Structural and Functional Characterization of the Intergenic Spacer Region of the rDNA in *Daucus carota*

Akihiro Suzuki\(^1,2\), Shigeyuki Tanifuji\(^1\), Yoshibumi Komeda\(^1\) and Atsushi Kato\(^1\)

\(^1\) Division of Biological Sciences, Graduate School of Science, Hokkaido University, Sapporo, 060 Japan

The intergenic spacer (IGS) region of rDNA in *Daucus carota* contains at least eight kinds of repeated sequence. One sequence may act as a genuine initiation site for a stable transcript and another may act as a spacer promoter that may be an entry site for proteins required for transcription.

**Key words:** *Daucus carota* — IGS — Promoter — rRNA gene — *trans* factor — Transcription.

In higher eukaryotes, genes for ribosomal RNA (rRNA) are organized in tandem arrays at the nucleolar organizing regions of chromosomes. Each repeating unit usually contains one coding sequence for each of the 18S, 5.8S and 25–28S rRNAs, and three spacer sequences. The intergenic spacer (IGS) region from the 3′ end of the gene for 25–28S rRNA to the 5′ end of the gene for 18S rRNA consists of untranscribed spacer and transcribed spacer regions. In many organisms, various types of repeating sequence have been found in tandem arrays in the IGS region, and variations in the numbers of such repeating sequences usually cause the length polymorphisms of rDNA repeating units (Appels and Honeycutt 1986). Many studies on the transcription of rRNA in animals have demonstrated that the IGS region contains promoter sequences for transcription of rRNA and that some of the repeating sequences found in IGS region have enhancer functions for the transcription of the rRNA (Sollner-Webb and Tower 1986). In *Drosophila* and *Xenopus*, some of the repeats in the IGS region were shown to have an enhancer function for the transcription of the precursor rRNA. They appear to act as entry sites for transcription factors (Coens and Dover 1982, Reeder 1983). The protein factors required for transcription have been analyzed in some species (Jantzen et al. 1990, McStay et al. 1991, O’Mahony and Rothblum 1991).

In plants, studies on the mechanisms of transcription of genes for rRNA have been initiated. The structure of the IGS and the consensus sequence surrounding the initiation site have been reported in several plants (Delcassos-Tremousaygue et al. 1988, Gerster et al. 1988, Vincentz and Flavell 1989, Kato et al. 1990, Perry and Palukaitis 1990, Polanco and Vega 1994, Da Rocha and Bertrand 1995). In *Arabidopsis*, a spacer promoter, reflecting the duplication of part of the promoter sequences, was reported (Gruendler et al. 1991). Furthermore, it was reported that the upstream repeated sequence containing a spacer promoter increased the rate of transcription from the promoter of a *Xenopus* ribosomal gene in injected frog oocytes, even though the enhancer activity was very limited in transfected protoplasts of *Arabidopsis* (Doelling et al. 1993). Protein fractions that interact with a certain sequence in IGS regions in a sequence-specific manner have been described for several plants (Echeverria et al. 1992, 1994, Jackson and Flavell 1992, Nakajima et al. 1992, Schmitz et al. 1989, Zentgraf and Hemleben 1992, Suzuki et al. 1995).

In *Arabidopsis*, it was reported that bending of DNA occurred as a result of binding of a specific protein and that the protein was related to the high mobility group (HMG) proteins (Christine et al. 1995).

Here, we present the complete nucleotide sequence of the IGS region of the rDNA from *Daucus carota*. Figure 1A shows the complete nucleotide sequence of 5,775 bp of the IGS region of rDNA cloned from *Daucus carota*. The borders between the spacer and the coding regions were inferred from sequences reported for other plants. From an examination of a dot matrix self-comparison of the IGS region, we found several repeated sequences. A diagrammatic representation of the organization of the IGS region is shown in Figure 1B. At least eight kinds of repeated sequence, designated A, B, C, D, E, F, G and H, respectively, were found in this region. In each repeated sequence, the extent of the sequence similarity ranged from 68% to 98% (Fig. 1C). The repeat designated G consisted of a tandem array of three complete copies and one truncated copy of a 455 bp sequence. The existence of this repeat sequence was reported previously (Taira et al. 1988). The difference in the number of iterations of this repeat sequence causes the length heterogeneity of the rDNA repeating unit of *Daucus carota* (Taira et al. 1988).

