Protist, Vol. 162, 154–167, January 2011
http://www.elsevier.de/protis
Published online date 22 July 2010

Protist

ORIGINAL PAPER

Introns, Alternative Splicing, Spliced Leader *trans*-Splicing and Differential Expression of *pcna* and *cyclin* in *Perkinsus marinus*

Huan Zhang[a,1], Christopher F. Dungan[b], and Senjie Lin[a]

[a]Department of Marine Sciences, University of Connecticut, Groton, CT 06340
[b]Maryland Department of Natural Resources, Cooperative Oxford Laboratory, Oxford, MD 21654

Submitted December 8, 2009; Accepted March 6, 2010
Monitoring Editor: Frank Seeber

To gain understanding on the structure and regulation of growth-related genes of the parasitic alveolate *Perkinsus marinus*, we analyzed genes encoding proliferating cell nuclear antigen (*pcna*) and cyclins (*cyclin*). Comparison of the full-length cDNAs with the corresponding genomic sequences revealed *trans*-splicing of the mRNAs of these genes with a conserved 21-22 nt spliced leader. Over 10 copies of *pcna* were detected, with identical gene structures and similar nucleotide (nt) sequences (88-99%), encoding largely identical amino acid sequences (aa). Two distinct types of *cyclin* (*Pmacyclin*1 and *Pmacyclin*2) were identified, with 66-69% nt and 81-85% aa similarities. *Pmacyclin*2 was organized in tandem repeats, and was alternatively spliced, giving rise to five subtypes of transcripts. For both *pcna* and *cyclin* genes, 6-10 introns were found. Quantitative RT-PCR assays showed that *pcna* and *Pmacyclin*2 expression levels were low with small variations during a 28-h time course, whereas *Pmacyclin*1 transcript abundance was 10-100 times higher, and increased markedly during active cell division, suggesting that it is a mitotic *cyclin* and can be a useful growth marker for this species. The gene structure and expression features along with phylogenetic results position this organism between dinoflagellates and apicomplexans, but its definitive affiliation among alveolates requires further studies.

Key words: Alternative splicing; cell cycle; differential expression; intron; *Perkinsus marinus*; spliced leader

Introduction

Perkinsus marinus is a pathogenic protist causing "dermo" disease in oysters in estuaries of the North and Central American Atlantic and Gulf of Mexico coasts. Other *Perkinsus* spp. cause similar, detrimental diseases in a wide variety of other marine molluscs worldwide, all of which yield serious losses for shellfish industries (Villalba et al. 2004). Based on the presence of micropores on the cell surface and an apical complex, *P. marinus* was previously considered to be a member of Apicomplexa (review by Perkins 1996), an exclusively parasitic lineage responsible for malaria and other infectious diseases in humans and animals. However, *P. marinus* was dinoflagellate-like in some of its cytological features such as flagellar spurs and closed mitosis (Perkins 1996). Furthermore, many recent phylogenetic studies showed that *P. marinus* is closer to dinoflagellates than to apicomplexans (e.g. review by Perkins 1996; Saldarriaga et al. 2003; Siddall et al. 1997). There is evidence that *P. marinus* contains a heterotrophically functional, vestigial plastid (Stelter et al. 2007). The

[1]Corresponding author; fax +18604059153
e-mail huan.zhang@uconn.edu (H. Zhang).

doi:10.1016/j.protis.2010.03.003

exact phylogenetic position of the genus *Perkinsus* still remains elusive (Perkins 1996). Although mitosis, cytokinesis, and the life cycle of *P. marinus* have been observed, an unambiguous picture of the mechanisms and dynamics of these processes remains obscure. In particular, there is no information on the structure and expression of cell cycle genes in this organism.

To address the inadequacy of molecular genetic information on *Perkinsus* spp., we chose to analyze genes coding for proliferating cell nuclear antigen (*pcna*) and cyclin (*cyclin*), two ubiquitous genes in eukaryotes that are critical in regulating the cell division cycle. A typical eukaryotic cell cycle consists of four discrete phases, S (DNA replication), M (mitosis), G1 (gap between M and S), and G2 (gap between S and M). At the G1/S transition, cyclin A and cdc2 kinase play a critical role; whereas at the G2/M transition, cyclin B and cdc2 determine whether the cell will progress to mitosis. In the S phase, PCNA (protein encoded by *pcna*) functions to enhance binding of DNA polymerase to the DNA template (hence called a "clamping" molecule), and enhances the efficiency of DNA synthesis (hence also called a processivity factor) (e. g. Burgers 1991; Kannouche et al. 2004; Kisielewska et al. 2005; Nigg 1995; Prelich et al. 1987; Sasaki et al. 1994). In this study, we analyzed the characteristics of genomic structure and post-transcriptional processing of *pcna* and *cyclin*. We isolated 10 slightly different copies of *pcna* from cells of a *P. marinus* in vitro culture. We also identified two distinct types of *cyclin* genes and isolated their corresponding transcripts (cDNAs). By comparing the genomic sequences with the corresponding cDNA we found that the transcripts of these genes were *trans*-spliced with a 21- or 22-nt spliced leader sequence added to the 5'-end. We also determined the expression level of the genes in samples taken during a 28-hour time period and found that one of the cyclins was highly expressed during active cell division, suggesting that it is likely a mitotic cyclin.

