NISR RESEARCH GRANT
2002 Research Grant

Structural and Functional Analyses of Novel Transposons of a Koji-mold, *Aspergillus oryzae*, and Their Application for Efficient Insertional Mutagenesis

Katsuya GOMI

Graduate School of Agricultural Science, Tohoku University

Aspergillus oryzae is an important filamentous fungus in the Japanese fermentation industry, such as sake, soy sauce, and miso manufacturing, as well as commercial enzyme production. Since no sexual life cycle has been found in *A. oryzae*, it is difficult to identify the genes involved in industrially important pathways by classical genetic analysis. Thus, there is a need to develop general methods to identify and clone genes from *A. oryzae* and related fungi. For this purpose, isolation of transposable elements from filamentous fungi could give a useful tool for gene tagging. The aim of the research is to isolate transposable elements with potential for gene tagging from *A. oryzae*.

Isolation and characterization of DNA transposons from *A. oryzae*

We found an EST clone highly homologous to *A. niger* transposable element, *Ant1*, in the EST database. A genomic library of *A. oryzae* RIB40 was screened using this EST clone as a probe. Nucleotide sequencing revealed that the gene contained an ORF of 981-bp encoding 326 amino acid residues. It showed extensive sequence homology (70% amino acid identity) to that of *Ant1*. Additionally, the encoded protein included the D35E motif common to the transposases of the I*S630-Tc1* superfamily. As the ORF was found in EST database, the gene most likely encodes a functional transposase and thus we designated it *tnpA* (transposase of *A. oryzae*). In general, transposase gene is located within the IS (insertion sequence) element that has TIRs at the both ends. However, the *tnpA* gene was not surrounded with typical TIRs, but such an IS element was found about 1-kb downstream of the *tnpA*. IS element is approximately 1.9-kb in length, and had 39-bp imperfect TIRs, which showed 72–82% nucleotide homology to that of *Ant1*. In addition, there was no ORF within the 1.9-kb IS element. To our knowledge, this is the first example among fungal DNA transposons that the IS element larger than 1-kb contains no ORF and transposase gene is located out of the element. Although it remains to be elucidated whether the element could transpose, we designated the element as *Aot1* (*Aspergillus oryzae* transposon 1). Southern blot analysis of *A. oryzae* RIB40 showed that *tnpA* gene and *Aot1* exist as three copies in the chromosome, two of which are located immediately downstream of the *amyB* and *amyC* genes. Northern blot hybridization analysis revealed that the *tnpA* gene was expressed in the submerged cultures of rich media, especially in the medium containing wheat bran extract.

In order to test the transposition activity of *Aot1*, we constructed *niaD* disruptants by insertion of the *Aot1* into the intron sequence of the resident *niaD* gene. Conidiospores of the resultant disruptants were inoculated onto a minimal medium and then putative revertants that could grow on the medium were isolated. To examine whether the revertants are generated by transposition of *Aot1* from the *niaD* gene is underway.

Isolation and characterization of retrotransposons from *A. oryzae*

We have also found an EST clone (JZ1636) highly homologous to the gene encoding reverse transcriptase (RT) of the LINE-like DNA element in a fungus, *Ascobolus immersus*. LINEs (long interspersed nuclear elements) and LINE-like elements are the most abundant classes of transposable elements with non-LTR (long terminal repeat). This EST clone seemed to be transcribed from the gene in an *A. oryzae* LINE-like element, and to be involved in its transposition through an RNA intermediate. Screening of the genomic library with the EST clone as a probe resulted in isolation of two positive phage clones, which are derived from different chromosome loci. One phage clone contained whole putative retrotransposon, *Aoret1* (*Aspergillus oryzae* retrotransposon), and the other had *Aoret2* about 6-kb downstream of *Aoret1*. Both putative retrotransposons contained two long ORFs encoded on the strand of the same direction.

The first ORF (ORF1) of *Aoret1* and *Aoret2* encoded a putative protein of 411 and 444 amino acid residues, respectively. These putative proteins contained cysteine-rich motifs near the C-terminal region, which are probably involved in binding DNA. The cysteine-rich motifs are commonly found in *gag*-like proteins encoded in retrotransposons of other organisms. Amino acid homology between the proteins of ORF1s of *Aoret1* and *Aoret2* was 42%. The second ORF (ORF2) of *Aoret1* encoded a protein of 1301 amino acids, and ORF2 of *Aoret2* was 1255 amino acids in length, both of which showed homology to the RT domains of various retrotransposons. Eight highly conserved regions are found in the RT domain, including a YXDD motif that

is considered to be a significant sequence for RT activity. The YXDD motif is FIDD and FVDD in the *Aoret1* ORF2 and the *Aoret2* ORF2, respectively. There was no recognizable protease motif, but a region homologous to endonuclease was found at N-terminus in the both proteins. In addition, a motif homologous to RNase H was also found at the C-terminal region in the protein encoded by ORF2 of *Aoret1*. The proteins of both ORF2s showed 37% identity each other.

Distribution of the retrotransposons in the *A. oryzae* strains and its related fungi

In order to determine the copy number of the retrotransposons in various strains of *A. oryzae* and its closely related fungi, Southern blot analysis was done with *Aoret1* or *Aoret2* as a probe. While most *A. oryzae* strains including RIB40 have only single or two copies but some strains such as RIB143, RIB326, RIB647, and RIB67 have multiple (approximately 10) copies of *Aoret1*, all strains examined have at least 10 copies of *Aoret2*. Interestingly, *A. flavus* that is a putative ancestor of *A. oryzae* have no detectable signals for *Aoret1* and *Aoret2*, and all *A. sojae* strains examined have multiple copies of both retrotransposons. These observations suggest that those retroelements might have transposed between *A. oryzae* and *A. sojae* in the koji-making processes, where both species were possibly cultured, in ancient era.

Expression analysis of the retrotransposons in *A. oryzae*

Northern blot analysis using a respective ORF2 of both retrotransposons as a probe revealed that major transcripts of approximately 5.5-kb in length were transcribed from both retrotransposons. The transcripts were somewhat larger than the expected size for ORF2, but coincided with the full-length transcripts used as a LINEs RNA transposition intermediate. This was proved by the observation that ORF1 also hybridized to the same size of the transcript. Thus if two proteins, *gag*-like protein and RT protein, would be produced, they may be translated from a single full-length transcript of the both retrotransposons. In addition, both retrotranspsons are transcribed extensively in YPD medium. To compare the expression level of the retrotransposon among the strains with low and high copy number of *Aoret1*, Northern blot analysis was done using the strains, RIB40 (2 copies) and RIB67 (approx. 10 copies). No difference in expression level of *Aoret1* was observed, suggesting that transcriptional gene silencing may occur or most of *Aoret1* may be defective in RIB67. Since various transposons have been activated under stress conditions, *A. oryzae* RIB40 was cultured at 42°C or in the presence of $CuSO_4$. Whereas heat shock treatment did not affect the expression level of *Aoret1*, treatment with $CuSO_4$ led to a significant increase in the *Aoret1* transcript. It should be investigated whether the transposition of *Aoret1* could be enhanced under such a stress condition.

Addendum – The *Aspergillus oryzae* genome sequencing project revealed the existence of another LINE-like retrotrans*poson, Aoret3*, and two novel LTR retrotransposons, A*oLTR1* and *AoLTR2* in the genome of *A. oryzae*.