The last repeat sequence, repeat H, consisted of two sequences of 228 bp, which were separated by a 1,909 bp se-
Fig. 1A
Fig. 1B

Fig. 1 (A) Nucleotide sequence of the IGS region of *Daucus* rDNA. The plasmid DER1 used in this study was described previously (Taira et al. 1988). Eight families of repeated sequences are indicated by underlining or boxes (A, B, C, D, E, F, G and H). Boldface A residues at positions 990 and 3,127 are putative sites of initiation of transcription. (B) Diagrammatic representation of the IGS region of *Daucus carota* rDNA. The copies of each repeated sequence are boxed and marked with alphabetical letters followed by numerals. sp, spacer promoter; gp, gene promoter. (C) Alignment of members of each family of repeated sequences shown in Figure 1A and sequence homology. Hyphens indicate the same nucleotides, and plus signs indicate gaps introduced to maximize homology. Two members of the families of repeats A, E and H were directly compared with each other in order to calculate sequence homology. In other cases, consensus sequences (C.S.) were determined and the sequence homology was calculated for each member by comparing it with the consensus sequence. X represents an undetermined nucleotide in a consensus sequence.

sequence that contained repeats D, E and F. Their sequences showed a high degree of conservation (98%). Furthermore, each unit of repeat H contained the TATATAGGG motif which is a sequence that is conserved around the site of initiation of transcription of plant rRNAs (Delcasso-Tremousaygue et al. 1988, Gerster et al. 1988, Kato et al. 1990). The sequence of repeat H corresponded to positions — 217 to +11 relative to the presumed site of initiation of transcription. Thus, the IGS region of the gene for rRNA of *Daucus carota* contains multiple putative promoter sequences, as does that of *Arabidopsis* (Gruendler et al. 1991), *Oryza* (Cordesse et al. 1993), *Xenopus* (Reeder 1983) and *Drosophila* (Coens and Dover 1982). In each of these rDNAs, it was reported that one of the reported promoter sequences is the true promoter while the others may act as enhancer sequences. The sequence H1 was designated a spacer promoter and the sequence H2 was designated a gene promoter, eventhough the true promoter sequence is unknown.

In order to determine the 5' site of the initiation of transcription of rRNA in *Daucus carota*, we carried out S1 nuclease protection analyses using both a fragment that contained a gene promoter (positions 3,001 to 3,289) and a fragment with the spacer promoter (positions 864 to 1,099) as probes. When the fragment containing the gene promoter was used as the probe, several bands were detected. Two major bands corresponded to the A residues at positions 990 and 986. One of them, that corresponding to the A residue at position 990, also corresponded to the A residue at position 3,127 of the gene promoter, as deduced from a comparison of the surrounding sequences. However, these bands were weak and disappeared when the concentration of S1 nuclease was increased. These results indicate that the gene promoter acts as the dominant cis element for the transcription, together with the possibility that the spacer promoter also acts as such an element in vivo. Similar results, with two such promoters being active in vivo, were reported for the transcription of the rDNA of *Arabidopsis thaliana* (Doelling et al. 1993).

In the case of *Arabidopsis*, the transcripts from both promoters were clearly detected. it is unclear why we failed to detect a clear transcript from the spacer promoter but the following explanation is suggested. The nucleotide sequences of the gene promoter and spacer promoter were highly conserved and this conservation extended from position —217 to position +11. The gene promoter had a long AT-rich sequence in its upstream region, while the AT-rich sequence located in the upstream region of the spacer promoter was short. AT-rich sequences upstream of the rDNA promoter may be important for the transcription of rRNA in plants (Echeverria et al. 1992). Thus, the difference in length of AT-rich sequences might be related to the transcriptional activity of rRNA in *Daucus carota*.

In order to study DNA-protein interactions, we performed gel-retardation assays using fragments derived from the gene promoter and spacer promoter regions. A crude nuclear extract was partially purified by chromatogra-
phy on a column of heparin-Sepharose. After absorption of the nuclear extract to the column, protein fractions were obtained by stepwise elution with a buffer that contained 100, 200, 300, 400 and 500 mM KCl. The protein fraction that was eluted in the presence of 300 mM KCl yielded two retarded bands (Fig. 3). Furthermore, the bands obtained when the probe DNA that contained the gene promoter was used were similar to the bands obtained when the probe DNA that contained the spacer promoter was used.

To confirm the sequence specificity of protein binding, competition analyses were performed. When both the gene promoter and spacer promoter fragments were used as probes, the intensities of the two retarded bands were strongly reduced by the addition of the unlabeled gene promoter or
Intergenic spacer of the rDNA in *Daucus carota*

Fig. 2 Identification of the site of initiation of transcription from the gene promoter (A) and the spacer promoter (B) by S1 nuclease mapping. The gene promoter probe was the HindIII-EcoRI fragment (HindIII: -126, EcoRI: +163), labeled at the 5' end of the coding strand. The spacer promoter probe was the HindIII-EcoRI fragment (HindIII: -126, EcoRI: +110), labeled at the 5' end of the coding strand. Numbering is relative to each putative initiation site, which is designated +1. Probe DNA was allowed to hybridize with 75 µg of carrot RNA (lanes 6, 7, 8, and 9) or yeast tRNA (lane 5), and then it was treated with 10 units (lane 6), 40 units (lanes 5 and 7), 160 units (lane 8) or 640 units (lane 9) of S1 nuclease at 37°C for 30 min. After denaturation, protected fragments were subjected to electrophoresis with DNA sequencing markers (lanes 1 and 10, G; lanes 2 and 11, G+ A; lanes 3 and 12, C+ T; lanes 4 and 13, C). The arrowhead in (A) indicates the endpoint of the clear protected band and the arrowheads in (B) indicate the endpoints of the faint protected bands.

