Characterization of Short Interspersed Elements (SINEs) in a Red Alga, *Porphyra yezoensis*

Wenbo Zhang, Xiaofei Lin, Suresh Peddigari, Katsuaki Takechi, Hiroyoshi Takano, and Susumu Takio

1Graduate School of Science and Technology, Kumamoto University, Kurokami, Kumamoto 860-8555, Japan
2Center for Marine Environment Studies, Kumamoto University, Kurokami, Kumamoto 860-8555, Japan

Key words: *Porphyra*; red algae; repetitive element; retrotransposon; short interspersed elements (SINEs)

Short interspersed element (SINE)-like sequences referred to as *PySN1* and *PySN2* were identified in a red alga, *Porphyra yezoensis*. Both elements contained an internal promoter with motifs (A box and B box) recognized by RNA polymerase III, and target site duplications at both ends. Genomic Southern blot analysis revealed that both elements were widely and abundantly distributed on the genome. Analysis revealed that both elements were widely and abundantly distributed on the genome. 3

7S RNA, 5S rRNA, and 5S rRNA.

At present, SINEs are known in many higher eukaryotes, including plants, vertebrates, and invertebrates. 2) The first plant SINE was identified in introns of the *Waxy* gene in *Oryza sativa*. 3) Since then, several plant SINE families have been identified. 4-8) Most have a common property of tRNA-derived SINEs, an internal promoter with motifs (A and B box) recognized by RNA polymerase III at the 5' end. Up to now, only one report on SINE elements among non-vascular plants, in the green alga *Chlorella*, has appeared. 9) Unlike the usual SINEs, the *Chlorella* SINEs lacked the A–B box. This finding raises a question about the distribution of SINEs among non-vascular plants.

The small unicellular red alga *Cyanidioschyzon merolae* is considered to belong to one of the most deeply branched taxa in the plant kingdom. Recently, complete nucleotide sequences of the nuclear genome of *C. merolae* were reported. 10) We have indicated that the unicellular red alga *Cyanidioschyzon* lacks reverse transcriptase of LTR retrotransposons, while the red macro alga *Porphyra yezoensis* has reverse transcriptase sequences related to those of *copia*- and *gypsy*-type retrotransposons. 11) Hence, to identify the evolution of retrotransposons in red algae, we have tried to isolate non-LTR retrotransposons from *P. yezoensis*; through these experiments, we found SINE-like elements. Here we report the structural properties and expression of these SINE elements.

In our preliminary experiments to isolate LINE-like elements from *P. yezoensis*, a 760-hp genomic PCR product (fragment 1) was isolated. Although it was an unsuccessful PCR product because no primer sequences were identified in both ends, two SINE-like elements (*PySN1* and *PySN2*) were identified in fragment 1 (Fig. 1A). The nucleotide sequences have been deposited with DDBJ database under accession nos. AB266617 (*PySN1*) and AB266618 (*PySN2*). The two SINE elements contained the sequence of target site duplication (TSD) at both the 5' and the 3' end of the elements, and that of the A and B box of the promoter for RNA polymerase III (Fig. 1B). The Genetyx search tRNA program suggested that both elements possess a degenerated cloverleaf tRNA secondary structure (Fig. 1C). Of a total of 21 nucleotide pairs within a 4-stem region, 9 and 5 pairs were matched in *PySN2* and *PySN1* respectively. When the same program was

1 To whom correspondence should be addressed. Tel: +81-96-342-3443; Fax: +81-96-342-3431; E-mail: stakio@gpo.kumamoto-u.ac.jp
applied to other plant SINEs, the nucleotide number of matching pairs in the stem region ranged from 7 to 9 in rice p-SINE1-1, Brassica S1-na2, and tobacco TS-N2, but did not show a cloverleaf structure in PySN1 and PySN2, although PySN1 is more degenerated than PySN2.

