Title	Study of transcriptional regulation of rRNA genes in higher palnts
Author(s)	Suzuki, Akihiro; 鈴木, 章弘
Degree Grantor	北海道大学
Degree Name	博士(理学)
Dissertation Number	甲第3722号
Issue Date	1995-12-25
DOI	https://doi.org/10.11501/3108189
Doc URL	https://hdl.handle.net/2115/51292
Type	doctoral thesis
File Information	000000292929.pdf

Study of Transcriptional Regulation of rRNA Genes in Higher Plants

1995

Akihiro SUZUKI

①

Study of transcriptional regulation of rRNA genes in higher plants

Akihiro SUZUKI

1995

Contents

page

Abstract 5

Introduction 7

Materials and Methods 14

Plasmid constructions of *Daucus carota* rDNA 14

Plasmid constructions of *Vicia faba* rDNA 14

Construction of DNAs with point-mutation 15

DNA sequencing 16

RNA isolation and S1 nuclease protection analysis 16

Preparation of templates for *in vitro* transcription 17

Preparation of S-100 fraction for *in vitro* transcription 17

In vitro transcription assay 18

Preparation of probe DNA for DNA binding assay 19

Preparation of probe DNA for southwestern screening 20

Preparation of the nuclear fraction 20

Fractionation of crude nuclear extracts 21

Gel-retardation assay 22

DNase I footprinting and methylation interference analysis 23

Analysis of binding stability 23

Screening of λgt11 cDNA library of *Vicia faba* 23

Analysis of β-galactosidase recombinant proteins by western transfers 24

Results 26

Structure of the IGS region of *Daucus carota* rDNA 26
Analysis of transcription initiation site in *Daucus carota* rDNA 27
In vitro transcription of *Vicia faba* rDNA 29
Gel-retardation assay using *Daucus carota* nuclear extract 31
Competition assay using gene promoter and spacer promoter sequence in *Daucus carota* rDNA 31
Gel-retardation assay using *Vicia faba* nuclear extract 32
Column fractionation of *Vicia faba* nuclear extracts 34
Sequences recognized by the three proteins fractions 36
Methylation interference analysis 37
The affinity of RBP-1, RBP-2 and RBP-3 to promoter sequences with mutation 38
Stability of binding of RBP-1, RBP-2 and RBP-3 to promoter sequence 39
The affinity of nuclear protein and rDNA promoter sequence in heterogeneous system 40
DNA binding proteins during germination 41
Molecular cloning of DNA binding protein that interact with the DNA sequence near the transcription initiation site 41

Discussion 43

Comparison of IGS organization between *Vicia faba* and *Daucus carota* 43

Comparison of sequence around transcription initiation site in plant rDNA 44

In vitro transcription system of rRNA gene in plant 45

Cis elements required for accurate initiation of transcription of rRNA genes in plant 47

Functions of gene promoter and spacer promoter in *Daucus carota* rDNA 49

Protein binding sequences in rDNA promoter 52

Relationships among RBP-1, RBP-2 and RBP-3 55

Molecular cloning of a DNA binding protein that interact with the DNA sequence near the transcription initiation site 59

Acknowledgments 61

References 62

Figure legends 83

Figure 98

Abstract

I have been studying the mechanisms of transcription of rRNA in higher plants. Here, I present the complete nucleotide sequence of IGS region of *Daucus carota* rDNA. The repeating unit of *Daucus carota* rDNA has two major size classes and this polymorphism was caused by reiteration of a 460 bp sequence. Further analysis in this paper shows the possibility that *Daucus carota* rDNA has plural promoter sequences. The results of S1 analysis and gel-retardation analysis show the possibility that one of the sequences may act as a spacer promoter, like *Xenopus*, *Drosophila* and *Arabidopsis* . In order to identify the promoter sequences, I used *Vicia faba* rDNA as a material. From the analysis using *in vitro* transcription system, I showed the possibility that the *cis* element required for transcription of *Vicia faba* rDNA may locate in downstream sequence from -363. This region contains the sequence which binds to the nuclear protein fraction with sequence-specific manner. I expanded this analysis of DNA-protein interaction. I fractionated the *Vicia faba* nuclear extracts using ion-exchange chromatography and affinity chromatography, and showed that the binding protein fraction divided to three kinds of protein fractions. They recognized similar sequences for binding, but their binding stability was different. Furthermore, I was able to isolate a cDNA clone encoded DNA binding protein that interact with the sequence near the transcription start site

of rRNA gene by southwestern technique. From these results, I discussed the general mechanisms of transcription of plant rRNA gene.

Introduction

In higher eukaryotes, ribosomal RNA genes (rDNA) are organized in tandem arrays at the nucleolar organizing regions of chromosomes. Each repeating unit usually contains one coding sequence for each of the 17-18S, 5.8S and 25-28S rRNAs and three spacer sequences between these coding sequences. The spacer region from 3’ end of the 25-28S rDNA to the 5’ end of the 17-18S rDNA consists of the non-transcribed spacer and transcribed spacer regions and it is called the intergenic spacer (IGS) region. The coding region for mature rRNAs are highly conserved among various organisms but the IGS regions are highly variable, in length and primary structure, among organisms and even among individuals. In many animals and plants, various types of repeating sequence were found as tandem arrays in the IGS region, and variation in the number of such repeating sequences usually causes the length heterogeneity of the IGS region, thus resulting in length polymorphism of rDNA repeating units [1].

In several plants, the complete nucleotide sequence of the IGS region was determined for wheat [4], rye [2], maize [55, 94], radish [19] mung bean [27, 81] tomato [65, 83], potato [10], *Arabidopsis thaliana* [30], *Avena sativa* [68], *Brassica rapa* [18], *Eruca sativa* [49], *Pisum sativum* [44] and *Vicia faba* [44]. In plants, the repeat units observed in IGS region, in most cases are tandemly arranged, but may also be interspersed throughout the IGS, as has been shown for cucumber [20]. From these results, the mechanisms of interspecific homogenization and a rapid divergence of rDNA after separation of species was proposed [4, 18, 44, 68, 70].

In animal, many studies have demonstrated that IGS contains promoter sequences of the rRNA transcription and that some of the repeating sequences found in IGS have an enhancer function for the transcription of the rRNA [41, 87, 107]. In human, it was reported that the promoter consists of two major elements: a upstream region, composed of several domains, that lies between nucleotides -234 and -107 relative to the transcription initiation site and affects transcription up to 100-fold and a core element that lies between nucleotides -45 and -20 and affects transcription up to 1000-fold [33, 43]. In mouse, the results were reported that sequences extending out to residue -140 greatly favor the initiation reaction, although the sequence from -39 to +9 contains sufficient information to initiate accurately [58]. In *Drosophila* and *Xenopus*, IGS contains duplications of the gene promoter, called spacer promoter, and they may stimulate transcription from the downstream gene promoter [15, 20, 29, 48].

Furthermore, in animal, some protein factors required to form the transcription initiation complex have been characterized and specific DNA-protein and protein-protein interactions have been analyzed [7, 14, 66, 86]. For vertebrate system, two factors may be required for effective transcription of rRNA. One of them, called UBF (upstream binding factor), binds to promoter sequence with sequence-specific manner. The cDNAs encoding UBF were isolated in human [42], rat [63], mouse [35] and *Xenopus* [3, 56]. These UBF are well conserved and contain different numbers of repeated HMG homology boxes. And the existence of two different copies of the UBF gene differing in the number of HMG boxes was reported. Another factor was called SL1 (selectivity factor 1), which recognized and bound to DNA-UBF complex [7]. This factor consists of TBP (TATA box-binding protein) and TAFs (TBP-associated factors) [16]. And the TAFs including SL1 were identified as three polypeptides whose molecular weights were 110, 63 and 48 kDa in human[16] and 95, 68 and 48 kDa in mouse[23]. Recently, the process of formation of SL1 was analyzed [17, 109]. In addition, some studies have demonstrated that SL1 and part of the *cis* element (SSE) cause a species-specific transcription [32] that is a special characteristic of rRNA gene promoters [6, 7, 77]. Recent study demonstrated that two TAFs (TAFI 63 or TAFI 68 and TAFI 48) mediated the binding to rDNA promoter and the largest TAF (TAFI 110 or TAFI 95) may play a role in the species-specific assembly of productive preinitiation complexes [75, 76]. Except the knowledge about general factors described above, the regulatory mechanisms of rRNA transcription were reported. In NIH3T6 fibroblasts, it was reported that the transcription of UBF gene may be a primary response for activation of rRNA transcription [28]. In U937 cells, Rb (retinoblastoma susceptibility gene) products inhibit the rRNA transcription by the manner which Rb interacts with UBF and inhibits the UBF activity [11].

