.s.* species. Vouchers are deposited at the Herbarium KTU (University of Silesia, Chorzów, Poland).

2.2. DNA amplification, cloning and sequencing

Genomic DNA was isolated from fresh leaf tissue using the CTAB method (Doyle and Doyle, 1987). DNA concentration was measured using a NanoDrop spectrophotometer (ND-1000, peqLab, Erlangen, Germany).

The 5S rDNA NTS was amplified using a primer pair anchored in the 5S rRNA genic region (Table S1; Maughan et al., 2006). PCR amplification was carried out using a GeneAmpPCR System 9700 thermocycler (Applied Biosystems, USA). The reaction mixture contained 30 ng of DNA, 0.4 μM of each primer (Genomed, Warsaw, Poland), 0.1 mM of each dNTPs (Sigma–Aldrich, Steinheim, Germany), 1× PCR buffer (including 2.0 mM $MgCl_2$) and 1U of Taq DNA polymerase (Sigma–Aldrich, Steinheim, Germany). The PCR program consisted of an initial denaturation step at 94 °C for 1.5 min, followed by 40 cycles of 35 s at 94 °C, 30 s at 55 °C, and 1 min at 72 °C, with a final elongation step of 5 min at 72 °C (Maughan et al., 2006). PCR products were gel-purified using QIAquick Gel Extraction Kit (Qiagen, Germany), and ligated into pGEM-T Easy vectors (Promega, Madison, USA) following the manufacturer's instructions. Five to fifteen randomly chosen recombinant colonies (for details see Phylogenetic analysis) were selected for plasmid DNA isolation using the standard miniprep method (Sambrook et al., 1987). The cloned fragments were sequenced using M13 universal primers and BigDye terminator v3.1 Cycle Sequencing Kit technology (Applied Biosystems) in 3730 × l DNA Analyzer (Applied Biosystems).

The nuclear ITS region (ITS1, the intervening 5.8S rDNA and ITS2) of the 35S rRNA gene was amplified as described earlier in Kolano et al. (2015). The four non-coding plastid DNA spacer regions, *rpl*32-*trn*L, *rps*16-*trn*K, *pet*L-*psb*E and *psb*D-*trn*T, were amplified using primers from Shaw et al., 2007 (Table S1). The PCR reaction contained 1× RedTaq PCR ReadyMix (Sigma–Aldrich, Steinheim, Germany), 0.5 μM of each forward and reverse primer (Genomed, Warsaw, Poland) and 50 ng of DNA template. Amplification was performed using GeneAmpPCR System 9700 (Applied Biosystems, USA) under conditions described earlier (*pet*L-*psb*E region: Garcia et al., 2012; three remaining spacer regions: Bloch et al., 2009). PCR products were treated with *E. coli* Exonuclease I and FastAP Termosensitive Alkaline Phosphatase (Thermo Fisher Scientific, Waltham, USA) according to the manufacturer's instruc-