To determine the abundance of PySN1 and PySN2 in the genome, Southern blot analysis was performed (Fig. 2). Total DNA was extracted from young blades with hexadecyltrimethyl ammonium bromide and chloroform. Total DNA (1 μg) was digested with restriction enzymes, and the digests were transferred to Biodyne B membranes (Poll, Boston, MA). Hybridization was performed with digoxigenin-labeled DNA probes (Roche Diagnostics, Mannheim, Germany), and the hybridized products were detected with a non-radioactive digoxigenin luminescent detection system (Roche Diagnostics), as described previously. When total DNA was digested with SacI or SpfI which had no restriction site in fragment 1, smear bands of a wide range of fragment sizes were observed in both PySN1 and PySN2 probes (Fig. 2B). These findings indicate that PySN1 and PySN2 are abundantly and randomly distributed on the genome. As shown in Fig. 1A, PySN1 and PySN2 are co-located within fragment 1. Hence, to determine whether the two elements exist as partners on the whole genome, total DNA was digested with Mval, which has one restriction site within each element and one site between two elements (Fig. 2A). A single band was generated by each probe. The sizes of the bands (230-bp by the PySN1 and 440-bp by the PySN2 probe) were identical to the predicted size of the digestion of fragment 1 with Mval (Fig. 2B). This finding suggests that two elements might exist as partners on the whole genome. The result with Mval digestion also suggests that fragment 1 exists as part of the repetitive unit and is abundantly distributed on the whole genome. The copy numbers of PySN1 and PySN2 were estimated by comparing the signal density after Southern hybridization of genome DNA and plasmid DNA harboring PySN1 or PySN2 (Fig. 2C). The copy number of each SINE in the genome was calculated based on the haploid genome size of P. yezoensis (260 Mb). The estimated copy numbers of PySN1 and PySN2 were 4,600 and 8,900 respectively.

PySN1- and PySN2-related transcripts were analyzed by 5’RACE (Fig. 3) and 3’RACE (Fig. 4). Total RNA was isolated from leafy gametophytes with a Sepasol RNA I Super kit (Nacalai Tesque, Kyoto, Japan). The cDNA was synthesized using 1 μg of total RNA with a SMART™ RACE cDNA Amplification kit (Clontech, Mountain View, CA). 5’RACE was performed using specific primers (primers 1 and 2, Fig. 3A), and RACE products were used for Southern blot analysis (Fig. 3B). PySN1 products were detected with PI probes, but
PySN2 products were barely detected with the P2 probe. These findings suggest that the two elements are independently expressed and that PySN1 was more abundantly expressed than PySN2.

To identify the mode of transcription of PySN1, 5'-RACE products ranging from 100 bp to 300 bp were cloned. All of the products (5RACE-1, 5RACE-2, and 5RACE-3) showed high sequence similarity to fragment 1 within the region of PySN1, but did not show any sequence similarity to fragment 1 in the 5'-flanking region of PySN1. To determine whether the same chimera element exists in the genome, genomic PCR was performed with primer 1 and primer 3, which are specific for the 5' end of 5RACE-1, and with the primer 1 and primer 4, which are specific for the 5' end of 5RACE-2 (indicated by arrows in Fig. 3C). The genomic PCR products 5RACE-1G and 5RACE-2G were closely related to 5RACE-1 and 5RACE-2 respectively. This finding indicates the presence of chimera sequences related to 5RACE-1 and 5RACE-2 in the genome. Further, a chimera transcript with PySN1 was found on the EST clone of Porphyra yezoensis (AU292891).

To determine the mode of poly-A tailing of PySN1 and PySN2, 3' RACE was performed with primer 5 and primer 6 (Fig. 4A). Southern blot analysis revealed that the PCR products with primer 5 were detected by the PySN1 probe but not by the PySN2 probe, and that the product with primer 6 was not detected by the PySN2 probe (Fig. 4B). These results suggest that the level of expression of PySN1 is higher than that of PySN2, which is consistent with the results from 5'RACE.

To determine the polyadenylation site of PySN1, PCR products ranging from 100 bp to 300 bp were sequenced (Fig. 4C). All of the PySN1 3'RACE products (8 clones) possessed a poly-A tail at a position near the 3' end of PySN1. Considering these findings and the results shown in Fig. 3, PySN1 appears to be used to provide polyadenylation signals at the 3' end of the chimera transcript with unrelated sequences.