In contrast to the knowledge of rRNA transcription in animal, little is known about the mechanisms of rRNA transcription in plants. From the analysis of IGS region, the consensus sequence surrounding the initiation site have been reported [19, 27, 44, 65, 95]. Despite the fact that a consensus sequence surrounding the initiation site was proposed as a possible *cis*-control element, the significance of this sequence is still unclear. Recently, the first report about plant rDNA promoter was presented that 98 bp sequences which lies from -92 to +6 was sufficient to represent functional *Arabidopsis* rRNA gene promoter [22]. This analysis was performed using transient expression system, and the function of enhancer sequence located in upstream region was also discussed. In order to analyze the mechanisms of transcription of plant rDNA, the *in vitro* transcription system of rDNA was developed using *Vicia faba* embryonic axes [102]. In this report, although accurate transcription was observed using several exogenous *Vicia faba* rDNA fragments as templates, promoter sequence has not been identified. Recently, another *in vitro* transcription system using tobacco cultured cells, was reported [26]. In their system, when the tobacco rDNA fragment was used as template, accurate transcription was observed, but using *Vicia faba* rDNA fragments as templates, no specific transcripts were detected [89]. This result suggested that rRNA transcription apparatus in plant is also species-specific like animal.

With respect to trans-factors, some protein fractions that interact with a certain sequence in IGS in a sequence-specific manner were reported from several plants [24, 25, 39, 61, 82, 108]. In maize rDNA, several proteins bind to both regions of promoter sequences (from -40 to -10) and upstream repeat sequences. Thus, the enhancer function of upstream repeated sequence was suggested [82]. In wheat rDNA, the binding of specific proteins to the -34 to +10 region was detected. And this region contains two binding sites, one located from -7 to +2 and another located from -34 to -14. These sites contain a consensus sequence, which was CATGG-GC-AAAAC [39]. In *Cucumis* rDNA, three proteins, whose molecular weights were 16, 22 and 24 kDa, interact to the promoter region. This binding region also contains two binding sites like wheat. In this case, it was reported that these proteins seem to interact with the single strand DNAs of these region [108]. In radish rDNA, the binding protein with dA/dT-rich sequences located upstream region of transcription initiation site was reported

[24]. Furthermore, another protein that may relate to processing of precursor rDNA was also reported [25]. In *Vicia faba* rDNA, the protein which binds to promoter region with sequence-specific manner was detected and the detailed analysis of necessary nucleotides for DNA-protein interaction was performed [61].

I have been studying the mechanisms of transcription of rRNA in land plants. Here, I present the complete nucleotide sequence of IGS region of *Daucus carota* rDNA. The repeating unit of *Daucus carota* rDNA has two major size classes [45] and this polymorphism was caused by reiteration of a 460 bp sequence [92]. Further analysis in this report shows the possibility that *Daucus carota* rDNA has plural promoter sequences. The results of S1 analysis and gel-retardation analysis show the possibility that one of the sequences may act as a spacer promoter, like *Xenopus*, *Drosophila* and *Arabidopsis* [90]. In order to identify the promoter sequences, I used *Vicia faba* rDNA as a material. From the analysis using *in vitro* transcription system [102], I showed the possibility that the *cis* element required for transcription of *Vicia faba* rDNA may locate in downstream sequence from -363. This region contains the sequence which binds to the nuclear protein fraction with sequence-specific manner [61]. I expanded this analysis of DNA-protein interaction. I fractionated the *Vicia faba* nuclear extracts using ion-exchange chromatography and affinity chromatography, and showed that the binding protein fraction divided to three kinds of protein fractions. They recognized similar sequences for binding, but their binding stability was different [91]. Furthermore, I was able to isolate a cDNA clone encoded DNA binding protein that interact with the sequence near the transcription start site of rRNA gene by southwestern technique. From these results, I discussed about the general mechanisms of transcription of plant rRNA gene.

Materials and Methods

Plasmid constructions of Daucus carota *rDNA*

The plasmid DER1 was used in this study [92]. It contains a 7.6 Kb *Eco*R I-*Eco*R I fragment encompassing the whole length of the IGS region of *Daucus carota* rDNA in the *Eco*R I site of plasmid pBR325. As shown in Figure 1, *Eco*R I-*Pst* I fragment (for plasmid DEP1) and *Pst* I-*Eco*R I fragment (for plasmid DEP2) were cloned into *Sma* I site of plasmid pIBI31. *Hin*d III-*Hin*d III fragment was cloned into *Hin*d III site of plasmid pIBI31 and the resultant plasmid was named plasmid DHH1. For gel-retardation assay using gene promoter and spacer promoter, five fragments were cloned. As shown in Fig. 11A, gp1 fragment (from -368 to -285), gp2 fragment (from -285 to -126), gp3 fragment (from -126 to +165), sp1 fragment (from -290 to -126) and sp2 fragment (from -126 to +90) were recloned into *Hin*c II site of plasmid pIBI31.

Plasmid constructions of Vicia faba *rDNA*

The DNA fragments used in this study are derived from plasmid VER1 containing a *Eco*R I/*Eco*R I fragment encompassing the whole *Vicia faba* IGS region in the *Eco*R I site of plasmid pBR325 [100]. As shown in Figure 2, *Mlu* I-*Mlu* I fragment (for plasmid pVrMM), *Mlu* I-*Hind* III fragment (for plasmid pVrMH) and *Dra* I-*Hin*d III fragment (for plasmid pVrDH) were cloned into *Sma* I site of plasmid pIBI31. Several fragments excised from the these plasmid with variable restriction enzymes were used in this study (Fig. 13).

Plasmid pVrMH was used to construct of 5' or 3' deletion mutant. This plasmid was linearized with *Pst* I and *Xba* I, digested with exonuclease III and mung bean nuclease and ligated [106]. The extent of deletion was determined by sequencing through the vector junctions and was indicated in the name of the plasmid: for example, 5'Δ-62 would be a promoter mutant deleted of sequences upstream of -62 and 3'Δ-34 would be a promoter mutant deleted of sequences downstream of -34 (Fig. 16).

Construction of DNAs with point-mutation

As shown in Figure 3, several 30 bp nucleotides from -35 to -64, which contain one or a few mutated nucleotides, were synthesized. Another 15 bp nucleotide, designated RP, 5'-GCCACCAAATTGGTC-3', which corresponds to the bottom strand from -44 to -30, was synthesized. Each of the 30 bp nucleotides and 15 bp nucleotides were annealed and a polymerase reaction was done using T7 DNA polymerase and dNTPs. Synthetic 35 bp double stranded DNAs were fractionated in 10% polyacrylamide gel electrophoresis and elution from the gel. The resultant double stranded DNAs were cloned into *Sma* I site of pIBI31. The insert DNA was checked by sequencing through the vector junctions.

DNA sequencing

For determination of the IGS region of *Daucus carota*, plasmid DEP1, DEP2 and DEP3 were used. Plasmids DEP1 and DEP2 were linearized with *Pst* I and *Xba* I, digested with exonuclease III and mung bean nuclease and ligated [106]. Plasmid DHH1 was digested with 4 base cutter restriction enzyme. The resultant fragments were randomly cloned into *Hinc* II site of plasmid pIBI31.

For determination of the nucleotide sequence of the cDNA (λgt11-NO.2) insert, the cDNA insert was isolated by the digestion with *Eco*R I. The resultant DNA fragment was digested with 4 base cutter restriction enzyme. The several fragments were randomly cloned into *Hinc* II site of plasmid pIBI31. The chain termination method was used for sequencing [80].

RNA isolation and S1 nuclease protection analysis

Total cellular RNA was prepared from seedling of *Daucus carota* according to the guanidium-CsCl method [52]. S1 mapping experiments were carried out as described by Berk and Sharp [9]. The *Hin*d III(at 864)-*Eco*T14 I(at 1099) fragment and the *Hin*d III(at 3001)-*Eae* I(at 3289) fragment were labeled at their 5' ends for using as probe DNAs. Sequencing ladders for size markers were produced by the chemical modification method of Maxam and Gilbert [54].

Preparation of templates for in vitro *transcription*

Plasmids pVrMH and several plasmids with 5'-deletion mutation (5'Δ-121, 5'Δ-91, 5'Δ-62, 5'Δ-54, 5'Δ-45 and 5'Δ-41) were used in the *in vitro* transcription study. As shown in Figure 9A and Figure 10A, These plasmids were digested with several restriction enzymes. And then the resultant DNA fragments were used as templates.