Results

Culture of *P. marinus*: Experimental *P. marinus* cultures contained cells in diverse stages of its palintomic proliferative cycle (Fig. 1). Mature trophozoites (A) and subdividing schizonts (B) were generally larger than immature progeny trophozoites (C), and recently subdivided schizonts frequently contained up to 64 or more cytokinetic, developing cells of small progeny trophozoites (D) (Sunila et al. 2001).

Spliced Leaders: Comparison of full-length cDNAs with corresponding genomic sequences revealed a short track of RNA at the 5' end of each transcript, indicative of spliced leader (SL) *trans*-splicing (Zhang et al. 2007). Two types of SL sequences were detected for *pcna* : PerkinsusSL1, 5'-ACCGTAGCCATCTTGGCTCAAG-3' (22 nt) and PerkinsusSL2, 5'-ACCGTAGCCATCTGGCTCAAG-3' (21 nt). The SL was 18 bp (found in 7 cDNAs) or 25 bp (in 1 cDNA) upstream of the predicted start codon (ATG), indicating a very short 5' untranslated region (UTR) in each case. Comparison between cDNA and genomic sequences showed that this SL was *trans*-spliced to the pre-mRNA of *pcna* immediately after an AG dinucleotide, the canonical splice receptor site (Fig. 2). Similarly, for the two types of *cyclin* genes obtained, SL was also *trans*-spliced to pre-mRNAs after the dinucleotide AG (Fig. 3). For *Pmacyclin*1, SL was detected at 24-bp upstream of the start codon ATG, again giving rise to a very short 5' UTR. For *Pmacyclin*2, SL was spliced at 136-387 bp before the start codon, indicating longer 5'-UTR than *pmacyclin*1.

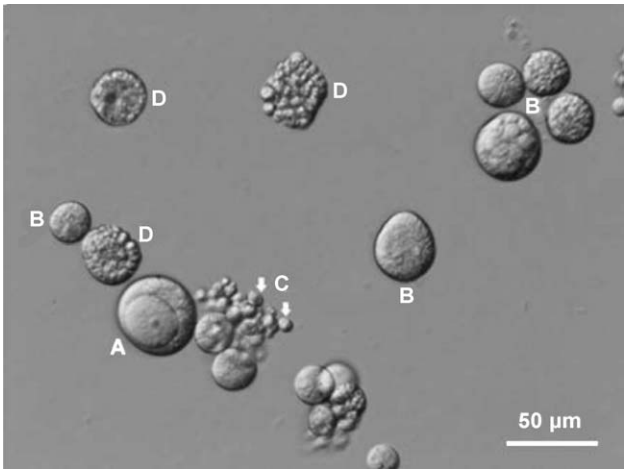

Figure 1. Cells in different stages of the *in vitro* cell cycle of the *Perkinsus marinus* isolate ATCC 50439 culture used in this study. **A**, large, mature trophozoite with a characteristic large, eccentric vacuole; **B**, internally subdividing schizonts; **C**, small progeny trophozoites (arrows) that will enlarge before dividing by schizogony; **D**, clusters of small, sibling progeny trophozoites.

Genomic Structure of *pcna*: Six slightly different genomic copies of *pcna* (gDNA) were retrieved, five of which were 98-99% identical in nucleotide sequence (nt) and 100% identical in the deduced 256 amino acid residues (aa), with the sixth (gDNA6) ~88% nt but 99.6% aa identical to the other 5 copies. Also six copies of cDNAs with 97-99% nt identity were found; the two most abundant copies (5 clones of cDNA1 and 11 clones of cDNA2) matched genomic clones gDNA1 and gDNA2 while the other less abundant four (cDNA3-6; one clone each) did not match genomic clones, revealing that there are at least 10 copies of *pcna* in the *P. marinus* genome.

BLAST search of the sequences against the recently released *P. marinus* genome sequence (Project ID: 12736) yielded two hits (AAXJ01000005 and AAXJ01000721) corresponding to the *pcna* gDNA1 and gDNA6 obtained in this study. No potential TATA box or polyA addition signal (AATAAA) was found at the expected positions in *pcna*. Comparison of *pcna* genomic sequences with corresponding cDNA sequences revealed that all the genomic clones had similar gene structure, containing 8 exons and 7 introns (Table 1, Fig. 2). The introns, 43-65 bp long, carried the canonical GT/AG boundaries.