As for the phylogenetic distribution of plant SINEs, no SINEs with a typical structure have yet been identified other than in seed plants. This is the first indication that tRNA-related SINEs are present in a red alga. In the public database of the complete whole genome of the unicellular red alga Cyanidioschyzon merolae, 253 sequences were annotated as interspersed repetitive elements, but all were unrelated to SINEs. Further, we did not find sequences related to PySN1 or PySN2 in the Cyanidioschyzon genome. An investigation of the phylogenetic distribution of sequences related to Porphyra SINEs among red algae and other algae should provide valuable information about the evolution of SINEs in plants.

The first step in the retroposition of SINEs is transcription by Pol III. 5'RACE sequences of Brassica S1 SINEs have suggested that a very small number of SINE elements possess a native structure that appears to be the founder (active) element. In this study, all 5'RACE products using PySN1 primer showed chimera transcripts with unrelated sequences. Therefore, the founder element of the PySN1 family, if present, is completely repressed under normal growth conditions.
Fig. 3. 5' RACE of PySN1 and PySN2.

A, Primers used for 5' RACE. B, PCR products were detected by ethidium-bromide staining (EtBr) and by Southern blot analysis using probe 1 (P1) or probe 2 (P2). C, Fragment 1, 5' RACE products (5RACE-1, 5RACE-2, and 5RACE-3), genomic PCR products (5RACE-1G and 5RACE-2G), and an EST clone of P. yezoensis (AU192891) were aligned. The number of clones with closely related sequences is indicated in parentheses. The PySN1 region (underlining), a chimerical region flanking the 5'-end (dashed line), and two TSDs of the PySN1 element (in white box) are indicated. The locations of boxes A and B are indicated in bold underline, and those for primers are indicated by arrows. Dots indicate identical sequences to fragment 1, and gaps have been introduced to maximize homology. 5RACE-1G was isolated by genomic PCR using primer 1 and primer 3, which are specific for the 5'-end of 5RACE-1. 5RACE-2G was isolated by genomic PCR using primer 1 and primer 4, which are specific for the 5'-end of 5RACE-2. To compare the nucleotide sequences of the chimera region of 5' RACE products and genomic PCR products, the nucleotides identical to 5RACE-1 and 5RACE-2 are indicated by black boxes and gray boxes respectively.

Fig. 4. 3' RACE of PySN1 and PySN2.

A, Primers used for 3' RACE. B, RACE products were detected by ethidium-bromide staining (EtBr) and by Southern blot analysis using probe 1 (P1) or probe 2 (P2). C, Fragment 1 and 3' RACE products (3RACE-1, 3RACE-2, and 3RACE-3) were aligned. The number of clones with identical sequences is indicated in parentheses.
As shown in Fig. 4, most 3’RACE products generated by the PySNI probe possessed a poly-A tail near the 3’ end of PySNI; however, the nucleotide position bound with a poly-A tail corresponded to a site of about 21 bases upstream of the 3’ end of PySNI. This finding indicates that the transcripts isolated by 3’RACE are not used for proliferation for deletion of the 3’ end. We speculate that PySNI families are used as the source for the functional polyadenylation signal of their chimera transcripts.

SINEs were considered to be genomic parasites having no benefit to the organism. Recently, however, it has been recognized that they are used for the benefit of the host. For instance, in the case of animal SINES, polyadenylation has been reported to be used for the generation of functional proteins such as the \( \gamma \) subunit of muscle phosphorylase kinase and leukemia inhibitory factor receptor.\(^{17,18}\) In this study, no sequences related to known proteins were identified in the 5’ flanking region of the chimera transcripts. To determine whether \textit{Porphyra} SINEs are used to generate functional proteins, further characterization of chimera transcripts is required.

Acknowledgments

The authors would like to thank Dr. Naotune Saga of Hokkaido University, Japan, for providing the gametozygotic \textit{P. yezoensis}. This work was supported by Grants-in-Aid for Scientific Research on Priority Areas (to H.T.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and from the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN, to H.T.).

References