Preparation of S-100 fraction for in vitro transcription

The preparation of soluble whole cell extract was performed according to the method of Yamashita *et al.* [102]. Cultured cells of carrot or embryonic axes of *Vicia faba* were pulverized with a mortar and pestle in liquid nitrogen and mixed in buffer A (10 mM HEPES-KOH pH 7.9, 10 mM KCl, 1.5

mM $MgCl_2$ 0.5 mM DTT) with a Dounce type homogenizer. After filtration through four layers of Miracloth, 0.1 volumes of buffer B (300 mM HEPES-KOH pH 7.9, 1400 mM KCl, 30 mM $MgCl_2$) was added to this filtrate. This lysate was centrifuged at 120,000 g for 15 min and then the supernatant was recentrifuged at 120,000 g for 1 hr. The resulting supernatant was carefully collected and dialyzed for 8 hr against two changes of 1000 ml of buffer C (20% glycerol, 20 mM HEPES-KOH pH 7.9, 100 mM KCl, 0.2 mM EDTA, 3 mM DTT). Insoluble materials were removed by centrifugation at 12000 g for 10 min. The final extracted volume was approximately 6-8 ml containing 15-20 mg/ml protein. This supernatant, designated whole cell extract, was quickly frozen in liquid nitrogen in small aliquots and stored at -80°C. Repetition of freeze thaw cycles was avoided before use.

In vitro *transcription assay*

The transcription reaction mixture contained, in a final volume of 30 μl, 12 mM HEPES-KOH pH 7.9, 100 mM KCl, 12 mM $MgCl_2$, 1.5 mM DTT, 10 mM creatin phosphate, 600 μM each of ATP, GTP and CTP, 15 μM UTP, 5 μCi [α-^{32}P]UTP (specific activity 800 Ci/mmol), 100 units human placenta RNase inhibitor, 100 μg/ml α-amanitin, 0.5-1.0 μg template DNA and 15 μl of whole cell extract. The assay mixture was incubated at 30°C for 45 min. The reaction was then terminated by adding 0.1 volumes of 10% sodium lauryl sulfate followed by two extractions with phenol/chloroform (1:1) mixture. After the nucleic acids were precipitated with ethanol, the pellet was dissolved in 5 μl of 50% formamide solution containing 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol. This sample was incubated at 80°C for 3 min and then loaded on a 50% urea-5% polyacrylamide gel. Following electrophoresis, the gel was dried and subjected to autoradiography.

Size markers were prepared from several *in vitro* labeled RNA fragments, using bacteriophage T3 RNA polymerase.

Preparation of probe DNA for DNA binding assay

For preparation of probe DNAs, plasmids pVrMM, pVrMH, pVrDH, 3'Δ-34, 3'Δ-59, 5'Δ-62 and 5'Δ-54 (Fig. 2 and Fig. 11) were digested with several restriction enzyme. An *Nsp* V-*Mlu* I fragment (-94 to -34 and a segment of a pIBI31 polylinker) was excised from the recombinant plasmid 3'Δ-34. In order to prepare nonspecific competitor DNA, plasmid pIBI31 was digested with *Hin*d III and *Eco*R I, and a polylinker segment was isolated.

In the case of *Daucus carota*, as shown in Figure 8A, gp1 fragment (-368 to -285), gp2 fragment (-285 to -126), gp3 fragment (-126 to +165), sp1 fragment (-290 to -126) and sp2 fragment (-126 to +90) were used as probes for gel-retardation assays. The gp2 fragment and sp1 fragment and the polylinker segments of pIBI31 were used as competitor DNAs.

The excised DNA fragments were labeled by a 3'-end-filling with the Klenow fragment of DNA polymerase and the [α-^{32}P]dCTP radioactive precursors.

Preparation of probe DNA for southwestern screening

For preparation of probe DNA, an *Nsp* V -*Mlu* I fragment (-94 to -34 and a segment of a pIBI31 polylinker) was excised from the recombinant plasmid 3'Δ-34. And this fragment was concatenated. The resultant DNA fragment was labeled with [α-P^{32}]dCTP by nick translation and used as probe.

Preparation of the nuclear fraction

Crude nuclear extracts were prepared according to the method described by Nakajima *et al.* [61]. Embryonic axes of *Vicia faba* seeds or seedlings of *Daucus carota* were ground with a motar and pestle in liquid nitrogen and then homogenized in buffer A (20 mM HEPES-KOH pH 7.9, 60 mM

KCl, 5 mM $MgCl_2$, 1 mM dithiothreitol, 0.4 M sucrose, 0.5% Triton X-100) with a Dounce type homogenizer. After filtration through four layers of Miracloth, the filtrate was centrifuged at 3000 X g for 5 min. The precipitate containing nuclei was suspended in buffer-B (the same as buffer-A, but without sucrose and Triton X-100). Then in order to lyse nuclei, 4M $(NH_4)_2SO_4$ and 1 M spermidine were added to give final concentrations of 0.4 M and 5 mM, respectively. After gentle stirring, the lysate was centrifuged at 12000 X g for 5 min. Proteins in supernatant were precipitated by the addition of an equal volume of 4M $(NH_4)_2SO_4$. Following centrifugation at 12000 X g for 15 min, the precipitate was resuspended in buffer-B and dialyzed for 4 hr against buffer-C (the same as buffer-B, but also containing 10% glycerol). Insoluble materials were removed by centrifugation at 12000 X g for 10 min. The supernatant, designated crude nuclear extracts, was quickly frozen in liquid nitrogen and stored at -80 °C.

Fractionation of crude nuclear extracts

DNA binding proteins were partially purified according to the method described by Suzuki *et al* [91]. Strategy of the fractionation of binding activity was summarized in Figure 4. For fractionation of proteins, the crude nuclear extract was dialyzed against buffer D (20 mM HEPES-KOH, pH 7.9, 60 mM

KCl, 5 mM 2-mercaptoethanol and 10% glycerol) and applied to a phosphocellulose column (P11; Whatman, England). After washing of the column with three column volumes of buffer D, proteins were eluted with buffer D with increasing concentrations of KCl. After fractionation, each fraction was dialyzed against buffer C (20 mM HEPES-KOH, pH 7.9, 60 mM KCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol and 10% glycerol) with three changes of 2000ml of buffer and stored at -80°C. For fractionation on a heparin column, P300 plus P400 (containing RBP-1 and RBP-3) or P500 (RBP-2) in the case of *Vicia faba,* or crude nuclear extracts in *Daucus carota,* were applied to an Econo-pac heparin cartridge (Bio-Rad, U.S.A.). After washing with three cartridge volumes of buffer C, proteins were eluted with buffer C that contained increasing concentrations of KCl. After fractionation, each fraction was dialyzed against buffer C with three buffer changes and stored at -80 °C. All steps were carried out at 4°C.

Gel-retardation assay

Binding reactions were carried out in buffer-C containing 0.05 pmol end-labeled probes (about 5 X 10^4 cpm) and 1μg of poly(dI-dC)•poly(dI-dC) by the addition of a nuclear protein fraction. The assay mixture was incubated at 25 °C for 30 min and loaded onto 5% polyacrylamide gels in 6.7 mM Tris-HCl pH

7.5, 1 mM EDTA and 3.3 mM sodium acetate [88]. Following electrophoresis at 5 V/cm, with constant recirculation, the gels were dried for autoradiography.

DNase I footprinting and methylation interference analysis

DNase I footprinting experiments were performed as described by Singh *et al.* [85] with slight modifications. After elution of free and protein bound DNA fragments, each sample was purified using DEAE-Sephacel column chromatography [79]. Methylation interference experiments were performed as described by Weinberger *et al.* [98].

Analysis of binding stability

The stability of binding of nuclear protein fractions to DNA was analyzed by the previously described methods [57], with the exception that the binding reactions were carried out for 20 min at room temperature and then the incubation was continued on ice. The experimental protocol is summarized in Figure 23A.

Screening of λgt11 cDNA library of Vicia faba

A cDNA library was constructed from embryonic axes of *Vicia faba* poly(A)+ RNA with cDNA synthesis and λgt11 cloning kits (Amersham). Screening of cDNA expression was performed as described [97] except that the binding buffer contained 20 mM HEPES-KOH (pH 7.9), 60 mM KCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.25% non-fat milk, poly(dI-dC)•poly(dI-dC) (10 μg/ml) and nick-translated probe DNA (0.01 μg/ml). After binding reaction, the filters were washed three times with aliquots of binding buffer except for probe DNA and poly(dI-dC)•poly(dI-dC). And then the filters were subjected to autoradiography.

Analysis of β-galactosidase recombinant proteins by western transfers

Y1089 lysogens harboring λgt11 or λgt11-NO.2 were isolated [36] and induced to express high levels of their respective β-galactosidase fusion proteins. Cells from 1.25 ml aliquots of the induced lysogen cultures were rapidly pelleted, and resuspended with 100 μl aliquots of SDS-PAGE loading buffer. After holding at 25 °C for 5 min, proteins in each sample (15 μl) were resolved by 10% SDS-PAGE. The separation gel was then soaked in 25 mM Tris, 190 mM glycine (pH 8.3), 20% (v/v) methanol for 30 min. Proteins were electrophoretically transferred onto a polyvinylidene difluoride (PVDF) membrane (0.45 μm) using the same buffer. After transfer, proteins bound to the filter were denatured by 6 M guanidine-HCl and renatured. To assay for DNA binding activity, the filters were blocked with blocking buffer contained 20 mM HEPES-KOH (pH 7.9), 60 mM KCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 5% non-fat milk. The DNA binding activities were examined by the methods described for screening of λgt11 cDNA library of *Vicia faba*.

