Fungal Genetics and Biology 49 (2012) 250–261

Contents lists available at SciVerse ScienceDirect

Fungal Genetics and Biology

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

PRP8 inteins in species of the genus *Botrytis* and other ascomycetes

Annika A.M. Bokor [a], Linda M. Kohn [b], Russell T.M. Poulter [a], Jan A.L. van Kan [c,*]

[a] Biochemistry Department, University of Otago, 710 Cumberland Street, Dunedin 9054, New Zealand
[b] Department of Ecology and Evolutionary Biology, University of Toronto, 3359 Mississauga Road North, Mississauga, Ontario, Canada L5L 1C6
[c] Laboratory of Phytopathology, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

ARTICLE INFO

Article history:
Received 2 November 2011
Accepted 11 January 2012
Available online 20 January 2012

Keywords:
Botrytis
Intein
Horizontal transmission
Genetic footprint

ABSTRACT

The mobile elements termed inteins have a sporadic distribution in microorganisms. It is unclear how these elements are maintained. Inteins are intervening protein sequences that autocatalytically excise themselves from a precursor. Excision is a post-translational process referred to as 'protein splicing' in which the sequences flanking the intein are ligated, reforming the mature host protein. Some inteins contain a homing endonuclease domain (HEG) that is proposed to facilitate propagation of the intein element within a gene pool. We have previously demonstrated that the HEG of the PRP8 intein is highly active during meiosis in *Botrytis cinerea*. Here we analysed the *Prp8* gene status in 21 additional *Botrytis* species to obtain insight into the mode of intein inheritance within the *Botrytis* lineage. Of the 21 species, 15 contained a PRP8 intein whereas six did not. The analysis was extended to closely related (Sclerotiniaceae) and distantly related (Ascomycota) taxa, focussing on evolutionary diversification of the PRP8 intein, including their possible acquisition by horizontal transfer and loss by deletion. Evidence was obtained for the occurrence of genetic footprints of previous intein occupation. There is no compelling evidence of horizontal transfer among species. Three distinct states of the *Prp8* allele were identified, distributed over different orders within the Ascomycota: an occupied allele; an empty allele that was never occupied; an empty allele that was presumably previously occupied, from which the intein was precisely deleted. The presence of the genetic footprint identifies 20 species (including *Neurospora crassa*, *Magnaporthe oryzae* and *Fusarium oxysporum*) that previously contained the intein but have lost it entirely, while only 18 species (including *Podospora anserina* and *Fusarium graminearum*) appear never to have contained a PRP8 intein. The analysis indicates that inteins may be maintained in an equilibrium state.

1. Introduction

Inteins (*in*ternal pro*teins*) are intervening protein sequences that autocatalytically excise themselves from a protein precursor without requiring host cofactors or accessory proteins (Kane et al., 1990; Kawasaki et al., 1997; Xu and Perler, 1996). Excision is a post-translational cellular process referred to as protein splicing in which the two sequences flanking the intein, the N- and C-exteins (*ex*ternal pro*teins*) are ligated, reforming a *bona fide* peptide bond to generate the mature host protein (Davis et al., 1992; Evans and Xu, 1999; Muir et al., 1998). There are two main types of inteins, the full-length inteins and mini-inteins. Full-length inteins carry both the N and C-terminal splicing domains and an additional homing endonuclease (Fig. 1) (Belfort and Roberts, 1997; Fsihi et al., 1996). Almost all intein homing endonucleases (HEGs) are members of the LAGLIDADG endonuclease family, after the sequence of the highly conserved active site.

The endonuclease domain encodes a highly specific homing endonuclease that is believed to facilitate the propagation of the intein element within a gene pool. It has been demonstrated in both *Saccharomyces cerevisiae* (Gimble and Thorner, 1992) and *Botrytis cinerea* (Bokor et al., 2010) that the intein HEG activity is restricted to meiosis. During meiosis, the HEG introduces a double-stranded DNA break in a cognate empty allele. The break triggers the homologous recombination repair system resulting in the copying of the intein sequence into the empty allele. The process results in biased allele segregation in the meiotic products, so called super-Mendelian inheritance of the intein (Gimble and Thorner, 1992; Bokor et al., 2010). Some inteins apparently carry inactive (degenerate) HEGs (Koufopanou et al., 2002). Mini-inteins carry the N and C-terminal splicing domains but lack the HEG (Fig. 1) (Derbyshire et al., 1997), and are therefore unable to propagate.

Since the discovery of the first intein, more than 400 inteins ranging from ~1 kb to 2.5 kb, have been identified in Archaea, Eubacteria and Fungi, in both organelles and nuclei (Anraku, 1997; Davis et al., 1994; Hodges et al., 1992; Liu and Hu, 1997; Pietrokovski, 1996; Wang and Liu, 1997), as compiled in the

* Corresponding author.
E-mail address: jan.vankan@wur.nl (J.A.L. van Kan).

