Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 598–611
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Allohexaploidy, introgression, and the complex phylogenetic history of *Elymus repens* (Poaceae)

Roberta J. Mason-Gamer *

The University of Illinois at Chicago, Department of Biological Sciences, MC 066, Chicago, IL 60607, USA

Received 28 June 2007; revised 11 January 2008; accepted 13 February 2008
Available online 10 March 2008

Abstract

The phylogenetic position of hexaploid *Elymus repens* within the tribe Triticeae (Poaceae) was examined using cloned sequences from the low-copy nuclear genes encoding phosphoenolpyruvate carboxylase (pepC) and β-amylase. A previous analysis of *E. repens* using data from the nuclear granule-bound starch synthase I (GBSSI) gene had yielded five phylogenetically distinct gene copies, two more than expected from hexaploidy alone. The three gene trees share three distinct *E. repens* clades, suggesting that *E. repens* contains three phylogenetically divergent genomes, contributed by *Hordeum*, *Pseudoroegneria*, and an unknown donor. The two additional GBSSI sequences, including one that was apparently derived from outside of the tribe, appear to reflect past introgression of GBSSI sequences into the *E. repens* genome. On all three trees, the *Hordeum*-like *E. repens* sequences are polyphyletic within *Hordeum*, and the trees are in conflict with regard to the placement of these sequences within *Hordeum*, highlighting multiple contributions from *Hordeum* to *E. repens*.

Keywords: *Elymus*; Triticeae; Polyploidy; Introgression; Phylogeny

1. Introduction

Over the last two decades, phylogenetic analyses of chloroplast and nuclear DNA data have clarified the relationships between many polyploid species and their diploid relatives. In some of the earliest of these, chloroplast DNA (cpDNA) restriction site data were successfully used to clarify polyploid origins (Doyle et al., 1990; Soltis and Soltis, 1989; Soltis et al., 1989a,b), and cpDNA trees remain important components of phylogenetic analyses of polyploids (Guggisberg et al., 2006; Johnson and Johnson, 2006; Sharbel and Mitchell-Olds, 2001; Tate and Simpson, 2003). The nuclear ribosomal rDNA repeat, including the internal transcribed spacer (ITS) region, has the added potential to identify multiple parental genome donors (Guggisberg et al., 2006; Hughes et al., 2002; Rauscher et al., 2002; Sang et al., 1995; Soltis and Soltis, 1991; Tate and Simpson, 2003; Widmer and Baltisberger, 1999), although the widespread phenomenon of ITS homogenization following polyploid formation (reviewed in Álvarez and Wendel, 2003) can limit the utility of ITS sequences in studies of polyploid origins. Single- and low-copy nuclear gene trees are now common in studies of polyploid phylogeny, and have been used to verify the existence of suspected polyploids (e.g., Smith et al., 2006; Winkworth and Donoghue, 2004), identify genome donors (e.g., Emshwiller and Doyle, 2002; Fortune et al., 2007; Ge et al., 1999; Helfgott and Mason-Gamer, 2004; Johnson and Johnson, 2006; Mason-Gamer, 2001; Petersen et al., 2006; Popp et al., 2005; Slotte et al., 2006), demonstrate multiple polyploid origins (Doyle et al., 2002), clarify combinations of polyploidy and hybridization or introgression (e.g., Cronn et al., 2003; Ferguson and Sang, 2001; Lihová et al., 2006; Mason-Gamer, 2004), and examine gene evolution or gene silencing in polyploids (Barrier et al., 2001; Cronn et al., 1999; Ford and Gottlieb, 2002; Hughes et al., 2002).

The present study clarifies the evolutionary history of the hexaploid species *Elymus repens* (L.) Gould, a