Results

Structure of the IGS region of Daucus carota rDNA

Figure 5A and 5B shows the complete nucleotide sequence of 5775 bp of the IGS of a cloned *Daucus carota* rDNA. The borders between the spacer and the coding regions were inferred from sequences reported for other plants. A diagrammatic representation of the organization pattern is shown in Figure 6. Eight kinds of repeated sequences denoted A, B, C, D, E, F, G and H, respectively, were presented in this region. Repeat A consisted of two sequences of approximately 150 bp, which were separated by insertion of a 10 bp sequence. They showed a high value of sequence similarity (91% identical), as shown in Figure 7. Repeat B consisted of three copies of approximately 35 bp sequences. Two of them, B2 and B3, were arranged in tandem and another sequence, B1, was separated by an inserted sequence of 168 bp. The values of their sequence similarity was in the range of 84% to 97% identical (Fig. 7). The inserted 168 bp sequence, between B1 and B2, consisted of a tandem array of three copies of an approximately 60 bp sequence, which was named Repeat C. The values of their sequence similarity were in the range of 80% to 98% identical (Fig. 7). Repeat D consisted of three sequences of approximately 180 bp, which were arranged in tandem, and their sequence similarity was in the range of 88% to 92% identical (Fig. 7). Repeat E consisted of two sequences of approximately 40 bp and their sequence similarity was 84% identical (Fig. 7). Repeat F consisted of four sequences of approximately 50 bp and located just behind the Repeat E. The values of their sequence similarity was in the range of 68% to 89% identical (Fig. 7). Repeat G consisted of a tandem array of three complete and one truncated copy of a 455 bp sequence unit. The existence of this repeat sequence was reported previously [92]. The difference of reiteration of this repeat sequence caused the length heterogeneity of the rDNA repeating unit of *Daucus carota* [92]. The last repeat sequence, Repeat H, consisted of two sequences of 228 bp, which were separated by 1909 bp sequences containing Repeat D, E and F. Their sequences showed high conservation (98% identical).

Analysis of transcription initiation site in Daucus carota *rDNA*

I carried out S1 nuclease protection analyses to determine the 5' starting site of *Daucus* rRNA transcription using various fragments derived from DPE2 as hybridization probes. The results revealed the occurrence of several protected fragments of different lengths (data not shown). The initiation site for the RNA transcript that gave rise to the major longest S1 fragment corresponded to the A at position 3127

(Fig. 8B). It is highly probable that the position 3127 is the initiation site for rRNA transcription because the surrounding sequences were similar to those of *Vicia*, maize, wheat and radish. On the other hand, this nucleotide located in Repeat H. The sequence of Repeat H corresponded to the position from -217 to +11 relative to the presumed transcription initiation site. Thus, IGS of *Daucus carota* rRNA gene may contain plural putative promoter sequences. The sequence of H1 was designated as a spacer promoter and H2 was designated as a gene promoter, although the correct promoter sequence is unknown.

For comparison of transcription activity between gene promoter and spacer promoter, I carried out S1 nuclease protection analyses using both fragments containing a gene promoter and spacer promoter as probes (Fig. 8A). When the fragment containing a gene promoter was used as the probe, a clear protected band was detected. The initiation site for the RNA transcript that gave rise to this protected fragment corresponded to the A residue at 3127 (Fig. 8B). On the other hand, when the fragment containing a spacer promoter was used as the probe, several bands were detected. Two major bands corresponded to the A residues at 990 and 986. One of them, the A residue at 990, agreed with the A residue at 3127 of the gene promoter from a comparison of the surrounding sequences. However, these bands were weak and disappeared when the units of S1 nuclease were increased (Fig. 8C). These results indicated that the gene promoter acts as the *cis* element of the major transcription.

In vitro *transcription of* Vicia faba *rDNA*

For determination of *cis* elements required rRNA transcription, I used *in vitro* transcription system using whole cell extract and exogenous rDNA fragments. At first, I prepared cell extract from suspension cultured cells of *Daucus carota*. However, any clear transcripts were not detected using several templates derived from plasmid DPE2. At second, I prepared cell extract from *Vicia faba* embryonic axes. Figure 9 shows the DNA fragments as templates derived from plasmid pVrMH. These fragments were composed of 478 bp rDNA fragments containing initiation sites (from -363 to +115) and a part of a vector sequence at each end. The total length of the template and expected length of transcript were also indicated (Fig. 9A). The results of this assay are shown in Figure 9B. The accurate transcripts were observed using each template, although additional bands were also detected. Therefore accurate transcription of the rRNA gene occurred in this system. Thus, I used this system for determination of *cis* elements required rRNA transcription.

Figure 10 shows the results of *in vitro* transcription using several fragments derived 5' deletion mutants of *Vicia faba* rDNA promoter as templates. Each reaction was carried out using control template (0.1 μg) which was composed of 478 bp rDNA fragment and each 5'-deleted template (0.5 μg) which has different 5'-end each other (Fig. 10A). The size of expected transcript from control template was 255 base and the expected size of transcript from each 5'-detected template was 381 bp. The results of this assay are shown in Fig. 10B. This results was not clear, since several bands appeared whose origins were not clearly located. However, I considered the following possibility. Since I can detect very weak 255 base band in all reactions (lanes 1-8), it is likely that accurate transcription was performed in each reaction. Using 5'del-363 fragment as a template, 381 base band was clearly observed (lane 1). When 5'del-121, -91 and -62 fragments were used as templares, I was able to detect very weak 381 base band (lanes 2, 3 and 4). However, when 5'del-45, -41 fragments and only vector fragment were used as template, 381 base band was not detected (lanes 6-8). Although the result using 5'del-54 was not clear, the size of detected band may be different with 381 base (lane 5). These results suggest that *cis* element required for transcription of *Vicia faba* rDNA may locate in downstream sequence from -363.

Gel-retardation assay using Daucus carota *nuclear extract*

In *Daucus carota*, gel-retardation assays were performed using crude nuclear extracts and fragments derived from gene promoter and spacer promoter regions (Fig. 11A). In the case of gene promoter, a retarded band was detected only when the gp2 fragment (-285 to -126 relative to the putative transcription initiation site at +1) was used for the probe (Fig. 11B, lanes 4-6). On the contrary, when other fragments were used as probes, I was unable to detect any retarded bands (Fig.11B, lanes 1-3 and lanes 7-9). In the case of spacer promoter, a retarded band was detected only when the sp1 fragment (-290 to -126 relative to the putative transcription initiation site at +1) was used for the probe (Fig. 11B, lanes 10-12). On the contrary, when other fragments were used as probes, I was unable to detect any retarded bands (Fig. 11B, lanes 13-15).

Competition assay using gene promoter and spacer promoter sequence in Daucus carota *rDNA*

When crude nuclear extracts were added in the binding reaction mixture, I failed to detect clear retarded bands (lanes 5, 6, 11 and 12 in Fig. 11B). Therefore, the specific binding proteins were partially purified by chromatography on a heparin-sepharose column. After absorption of the nuclear extracts, protein fractions were obtained by stepwise elution with a buffer containing 100, 200, 300, 400 and 500 mM KCl. Using the protein fraction which was eluted with 300 mM KCl, two retarded bands were detected (Fig. 12). Furthermore, these detected bands, when gp2 fragment was used as a probe, were similar to the bands when sp1 fragment was used as a probe. To confirm the sequence specificity of protein binding, competition analyses were performed. When both the gp2 and sp1 fragments were used as probes, two retarded bands were strongly attenuated by the addition of the gp2 or sp1 fragment as a competitor (lanes 3-6 in Fig. 12). In contrast, these retarded bands were not affected by the addition of a 100-fold excess of another DNA fragment as a competitor (lane 7 in Fig. 12). These results indicated that these DNA-protein(s) interactions are specific to the DNA sequences containing the gene promoter and spacer promoter.

Gel-retardation assay using Vicia faba *nuclear extract*

In the case of *Vicia faba*, Kato *et al.* have already determined the entire nucleotide sequence of the IGS of a *Vicia faba* rRNA gene [44]. As shown in Figure 13, One of them, named 'Repeat-C', is highly conserved among *Vicia faba* rRNA gene units [46] and the difference of the reiteration of Repeat C causes length heterogeneity among the rRNA gene units [100]. The transcription initiation site, determined by S1 nuclease analysis, is located at a position 2013 bp downstream from the 3' end of the 25 S coding region [44]. By *in vitro* transcription system using variable fragments derived from 5' deletion mutants of *Vicia faba* rDNA promoter as templates, I could roughly determine *cis* element required for accurate transcription of rDNA. Therefore, I carried out gel-retardation assay using nuclear extract of *Vicia faba*.