spacer promoter fragment as a competitor (lanes 3-6 in Fig. 3). By contrast, these retarded bands were unaffected by the addition of a 100-fold excess of another DNA frag-

Fig. 3 Gel-retardation assay using a gene promoter and spacer promoter as probes with various competitors. The HindIII-HindIII (−285 to −126) fragment, designated the gene promoter (gp), was used as a probe in the left panel and the RsaI-HindIII (−290 to −126) fragment, designated the spacer promoter (sp), was used as a probe in the right panel. Numbering is relative to each putative initiation site, designated +1. Lane 1, no nuclear extract; lane 2, no competitor; lanes 3 and 4, the same fragment as the gene promoter probe was used as the competitor; lanes 5 and 6, the same fragment as the spacer promoter probe was used as the competitor; lane 7, the polylinker segment of pBl31 (31) was used as the competitor. Each competitor was present at 10 times (lanes 3 and 5) or 100 times (lanes 4, 6 and 7) the level of the probe.

ment as a competitor (lane 7 in Fig. 3). These results indicated that the DNA-protein(s) interactions were specific to the DNA sequences that contained the gene promoter and spacer promoter. Furthermore, it appeared that the same protein fraction interacted with both the gene promoter and the spacer promoter sequences. The function of the protein fraction is unknown. However, the proteins in this fraction might be components of the transcriptional apparatus. Thus, there is a possibility that the spacer promoter acts as the site of entry for the protein factor(s) required for formation of the initiation complex for transcription of rRNA, even though the transcriptional activity from the spacer promoter seemed very low. A similar hypothesis has been proposed for transcription of rRNA in maize (Schmitz et al. 1989). These hypotheses remain to be confirmed but they should not be ignored when enhancer activity during transcription of plant rRNA is considered.

This study was supported by grants from the Ministry of Education, Science and Culture of Japan.

References


Cordesse, F., Cooke, R., Tremousaygue, D., Grellet, F. and Delseny, M.

Coens, E.S. and Dover, G.A. (1982) Multiple Pol I initiation sequences of


238


Echeverria, M., Delcasso-Tremousaygue, D. and Delseny, M. (1992) A nu-


Delcasso-Tremousaygue, D., Grellet, F., Panabieres, F., Ananiev, E.D.


The structure of the large spacer region of the rDNA in Vicia faba and


Isolation of RNA using guanidium salts. Methods Enzymol. 152: 219-


Ribi are both required for formation of a stable polymerase I promoter

complex in X. laevis. EMBO J. 10: 2297-2303.


tion of nucleotide sequences that interact with a nuclear protein fraction


O'Mahony, D.J. and Rothblum, L.I. (1991) Identification of two forms of


USA 88: 3180-3184.

Perry, K.L. and Palukaitis, P. (1990) Transcription of tomato ribosomal


221: 102-112.

Pikaard, C.S., McStay, B., Schultz, M.C., Bell, S.P. and Reeder, R.H.

(1989) The Xenopus ribosomal gene enhancers bind an essential poly-

merase I transcription factor xUBF. Genes Dev. 3: 1779-1788.

Polanco, C. and Vega, P.D.L. (1994) The structure of the rDNA intergenic

spacer of Avena sativa L.: a comparative study. Plant Mol. Biol. 25: 751-

756.


351.


binding of nuclear proteins to the promoter region of a maize nuclear


Smith, S.D., Oriahi, E., Lowe, D., Yang-Yen, H.-F., O'Mahony, D.,

Rose, K., Chen, K. and Rothblum, L.I. (1990) Characterization of fac-

tors that direct transcription of rat ribosomal DNA. Mol. Cell. Biol. 10:

3105-3116.


fractions that bind to a narrow region of the promoter of the ribosomal


size classes of carrot rDNA repeating units is due to reiteration of se-

quence of about 460 bp in the large spacer. Mol. Gen. Genet. 213: 170-

174.

Vinceit, M. and Flavell, R.B. (1989) Mapping of ribosomal RNA tran-


Zentgraf, U. and Hemleben, V. (1992) Complex formation of nuclear pro-

teins with the RNA polymerase I promoter and repeated elements in the

external transcribed spacer of Cucumis sativus ribosomal DNA. Nucl.

Acids Res. 20: 3685-3691.

(Received August 22, 1995; Accepted January 10, 1995)