Available online at www.sciencedirect.com
SCIENCE DIRECT®
Organisms, Diversity & Evolution 5 (2005) 275–283
ORGANISMS DIVERSITY & EVOLUTION
www.elsevier.de/ode

Allopolyploid evolution in Geinae (Colurieae: Rosaceae) – building reticulate species trees from bifurcating gene trees

Jenny E.E. Smedmark[a,b,*], Torsten Eriksson[a,b], Birgitta Bremer[a,b]

[a]*Bergius Foundation, Royal Swedish Academy of Sciences, Box 50017, SE-104 05 Stockholm, Sweden*
[b]*Department of Botany, Stockholm University, SE-106 91 Stockholm, Sweden*

Received 17 September 2004; accepted 27 December 2004

Abstract

A previous phylogenetic study of paralogous nuclear low-copy granule-bound starch synthase (GBSSI) gene sequences from polyploid and diploid species in Geinae indicated that the clade has experienced two major allopolyploid events in its history. These were estimated to have occurred several million years ago. In this extended study we test if the reticulate phylogenetic hypothesis for Geinae can be maintained when additional sequences are added. The results are compatible with the hypothesis and strengthen it in minor aspects. We also attempt to identify extant members of one of the inferred ancestral lineages of the allopolyploids. On the basis of previous molecular phylogenies, one specific group has been proposed to be the descendants of this taxon. However, none of the additional paralogues belong to this ancestral lineage. A general method is proposed for converting a bifurcating gene tree, with multiple paralogous low-copy gene sequences from allopolyploid taxa, into a reticulate species tree.

Keywords: Geinae; Rosaceae; Allopolyploidy; Reticulate phylogeny; GBSSI

Introduction

The group Geinae Schulze-Menz has been suggested to have been shaped by allopolyploidy (Gajewski 1957, 1958), an evolutionarily important process among plants (Stebbins 1971; Grant 1981) involving interspecific hybridisation followed by chromosome doubling. To address this question, a phylogenetic study of low-copy granule-bound starch synthase (GBSSI) gene sequences from species in this clade was performed (Smedmark et al. 2003). When reconstructing allopolyploid speciation with single- or low-copy nuclear gene sequences, paralogous copies from a putative allopolyploid species are analysed phylogenetically together with sequences from closely related species. This method provides a direct reconstruction of phylogenetic relationships of parental lineages of allopolyploid taxa (Sang and Zhang 1999). In an allopolyploid, there are homoeologous copies of the gene, contributed by different ancestral taxa. These gene copies will be sister to orthologues in the ancestral taxa in a phylogenetic tree, rather than to each other. Autopolyploids, on the other hand, have paralogous loci that will be more closely related to each other than to orthologous sequences in closely related species. Several studies have used this method successfully to infer complex reticulate relationships among plant species, e.g., in *Paeonia* (Sang

*Corresponding author. Current address: Swedish Museum of Natural History, Department of Phanerogamic Botany, Box 50007, SE-104 05 Stockholm, Sweden.
E-mail address: jenny.smedmark@nrm.se (J.E.E. Smedmark).

1439-6092/$ - see front matter
doi:10.1016/j.ode.2004.12.003

and Zhang 1999), *Gossypium* (Cronn and Wendel 2004), *Glycine* (Doyle et al. 2004), and *Elymus* (Mason-Gamer 2004). The study of Geinae also provided strong evidence for reticulate historical relationships between lineages (Smedmark et al. 2003). A hypothesis suggesting two allopolyploidisation events with major impact on extant species diversity in the group was formulated, based on the GBSSI-1 gene tree. The analysis included paralogous copies from seven polyploid species in Geinae, along with copies from two diploid species in this group. These diploids had been shown to belong to the same clade (Smedmark and Eriksson 2002), which is the only one in Geinae known to contain extant diploids. This previous study only identified one of the two original ancestral lineages of the polyploids. Although the phylogenetic position of the other lineage can be inferred from the gene tree, it turned out that no extant representatives were included in the study. With reference to a previous phylogenetic study with wider taxon sampling (Smedmark and Eriksson 2002), a specific clade, the sister group of the remainder of Geinae, was suggested to be potential descendants of this second parental lineage. In this extended study we test whether the hypothesis proposed by Smedmark et al. (2003) about reticulate relationships within Geinae holds true with increased sampling of GBSSI paralogues. We also attempt to identify the unknown ancestral lineage of the allopolyploids.

It is worth noting that, despite the fact that low-copy gene sequences have been used successfully in several studies to infer the ancestry of polyploids, no method for converting bifurcating gene trees into reticulate species tree has been published. We here describe the method that we have used, a method that seems to be generally applicable to this type of problems.

Materials and methods

Samples, DNA extraction, amplification, cloning, and sequencing

Four species in Colurieae and two Rosoideae species outside this clade were selected (Table 1) to be incorporated in an existing data set (Smedmark et al. 2003). One species from the suggested paternal lineage was included, the decaploid *Oncostylus leiospermus*, as well as the closest relatives of Geinae, *Sieversia* Willd. and *Fallugia* Endl., and the tetraploid *Novosieversia glacialis*. Based on data from the *trnL–trnF* region (Smedmark and Eriksson 2002), the last of the above species was placed in a group that would correspond to the hypothesised allopolyploid clade.

DNA extraction, PCR amplification, and cloning followed the procedures described by Smedmark et al. (2003). Sampling was limited to available living specimens due to difficulties in amplifying GBSSI using extractions from dried material. The amplification of GBSSI-1 for *N. glacialis*, *O. leiospermus*, and *Sieversia pentapetala* was carried out with the primers 1F1C and 9R1C (Smedmark et al. 2003), whereas *Fallugia paradoxa*, *Filipendula vulgaris*, and *Fragaria vesca* were amplified with the primers 1F and 9R (Alice 1997).

Sequencing reactions were performed with a DYEnamic ET termination cycle sequencing premix *kit*

Table 1. List of taxa

Species	Voucher	Origin	Clone	No. clones screened	EMBL accession
Fallugia paradoxa (D. Don) Endl.	T. Eriksson No. 796 (SBT)	USA (Colorado)	*F. paradoxa* 2–2	7	AJ871485
Sieversia pentapetala (L.) Greene	T. Eriksson No. 749 (SBT); cult. Göteborg Botanic Garden	Unknown	*S. pentapetala* 2–2	5	AJ871484
Novosieversia glacialis (Adams) F.Bolle	A. Batten	USA (Alaska)	*N. glacialis* 1–2	22	AJ871488
Oncostylus leiospermus (Petrie) F. Bolle	M. Chase; cult. Royal Botanic Gardens Kew	New Zealand	*O. leiospermus* 1–3, *O. leiospermus* 1–6, *O. leiospermus* 2–2	14	AJ871490, AJ871489, AJ871491
Fragaria vesca L.	T. Eriksson & J.E.E. Smedmark 43 (SBT)	Sweden	*F. vesca* 3–2	2	AJ871486
Filipendula vulgaris Moench.	T. Eriksson 821 (SBT)	Sweden	*F. vulgaris* 3–3	1	AJ871487

The list follows the classification of Bolle (1933) and includes information on vouchers, origins, clones included in the analyses, numbers of GBSSI-1 clones screened, and EMBL accession numbers of the sequences (for information on *Rosa multiflora* and *Rubus odoratus* see Evans et al. 2000, for the remaining sequences see Smedmark et al. 2003).