Three kinds of fragments derived from the IGS region, including the initiation site (Fig. 13), were used as probes for gel-retardation assays. As shown in Figure 14, a retarded band was observed only when a *Dra* I-*Eco*T 14I fragment (-122 to +96 relative to the transcription initiation site) was used for the probe (Fig. 14, lanes 5-8). This retarded band could not be obtained when the nuclear fraction was heated or SDS was added to the binding reaction mixture (data not shown). Thus, this retarded band may be produced by DNA- protein(s) interaction, although whether this interaction is due to a single protein or complex of several protein is not known. Using other fragments, on the contrary, I was unable to detect any retarded bands (Fig. 14, lanes 1-4 and 9-12).

Sequence specificity of this DNA-protein(s) interaction was confirmed by competition experiments. When the competitor DNA was identical to the probe, the retarded band was strongly attenuated (Fig. 15, lanes 3-5). On the other hand, there was little competition even when a 100-fold excess of *Escherichia coli* DNA or *Mlu* I-*Dra* I fragment (-363 to -123) was added. These results support the hypothesis that this DNA-protein(s) interaction is specific for the DNA sequence. By the addition of the upstream repeat sequence (Repeat C) as the competitor, there was slight competition (Fig. 15, lanes 12-14).

In order to define more precisely the sequence involved in the interaction with protein fraction, shorter probes prepared from deletion mutants were examined in the gel-retardation assay (Fig. 16). When both fragments excised from 3'Δ-34 and 5'Δ-62 were used as probes, retarded bands were observed (Fig. 16, lanes 1-6). However, using other fragments excised from 5'Δ-54 and 3'Δ-59, I was unable to detect any retarded bands (Fig. 16, lanes 7-12). These results show that the nuclear protein fraction may interact with the sequence from -62 to -34.

Column fractionation of Vicia faba *nuclear extracts*

In order to investigate DNA-protein interactions more precisely, the specific DNA-binding proteins were partially purified from the crude nuclear extracts by phosphocellulose column chromatography (strategy of fractionation of binding activity was summarized in Figure 4). After washing of the column with three volumes of buffer D, protein fractions were eluted from the column by buffer D that contained 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1000 mM KCl (the eluted fractions were designated P100, P200, P300, P400, P500, P600, P700, P800, P900 and P1000, respectively). Gel-retardation assays using these protein fractions which *Nsp* V-*Mlu* I (WT) fragment, which contained the region from -94 to -34, were carried out. As shown in Figure 17, several fractions gave clear retarded bands. To confirm the specificity of the binding ability, competition experiments were performed (Fig. 18). The WT fragment and the *Hin*d III-*Eco*R I fragment from pIBI31 were used as competitors in binding reactions. When the competitor DNA was identical to the probe DNA (WT fragment), the intensities of bands A, B, C and D, indicated by arrowheads in Figure 18, were considerably reduced (lanes 7, 8, 12, 13, 17, 18, 22 and 23). These bands were not, however, affected by the addition of a nonspecific competitor (lanes 9, 10, 14, 15, 19, 20, 24 and 25). These results suggested that the DNA-protein(s) interactions were specific for the DNA sequence of the WT fragment. I refer to the protein fractions that generated retarded bands B, C and D as RBP-1, RBP-2 and RBP-3 [rRNA gene promoter-binding protein(s)], respectively. The DNA-binding protein fraction that Nakajima *et al.* reported previously [61] corresponds to RBP-1 since the intensitiy of band B was the strongest of all the retarded bands and the mobility of band B resembled that of the previously reported complex.. In addition, band A was not stably generated in my system, so I was not able to analyze the corresponding complex in detail.

P300 plus P400 and P500 were separately collected by chromatography on a heparin column. Protein fractions were obtained from the column by stepwise elution with buffer C that contained 100, 200, 300 and 400 mM KCl (designated H100, H200, H300 and H400, respectively). As shown in Figure 19A, I performed gel-retardation assays with these fractions derived from P300 plus P400. RBP-1 (band B) was eluted with buffer C that contained 200 mM KCl, and RBP-3 (band D) was eluted with buffer C that contained 300 mM KCl. In the case of P500 (Fig. 19B), RBP-2 (band C) was eluted in P500/H300 and P500/H400 (lanes 8-11). These data suggest that RBP-1, RBP-2 and RBP-3 are chromatographically distinguishable. In lanes 8 and 9 (Fig. 19B), the signals (free and retarded band) are weak. There is a possibility that nuclease activity was present in the P500/H300 fraction. The results of column fractionations were summarized in Figure 4.

Sequences recognized by the three proteins fractions

To identify the region of the WT fragment that binds the nuclear proteins, I performed DNase I footprinting experiments. In the case of the top strand, as presented in Figure 20A, with DNA extracted from band B (lane B), the region from -61 to -53 was protected from DNase I by the binding of RBP-1. In the case of bands C (lane C) and D (lane D), the region from -61 to -54 was protected in each case. In the case of the bottom strand, the DNA extracted from bands B, C and D was protected over the same region, from -59 to -52 (Fig. 20B lanes B, C and D). The intensities of signals due to the residues located from -52 to -47 (Fig. 20A, lane D) and the A residue at position -50 (Fig. 20B, lanes B, C and D) were enhanced. These data suggest that the fragment was easily digested by DNase I at this position because of the binding of the protein fractions. Consequently, it appears that RBP-1, RBP-2 and RBP-3 bind independently to nearly the same region near the site of initiation of transcription.

Methylation interference analysis

To obtain further information about the binding site(s), I performed methylation interference analysis. In the case of the top strand (Fig. 21A), when DNA was extracted from bands B, C and D (lanes B, C and D), the intensities of the signals corresponding to the G residue at position -61 were remarkably reduced as compared with the free probe. In the case of the bottom strand (Fig. 21B), the intensities of the signals corresponding to the G residue at position -56 (lanes B, C and D) was reduced. These results suggest that these two G residues have important roles in the binding of RBP-1, RBP-2 and RBP-3.

The affinity of RBP-1, RBP-2 and RBP-3 to promoter sequences with mutation

The results of the DNase I footprinting experiments indicated that three protein fractions bound to a DNA element near the site of initiation of transcription, and apparently important G residues were identified by methylation interference analysis. Therefore, I introduced point mutations in the putative binding region (from -62 to -52, see Fig. 22A) and each mutated DNA was examined for its ability to bind to the three protein fractions. Gel-retardation assays were carried out using each mutated DNA as probe (Fig. 22B and 22C). When M-59, M-57, M-56, M-55 or M-54 was used as the probe, the three retarded bands were not detected (lanes 5, 7, 8, 9 and 10 in panels B and C). When M-60, M-58 or M-53 was used as the probe, the intensities of the three bands were very low (lanes 4, 6 and 11).

A nuclear protein in radish that binds to dA/dT-rich sequences in promoter and upstream regions has been reported (Echeverria et al., 1992). Therefore, adenine nucleotides from -51 to -47 were replaced by guanine nucleotides, and the mutated DNA was used in gel-retardation assays. When $M^{-51}/_{-47}$ was used as a probe, the three bands were not detected (lane 13 in panels B and C). These results supported the results of competition experiments with mutated DNA (Fig. 22D and 22E).

The data obtained with mutated promoter sequences suggest that the five residues at positions -59, -57, -56, -55 and -54, and a stretch of adenine residues from -51 to -47 were essential for DNA-protein(s) interactions. Furthermore, the residues at positions -60, -58 or -53 affected the stability DNA-protein(s) interactions with RBP-1, RBP-2 and RBP-3.

Stability of binding of RBP-1, RBP-2 and RBP-3 to promoter sequence

In order to characterize more precisely the interactions of RBP-1, RBP-2 and RBP-3 with DNA, I carried out an analysis of binding stability. Binding reactions were performed as described above, and competitor DNA that was identical to the probe DNA was added at the indicated times (Fig. 23A, see A-G) during the incubation. When P300 plus P400 was used for the binding reaction, the intensities of bands B and D were strongly reduced independently of the duration of incubation with the competitor (Fig. 23B, lanes A-G). This results suggests that the stability of binding to DNA of RBP-1 and RBP-3 was extremely low. In the case of P500, when the duration of incubation was increased (from G to B in panel A), the retarded band C gradually disappeared (Fig. 23C, lanes G-B). These results indicate that the complex of probe DNA and RBP-2 was much more stable than that of DNA and RBP-1 or DNA and RBP-3.