1087-1845/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.fgb.2012.01.001

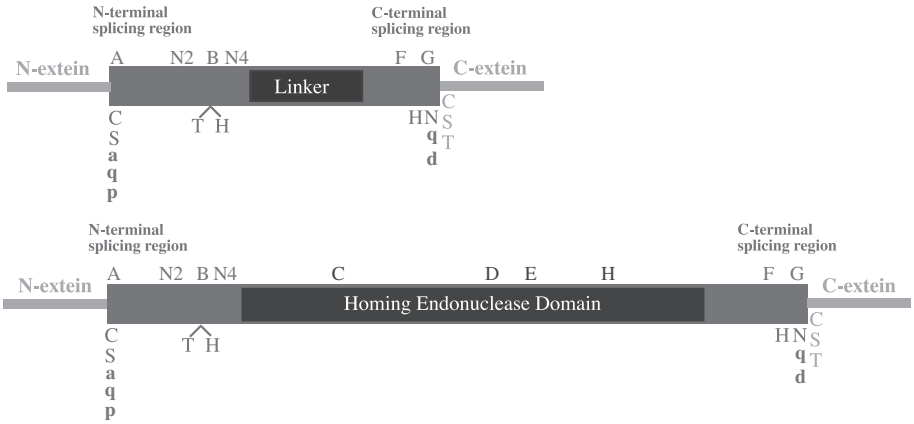

Fig. 1. *Intein structure*. The host protein (extein) disrupted by the intein is indicated. Mini-inteins (top) have a short linker domain separating the splicing domains and full-length inteins (bottom) have an additional homing endonuclease domain. Highly conserved intein residues are indicated below the inteins (lower case – amino acids in polymorphic inteins that may splice by a modified mechanism) and the motifs (A–H, N2 and N4) are displayed above the intein structures.

database InBase (http://www.neb.com/neb/inteins.html; Perler, 2002). In many cases inteins are found within genes that encode proteins involved in nucleic acid metabolism or DNA replication (Perler et al., 1997; Liu, 2000). The intein is usually integrated within a highly conserved region of the 'host' protein (Pietrokovski, 2001).

Most reports of inteins have been in the Archaea and Eubacteria, but many have also been identified in the nuclear genomes of fungi. The first intein described was the full-length VMA1 intein in *S. cerevisiae* (PI-SceI) (Hirata et al., 1990). The second nuclear intein was a mini-intein in the *Prp8* gene of *Cryptococcus neoformans* (Butler et al., 2001), followed by several other PRP8 inteins (Liu and Yang, 2004; Butler et al., 2006; Bokor et al., 2010). Inteins have now been found in several specific genomic locations in fungi (Butler et al., 2005) and other microorganisms (Goodwin et al., 2006).

Given that HEG-containing inteins show super-Mendelian inheritance during meiosis it might be expected that, if an intein-encoding allele is present in a gene pool, the frequency of the intein will increase over time, eventually reaching fixation. This expectation assumes that the presence of the intein is selectively neutral or near neutral. However, this prediction presents an enigma. Once an intein reaches fixation there will be no remaining targets (empty alleles). The HEG would therefore be unable to initiate recombination (homing). In the absence of any selection for activity, the HEG would be predicted to degenerate and eventually be lost. The enigma is that nuclear inteins with active HEGs still occur in various organisms (Gimble and Thorner, 1992; Bokor et al., 2010). One hypothesis to reconcile these observations (Goddard and Burt, 1999) proposes that HEG-containing inteins have a 'life-cycle' in which the HEG-containing intein invades a new species through horizontal transfer, that the intein then increases in frequency eventually coming to fixation followed by degeneration and eventual loss of both the HEG and the intein. Mini-inteins could be viewed as inteins that are in the process of being reduced and eventually perhaps lost. This model assumes that inteins remain active during evolution as a result of repeated horizontal transfers. This cyclic model is not a mathematical necessity; it is possible to have all three states (inteins with active HEGs, inteins with inactive HEGs and unoccupied alleles) persisting indefinitely in a species, either at a stable equilibrium or in a stable low amplitude limit cycle (Burt and Koufopanou, 2004). The distinctive characteristic of such equilibrium models is the assumption that the presence of the intein is not selectively neutral but is slightly disadvantageous. The two models, in which active inteins are maintained by horizontal transmission (cyclic model) or mutation and selection (equilibrium model), make distinct predictions about the similarity of inteins in related species and the probability of an intein being at fixation. One fungal intein, the VMA1 intein of Saccharomycete yeasts has been the subject of such an analysis (Koufopanou et al., 2002). The present analysis attempts to compare the similarity of the PRP8 inteins of various filamentous fungi, especially members of the genus *Botrytis*. The PRP8 intein is the most widely distributed of any nuclear intein, reported from diverse ascomycetes, basidiomycetes and chytridiomycetes.

We have previously demonstrated that the HEG of the PRP8 intein is highly active in natural populations of *Botrytis cinerea* (Bokor et al., 2010). In addition to *B. cinerea*, the *Botrytis* genus contains 21 recognised species as well as one allodiploid hybrid species, *B. allii* (Staats et al., 2005). The rationale of the present study is that the *Prp8* gene status in other *Botrytis* species could give an insight into the mode of intein inheritance within the evolutionary time span of speciation in the *Botrytis* lineage with comparisons to the more distantly related lineages in the family, Sclerotiniaceae. This intermediate evolutionary timescale presents a valuable framework for capturing patterns of invasion, fixation and extinction. We anticipate that with this sample, vertical transmission will result in a molecular phylogeny in which the phylogeny of the host species and the phylogeny of inteins are congruent. Conversely, horizontal transmission is likely to result in the two phylogenies being dissimilar.

2. Materials and methods

2.1. Polymerase chain reaction

All sequences were amplified using Taq Polymerase (Roche) and the following PCR conditions: initial denaturation at 94 °C for 2 min; denaturation at 94 °C for 1 min; annealing at 45–52 °C (depending on primer combination) for 1 min; extension at 72 °C for 2 min. Steps 2 through 4 were repeated 34 times followed by a final extension at 72 °C for 7 min. The primers used are listed in Table 1.

2.2. Cloning and sequencing of PRP8 inteins from Botrytis species and related genera

DNA from *Botrytis* was isolated as described in Staats et al. (2005). DNA from Sclerotiniaceae other than *Botrytis* was isolated as described in Andrew and Kohn (2009). Overlapping amplicons of the full-length PRP8 inteins were cloned using the pGEM-T Easy vector system (Promega) following the manufacturer's instructions. The clones were sequenced, and the sequences of the PRP8