* Fax: +1 312 413 2435.
E-mail address: robie@uic.edu

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.02.008

widespread, morphologically variable species in the wheat tribe (Triticeae), using sequence data from low-copy nuclear genes. This native of Europe and Asia has become a problematic weed throughout much of northern North America since its introduction about a century ago (Batcher, 2002). The species is allohexaploid ($2n = 42$), with generally normal, diploidized meioses in which the chromosomes form 21 bivalents (Cauderon, 1958; Dewey, 1961). Cytogenetic studies suggest that two of the three haploid chromosome sets are closely related, with the third set distinct from the other two (Cauderon, 1958; Dewey, 1961). The identities of the genome donors have been investigated using cytogenetic (Assadi and Runemark, 1995; Cauderon and Saigne, 1961; Dewey, 1970b, 1976) and GISH (Ørgaard and Anamthawat-Jónsson, 2001) techniques, but the results are not in complete agreement. Cytogenetic studies of hybrids between *E. repens* and two different *Pseudoroegneria* species (Dewey, 1970b, 1976) suggested that the two similar genome sets of *E. repens* were derived from *Pseudoroegneria* (genome designation **St**), while in a more recent GISH-based study (Ørgaard and Anamthawat-Jónsson, 2001), *P. spicata* probes did not hybridize strongly to the *E. repens* genome, leading the authors to question whether *Pseudoroegneria* is a genome donor. In the same GISH study, *Hordeum californicum* probes hybridized to the *E. repens* genome, indicating the presence of an **H** genome, in agreement with some cytogenetic studies (Cauderon and Saigne, 1961), but Dewey (1984) doubted the presence of the **H** genome on morphological grounds. The cytogenetic analysis of Assadi and Runemark (1995) was one of few that simultaneously indicated both **St** and **H** genomes in *E. repens*: based on chromosome pairing in hybrids between *E. repens* and the presumed **StStStStHH** hexaploid *E. transhyrcanus*, they concluded that *E. repens* shared the **StStStStHH** genome complement. Although the analyses of the *E. repens* genome content are not in complete agreement, they collectively suggest that the two similar genome sets were derived from *Pseudoroegneria* (genome designation **St**) and the third from *Hordeum* (genome **H**). However, studies attempting to ascertain polyploid genome content through observations of meiotic chromosome behavior in hybrids do have some shortcomings. First, they are based on the assumption that chromosome pairing in such hybrids is a direct indicator of relatedness among chromosomes, an assumption that has long been questioned (e.g., Darlington, 1932; de Wet and Harlan, 1972; Seberg and Petersen, 1998). Second, studies of hybrid meioses typically include only a limited number of hybrid combinations. In these cases, patterns of chromosome pairing in hybrids between polyploids and diploids can potentially identify candidate genome donors, but they cannot necessarily provide positive identifications if there are other, related taxa that have not been tested.

The hypothesis that *E. repens* has an **StStStStHH** genome complement was recently examined for six individuals using molecular data from the chloroplast genome and from the single-copy nuclear gene encoding granule-bound starch synthase I (GBSSI or *waxy*; Mason-Gamer, 2004). The chloroplast data identified *Pseudoroegneria* as the maternal genome donor, and the GBSSI data revealed **St** and **H** genome contributions from *Pseudoroegneria* and *Hordeum*, respectively. There were also three additional, phylogenetically distinct GBSSI gene copies within *E. repens*, representing apparent contributions from *Taeniatherum* (genome designation **Ta**), from an unknown donor within the Triticeae, and from a second unknown-donor outside of the tribe. Based on these results, a combination of allohexaploidy and introgression was hypothesized, but it remained unclear which, if any, of the three unexpected gene copies represented an entire third genome set, or which were acquired through introgression. In the present study, phylogenetic analyses of *E. repens* sequences from two additional, unlinked single-copy nuclear loci are used to further clarify the evolutionary history of *E. repens*. Under the assumption that *Pseudoroegneria* and *Hordeum* were donors to *E. repens*, three hypotheses are considered: (1) the basic *E. repens* genome complement is **StStStStHH**, as suggested by the cytogenetic studies, and all three additional GBSSI gene copies have been acquired through introgression; (2) *Taeniatherum* is a genome donor (giving *E. repens* an **StStHHTaTa** complement), and the GBSSI gene copies from both of the unknown donors were acquired through introgression; or (3) the Triticeae unknown-donor sequence represents a complete genome (giving an **StStHHXX** complement, where the origin of **X** is unknown), and the GBSSI gene copies from both *Taeniatherum* and the more distant unknown donor were acquired through introgression. While other hypotheses are conceivable, the three listed here seemed the most reasonable. The ability to distinguish among the hypotheses using comparisons among gene trees is based on the assumption that gene copies present on all three gene trees represent contributions of entire genomes, while copies that are unique to just one gene tree probably represent introgression.

2. Materials and methods

2.1. Molecular data sets

Separate analyses were run for the three nuclear gene data sets, each of which includes sequences from between five and six *E. repens* individuals (Table 1). The GBSSI *E. repens* sequences were previously published (Mason-Gamer, 2004), and pepC and β-amylase *E. repens* sequences were obtained for this analysis. These were analyzed with a broadly representative sample of monogenomic, mostly diploid species from throughout the Triticeae. Most of the sequences from the monogenomic species were previously published, but all three data sets were augmented with new sequences from *Hordeum* and *Pseudoroegneria*, the two presumed genetic contributors to *E. repens* (Table 2). Based on genomic studies of grass crop species