The affinity of nuclear protein and rDNA promoter sequence in heterogeneous system

In order to investigate DNA-protein interactions in heterologous system, I carried out gel-retardation assay. As shown in Figure 24, when WT fragment (which contained the region from -94 to -34 of *Vicia faba*) and 300 mM KCl heparin column fraction of *Daucus carota* were added in the binding reaction mixture, I was unable to detect any retarded bands (lanes 2-4). On the contrary, when crude nuclear extract of *Vicia faba* and gp2 fragment or sp1 fragment were added in the binding reaction mixture, I detected one retarded band, respectively. To confirm the specificity of these retarded bands, I carried out competition assay. When competitor was identical to the probe DNA (lanes 6, 7, 11 and 12), retarded bands disappeared. And when *Hin*d III-*Eco*R I fragment from pIBI31 was used as a competitor, retarded bands also disappeared. These results suggested that the DNA-protein(s) interactions were not specific for the DNA sequence of the probe DNA.

DNA binding proteins during germination

In *Vicia faba,* it takes 29 hours to reach to the maximum level of transcription of rRNA gene after soaking of seeds. If the binding abilities of RBP-1, RBP-2 and RBP-3 are interchangeable during germination, it is possible that these protein fractions are closely related to the initiation of transcription of rRNA gene. So, I prepared the crude nuclear extracts from 6 stages during germination. As shown in Figure 25, I was able to obtained at least three retarded bands (RBP-1, RBP-2 and RBP-3) in all stages. These results suggest that binding activities of three proteins exist through the germination of *Vicia faba.*

Molecular cloning of DNA binding protein that interact with the DNA sequence near the transcription initiation site

To study the structure and function of DNA binding proteins that interact with the sequence located about 60 bp upstream from the transcription initiation site, I attempted to isolate its cDNA clone from a *Vicia faba* cDNA library in λgt11.

From a screen of 1 x 10^6 phage plaques, about 20 phages that bound to the probe were isolated after three rounds of screening. For one of them, termed λgt11-NO.2, I determined the specificity of the DNA binding protein encoded by the recombinant phage in extracts of an induced phage lysogen. Lysogen extracts were subjected to SDS-PAGE and transferred to polyvinylidene difluoride (PVDF) filters. After a denaturation-renaturation procedure, the filters were treated with the ^{32}P-labeled probe. A single fusion protein band of ~140 kD was detected in the lysogen extract from λgt11-NO.2 (Fig. 26, lane 2), whereas no such band was found in a lysogen extract from nonrecombinant λgt11 (Fig. 26, lane 1).

Sequence specificity of this DNA-protein(s) interaction was confirmed by competition experiments. When the competitor DNA was contained a region from -94 to -34, the band was strongly attenuated (Fig. 26, lanes 4, 6). On the other hand, there was little competition even when a 50-fold excess of nonspecific competitor DNAs (Fig. 26, lanes 8, 10) was added. These results suggested that this DNA-protein(s) interaction is specific for the DNA sequence.

The nucleotide sequence of the cDNA insert was determined (Fig. 27). The deduced amino acid sequence indicated that λgt11-NO.2 encodes a 295-amino acid polypeptide in frame with the β-galactosidase of λgt11.

Discussion

Comparison of IGS organization between Vicia faba *and* Daucus carota

I present the complete nucleotide sequence of 5775 bp of the IGS of a cloned *Daucus carota* rDNA. I identify eight kinds of repeated sequences. Figure 28 summarized the overall structure of the IGS region of the *Daucus carota* and *Vicia faba* rDNA. In both species, several sequences are repeated to various extents. Though the number of repeated sequences are different in *Daucus carota* and *Vicia faba*, respectively, the IGS region of two species are similar with each other in the patterns of their molecular organization. This organization is similar to that observed with the large spacer in wheat [4]. In both species, there is one repeated sequence which caused the length heterogeneity of the rDNA repeating unit by the difference of reiteration of this repeat sequence. However, these two species have different features. There is a heterogeneity in length of ETS (external transcribed spacer) in *Daucus carota*, on the other hand, there is no heterogeneity in *Vicia faba*. Furthermore, there are two promoter-like sequences in *Daucus carota*. Each unit of Repeat H contained the TATATAGGG motif which was a conserved sequence around the transcription initiation site of plant rRNA [19, 27, 44, 65, 68, 95]. The sequence of Repeat H was corresponded to the region at -217 to +11 relative to the presumed transcription initiation site. Thus, IGS of *Daucus carota* rRNA gene contains plural putative promoter sequences, like *Arabidopsis* [30], mouse [67], *Xenopus* [71] and *Drosophila* [15]. In these rDNAs, it was reported that one of the reported promoter sequence is genuine and the others may act as enhancer sequences.

Comparison of sequence around transcription initiation site in plant rDNA

Repeat H contained the TATATAGGG motif which was a conserved sequence around the transcription initiation site of plant rRNA [19, 27, 44, 65, 68, 95]. So, I investigated whether another conserved sequence exist or not in the promoter region of rRNA gene in other higher plants. As shown in Figure 29A, I found two consensus sequences TATG and CAGG. In plant RNA polymerase I transcription, sequences around the initiation site are conserved in higher plants. However, the proteins that interact with those sequences have not been identified except for a few species and a commonly defined *cis* control element in a wide range of plant species has not been clearly defined. It is possible that these sequences have the same function on rRNA transcription in these plants.

Furthermore, I found purine-rich regions lying downstream of these consensus sequences, except for the case of maize rRNA genes. I therefore speculate that these consensus sequences are commonly required for the interaction with DNA-binding protein(s) on rRNA transcription in plants. Therefore, it is possibly that this consensus sequence seems to be important for initiation of transcription of rDNA.

Furthermore, I searched for similar sequences with this consensus sequence in the promoter region of other species. As shown in Figure 29B, I found the TATG/CAGG-like sequence in human [7], mouse [14], rat [86], *Xenopus* [66], *Acanthamoeba* [38, 47], yeast [12] and *Physarum* [84]. In these species except for *Physarum polycephalum*, the sequence containing the TATG/CAGG motif is important for DNA-protein interaction [7, 14, 60, 67, 38, 12]. Consequently, it is possible that this sequence or its homologue may be a *cis* control element as a contact site for rRNA transcription factor(s).

In vitro *transcription system of rRNA gene in plant*

The availability of *in vitro* system for the study of transcription has been of prime importance in elucidating the mechanisms of gene regulation. In studies of animals, several systems containing protein extract from cells have been used in attempts to determine the *cis* elements and *trans* factors required for accurate initiation of transcription by RNA polymerase I, II or III [31, 53, 78].

On the other hand, there has been little success in the development of *in vitro* transcription system from higher plant cells. Recently, *in vitro* transcription systems for RNA polymerase II [103, 104, 74, 26] and RNA polymerase III genes [26] were developed. Furthermore, Yamashita *et al.* described an *in vitro* transcription system containing a soluble whole cell extract from *Vicia faba* embryonic axes [102]. Accurate transcription being dependent on exogenously added *Vicia faba* rDNA fragment was observed in this system. Moreover, the resistance to α-amanitin showed that this transcription occurred with RNA polymerase I.

In this study, I attempt to develop *in vitro* transcription system for RNA polymerase I in *Daucus carota.* Though I used carrot cultured cells for preparation of whole cell extract, I could not develop it. There are some reason of my failure. For the first reason, the whole cell extract from carrot cultured cells may involve more digestive enzyme comparing with the case of *Vicia faba.* Another possibility is that there are several secondary metabolic products in carrot cultured cells. Therefore, it might be necessary for me to fractionate the whole cell extract.

In animal, RNA polymerase I transcription system is known to be species specific. In order to study whether plant rRNA gene transcription mechanisms is species specific or not, I carried out the *in vitro* transcription using whole cell extract of *Vicia faba* and cloned *Daucus carota* rDNA fragments. However, any significant transcripts were not detected (data not shown). Thus, in plant, RNA polymerase I transcription system may be also species specific. But, this result was preliminary, since we have failed to develop the *in vitro* transcription system using *Daucus carota* cell extract.

Cis *elements required for accurate initiation of transcription of rRNA genes in plant*

For determination of *cis* elements required for rRNA transcription, I used *in vitro* transcription system using whole cell extract and exogenous rDNA fragments (Fig. 10). The results was not clear, since several bands appeared whose origins were not clearly located. However, I considered the following possibility. Since I can detect very weak 255 base band in all reactions (lanes 1-8), it is likely that accurate transcription was performed in each reaction. Using 5'del-363 fragment as a template, 381 base band was observed (lane 1). When 5'del-121, -91 and -62 fragments were used as templates, I was also able to detect very weak 381 base band (lanes 2, 3 and 4). However, when 5'del-45, -41 fragments and only vector fragment were used as template, 381 base band was not detected (lanes 6-8). Although the result using 5'del-54 was not clear, the size of detected band may be different with 381 base (lane 5). So, there is a possibility that *cis* element required for transcription of *Vicia faba* rDNA may locates in downstream from -363, and it is likely that there are two *cis* elements in this region. One element may be located within the region from -363 to -121 and another element may be located in downstream from -62. From the results of DNase I footprinting analysis (Fig. 20), methylation interference analysis (Fig. 21) and gel-retardation assays using mutated DNAs (Fig. 22), the region of proximal *cis* element (downstream from -62) corresponded to the binding site of RBP-1, RBP-2 and RBP-3. So, the results of run-off transcription assay for determination of *cis* element may be supported. There have been few studies of the *cis* elements that regulate the transcription of rRNA. Doelling *et al.* [22] reported that the region from -91 to +22 relative to the site of initiation of transcription was sufficient to represent the functional promoter of an rRNA gene of *Arabidopsis* in an *Arabidopsis* transient expression system. From the comparison of the *cis* elements in *Vicia faba* and *Arabidopsis thaliana*, I found that abovementioned consensus sequence TATG and CAGG were contained in both *cis* elements (Fig. 29A). Therefore, it is possibly that this consensus sequence seems to be important for initiation of transcription of rDNA.

In animals and yeast systems of rDNA transcription, CPE (core promoter element) that is about 60 bp DNA sequence located near the transcription start site (+1) is essential for basal transcription and upstream element located upstream of CPE act as enhancer [50, 99, 12]. From the comparison of location of *cis* element, this feature may be common in RNA polymerase I transcription system.

Functions of gene promoter and spacer promoter in Daucus carota rDNA

In order to determine the site of transcription initiation, I carried out S1 nuclease protection analyses. In the case of gene promoter, a clear protected band was detected. The results of S1 analyses indicated that the gene promoter acts as the *cis* element of the major transcription, but there is a possibility that the spacer promoter also acts as the same element *in vivo*. Similar results in which both promoters were active *in vivo*, were reported in the transcription of *Arabidopsis* rDNA [22].

In the case of *Arabidopsis*, the transcripts from both promoters were detected clearly. The reason why we failed to detect a clear transcript from the spacer promoter is unknown, but the following reason is suggested. The nucleotide sequences of the gene promoter and spacer promoter were highly conserved and this conserved sequence spread from -217 to +11. The gene promoter has a long AT-rich sequence in its upstream region, although the AT-rich sequence located in the upstream region of the spacer promoter is short in length. AT-rich sequences upstream from the rDNA promoter may be important for rRNA transcription in plants [24]. Thus, the difference in the length of AT-rich sequences may be related to the transcription activity of rRNA in *Daucus carota.*

On the other hand, when I carried out gel-retardation assay using fragments derived from gene promoter (gp2 fragment) and spacer promoter (sp1 fragment), I could obtain the similar band pattern. And in competition analysis, two detected bands were disappeared by the addition of gp2 fragment or sp1 fragment. These results indicated that these DNA-protein(s) interactions are specific to the DNA sequences containing the gene promoter and spacer promoter. Furthermore, it is possible that the same protein fraction may interact with both DNA sequences of the gene promoter and the spacer promoter. Though the function of these proteins is unknown, it is possible that they play roles for transcription of rDNA because of the sequence specificity of binding. For example, spacer promoter may have an enhancer function for the transcription of the precusor rRNA according to the manner in which they acted as entry sites of a transcription factor. When I carried out S1 nuclease protection analysis, I could obtain the weak protected band (Fig. 8). Therefore, I suggested that precursor rRNA is actually transcribed from the spacer promoter, but transcripts from the spacer promoter is very unstable. In fact, when I used the probe which have short annealing region in S1 nuclease protection analysis, I could detect a faint band. However, when I used the probe which have long annealing region, I could not detect any protected bands. These results supported my idea.

In *Drosophila*, *Xenopus* and mouse, some of the repeats in IGS were shown to have an enhancer function for the transcription of the precursor rRNA according to the manner in which they acted as entry sites [15, 71, 67]. So, it is possible that some of the repeated sequences may act as a spacer promoter, like *Xenopus*, *Drosophila* and *Arabidopsis*. In *Arabidopsis* in fact, it was reported that the upstream repeated sequence containing a spacer promoter increased transcription from a *Xenopus* ribosomal gene promoter in injected frog oocytes, although the enhancer activity only had a small effect on transcription in the transfected protoplast of *Arabidopsis thaliana* [22].

In the case of *Vicia faba*, the results of the competition assay using Repeat C (Fig. 15), which was located in the external spacer region (ESR), suggest that the same nuclear protein(s) interacts with promoter sequences and repeat sequences in the ESR. Therefore, I speculate that Repeat C located in ESR functions as the spacer promoter as is the case for *Xenopus* rRNA gene [66, 73]. It resembles the model proposed for maize rRNA gene [82]. Repetitive elements in ESRs are generally thought to play an important role in the regulation of transcription by trapping transcription factors.

Protein binding sequences in rDNA promoter

Figure 30 shows a schematic representation of the binding properties of RBP-1, RBP-2 and RBP-3. I identified the binding regions for RBP-1, RBP-2 and RBP-3 by DNase I footprinting experiments (Fig. 20). The brackets correspond to the regions protected by DNase I digestion. Furthermore, I determined the residues which are required for stable complex formations by methylation interference experiments (Fig. 21), and also by gel-retardation assays with mutated DNA (Fig. 22). Nucleotides with arrowheads in the sequence indicate essential G residues for the binding of the three protein fractions determined by methylation interference experiments (Fig. 21). From the gel-retardation assays with mutated DNA (Fig. 22), the nucleotides with white lettering show essential residues for binding of RBP-1, RBP-2 and RBP-3, and the closed circles in the sequence indicate the residues that have strong effects on the stability of DNA-protein(s) interactions. Consequently, RBP-1, RBP-2 and RBP-3 can bind to nearly the same region and required the same residues for their respective stable complex formations.

In higher plants, some protein fractions that bind to the upstream sequences of rRNA genes have been reported. In maize, a protein that binds to the promoter region of the rRNA gene has been reported [82]. The binding region is located from -40 to -9 relative to the site of initiation of transcription. Comparison of the binding site in *Vicia faba* and maize shows two differences. As shown in Figure 31, the location of the binding site corresponding to the initiation site was different and the DNA sequences around the binding site were mostly non-homologous. However, despite these differences, it was found that both binding sites shared two consensus sequences, TATG and CAGG. Each internal sequence between these consensus sequences is very different (TCTA in *Vicia faba*, CCCCT in maize). In the case of radish, dA/dT-rich sequences were identified as nuclear protein(s)-interacted sequences [24]. The *Vicia* and radish DNA-binding protein(s) may be distinct. However, the functions of these proteins are still unknown. Therefore, whether the difference is caused by the species specificity or different functions is a further problem.

I then searched for similar sequences with this consensus sequence in the promoter region of other plant rRNA genes (Fig. 29A). It was found that this consensus sequence was conserved in all of them except for maize. Although DNA-protein(s) interactions in the promoter region have not been reported in them, it is possible that these sequences have the same function on rRNA transcription in the binding sites of proteins in these plants. Furthermore, I found purine-rich regions lying downstream of these consensus sequences, except for the case of maize rRNA genes. Though there are consensus sequences in rRNA gene promoter of *Daucus carota*, when gel-retardation assay using gp3 or sp2 and crude nuclear extracts, I was not able to detect any retarded bands. It is possibly that there is a binding activity in carrot for the consensus sequence, but my gel-retardation system is not suitable for detection of the activity. I therefore speculate that these consensus sequences, which constitute a bi-partite structure, are commonly required for the interaction with DNA-binding protein(s) on rRNA transcription in plants.

My conclusion that the TATG(N)XCAGG sequence is important for DNA-protein(s) interaction in the rRNA gene promoter region, was partially supported by the methylation interference experiments and gel-retardation assays using synthetic oligonucleotides containing point mutations. The methylation of G residues (-61 in the top strand and -56 in the bottom strand) in the consensus sequence impairs protein(s) binding. On the other hand, from the analyses using point mutations, different results were obtained. Two G residues at positions -54 (top strand) and -59 (bottom strand) are important nucleotides, although the methylation of these G residues does not interfere with protein(s) binding. Furthermore, G-A transition at position -61 has no effect on protein(s) binding, although the methylation of this nucleotide impairs it. The difference of base transition from base modification may cause this discrepancy. Although I do not have more results, six residues at positions -59, -57 -56, -55 -54 and adenine stretch from -51 to -47 are essential nucleotides for these DNA-protein(s) interactions. Other nucleotides, T (-60), T (-58), and G (-53) may relate to the stability of DNA-protein(s) interaction.

Relationships among RBP-1, RBP-2 and RBP-3

As shown in Figure 30, I revealed the characteristics for binding of RBP-1, RBP-2 and RBP-3. I identified the binding regions for RBP-1, RBP-2 and RBP-3 by DNase I footprinting experiments (Fig. 20). Furthermore, I determined the residues which are required for stable complex formations by methylation interference experiments (Fig. 21), and also by gel-retardation assays with mutated DNA (Fig. 22). Consequently, RBP-1, RBP-2 and RBP-3 can bind to nearly the same region and required the same residues for their respective stable complex formations.

The binding properties of the three protein fractions are very similar. To characterize the binding properties more precisely, I analyzed binding stability. When incubation with a competitor was carried out at room temperature, the intensities of retarded bands B, C and D were strongly reduced, independently of the duration of incubation with competitor (data not shown). Therefore, incubation with the competitor was carried out on ice. The results in Figures 23B and 23C show that DNA-binding stability of RBP-1 and of RBP-3 was extremely low and the complex of probe DNA and RBP-2 was much more stable than that of probe DNA and RBP-1 or RBP-3. Putnam *et al.* reported that xUBF, a factor that binds to a *cis* element within the rRNA gene promoter of *Xenopus*, has a high rate of dissociation from DNA, and that xUBF has a half-life of about 1 min [69]. Since our three protein fractions and xUBF have high rates of dissociation from DNA, there is a possibility that this property might be common to all *trans*-acting factors that participate in transcription by RNA polymerase I.

From the next results, 1) The mobility of bands B, C and D were different in the gel-retardation assay, 2) they could bind to almost the same region, 3) they required the same residues for stable binding to DNA, 4) they were chromatographically distinguishable, 5) DNA-binding stability of RBP-1 and of RBP-3 was extremely low and the complex of probe DNA and RBP-2 was much more stable than that of probe DNA and RBP-1 or RBP-3, I can suggest that three possibilities about the relationship among RBP-1, RBP-2 and RBP-3. It is possible that the three retarded bands reflect distinct proteins or complexes of protein bound to the probe DNA. The second possibility is that a certain retarded band was formed by the interaction of an additional protein(s) with a DNA-protein(s) complex. The third is that the three retarded bands were created by modification, such as phosphorylation, of the binding protein(s). With respect to the second possibility, in vertebrate transcription systems that involve RNA polymerase I, two kinds of transcription factor have been identified. One transcription factor, UBF, binds to *cis* elements upstream of the rRNA gene. The DNA-UBF complex is then recognized by another transcription factor, SL1, that consists of TBP and TAFs [16]. To examine this possibility, we carried out gel-retardation assays, in which we added two or three protein fractions together in the same binding reaction, but we failed to detect any additional retarded bands (data not shown). The functions of transcription factors are known to be affected by phosphorylation. For example, *trans*-acting factors such as CREB [101], Sp1 [40] and UBF [97, 64] are phosphorylated and their binding properties are affected as a result. In higher plants, HBP-1b, a factor that binds to a hexameric motif of the wheat gene for histone H3, is affected by phosphatase treatment which alters its ability to bind to a DNA element *in* *vitro* [93]. I treated several fractions with phosphatase, but failed to detect any different patterns in the gel-retardation assay (data not shown). However, the abovementioned possibilities cannot be entirely ruled out by the results of these experiments.

In higher plant, there are several reports concerned DNA-protein(s) interactions in rDNA transcription. In cucumber, the same proteins recognize two binding sites in the promoter region of rRNA gene, namely the upstream binding element between -164 and -105 and the core promoter between -41 and +16. The three binding proteins (16, 22 and 24 kDa, respectively) have been isolated [108]. In radish, an NER (nuclear extracts from young radish leaves) protein fraction bound to a repeated sequence (located at -1077 to -740 from the transcription start site), and a region (-120 to -55) containing the putative promoter has been reported [24]. Recently, Echeverria *et al.* reported the presence of a protein factor, designated NF B, in radish [25]. This factor binds to a unique sequence located 164 bp downstream of the transcription start site and this region corresponds to the site at which pre-rRNA is processed. A putative role was proposed for NF B as a factor that couples transcription and processing of pre-rRNA is discussed. In wheat, the presence of a DNA-binding protein bound to the promoter of the rRNA gene and upstream repeats has also been reported [39]. From their DNA-binding specificities, these protein factors appear to be different from RBP-1, RBP-2 and RBP-3. Although the function of our protein fractions is still unknown, it seems possible that the three protein fractions might function cooperatively to regulate transcription by RNA polymerase I.

In *Vicia faba*, it is known that it takes 29 hours to reach to the maximum level of transcription of rRNA gene after soaking of seeds (unpublished data). If the binding abilities of RBP-1, RBP-2 and RBP-3 are interchangeable during germination, it is possible that these protein fractions are closely related to the initiation of transcription of rRNA gene. Therefore, I carried out gel-retardation assay using crude nuclear extract from several stages of germination. As shown in Figure 25, I was able to obtained at least three retarded bands (RBP-1, RBP-2 and RBP-3) in all stages. These results suggest that binding activities of three proteins exist through the germination of *Vicia faba*. It is likely that the activation of transcription of rDNA is independent of presence or absence of these factors.

Molecular cloning of a DNA binding protein that interact with the DNA sequence near the transcription initiation site

The nucleotide sequence of the insert of λgt11-NO.2 was determined (Fig. 27) and the deduced amino acid sequence indicated that λgt11-NO.2 encodes a protein of 295 amino acids in length.

A computer search for related protein sequences revealed definite similarity between this protein and nucleic acid binding protein, such as ribonucleoprotein [51, 5], nucleolin [105] and FMV3bp [21]. Amino acids 153 to 158 and 249 to 254 contain the regions homologous to nucleolin, other ribonucleoprotein and FMV3bp. These regions have been implicated in nucleic acid binding [62].

In cucumber, the same proteins recognize two binding site in the promoter region of rRNA gene, namely the upstream binding element between -164 and -105 and the core promoter between -41 and +16. And the three binding protein have been isolated [108]. These proteins also seem to interact with the single strands of the respective DNA regions suggesting an effect on transcriptional regulation while the DNA is transcribed and, therefore, is single-stranded. Although whether the protein encoded by λgt11-NO.2 can bind to a single-stranded DNA or not is unknown, this protein have two nucleic acid binding domains. So, there is a possibility that the protein encoded by λgt11-NO.2 can bind to single-stranded DNA and may play an important role in the RNA polymerase I transcription system.

Acknowledgments

I wish to express my sincere appreciation to Associate Professor Atsushi Kato, Emeritus Professor Shigeyuki Tanifuji and Professor Yoshibumi Komeda of Hokkaido University for their invaluable suggestions and constructive criticisms.

My sincere thanks are also due to Drs. Fumio Takaiwa and Seiichi Toki of National Institute of Agrobiological Resources, and Assistant Professor Taku Takahashi of Hokkaido University for their valuable suggestions and criticisms. Special thanks must be accorded to all members of the laboratory of Hokkaido University for useful discussions on the results.

References

1. Appels R, Honeycutt RL: rDNA: Evolution over a billion years. In: Dutta SK (ed) DNA systematics, 2: 82-135. CRC Press Boca Raton FL (1986).

2. Appels R, Molan LB, Gustafson JP: The structure of DNA from the rye (*Secale cereale*) NOR R1 locus and its behavior in wheat backgrounds. Can J Genet Cytol 28: 673-685 (1986).

3. Bachvarov D, Moss T: The RNA polymerase I transcription factor xUBF contains 5 tandemly repeated HMG homology boxes. Nucl Acids Res 19: 2331-2335 (1991).

4. Barker RF, Harbert NP, Jarvis MG, Flavell RB: Structure and evolution of the intergenic region in a ribosomal DNA repeat unit of wheat. J Mol Biol 201: 1-17 (1988).

5. Bar-Zvi D, Shagan T, Schindler U, Cashmore AR: RNP-T, a ribonucleoprotein from *Arabidopsis thaliana*, contains two RNP-80 motifs and a novel acidic repeat arranged in an α-helix conformation. Plant Mol Biol 20: 833-838 (1992).

6. Bell SP, Jantzen H-M, Tjian R: Assembly of alternative multiprotein complexes directs rRNA promoter selectivity. Genes Devel 4: 943-954 (1990).

7. Bell SP, Pikaard CS, Reeder RH, Tjian R: Functional cooperativity between transcription factors UBF1 and SL1 mediates human ribosomal RNA synthesis. Science 241: 1192-1197 (1988).

8. Bell SP, Pikaard CS, Reeder RH, Tjian R: Molecular mechanisms governing species-specific transcription of ribosomal RNA. Cell 59: 489-497 (1989).

9. Berk AJ, Sharp PA: Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell 12: 721-732 (1977).

10. Borisjuk N, Hemleven V: Nucleotide sequence of the potato rDNA intergenic spacer. Plant Mol Biol 21: 381-384 (1993).

11. Cavanaugh AH, Hempel WN, Taylor LJ, Rogalsky V, Todorov G, Rothblum LI: Activity of RNA polymerase I transcription factor UBF blocked by Rb gene product. Nature 374: 177-180 (1995).

12. Choe SY, Schultz MC, Reeder RH: *In vitro* definition of the yeast RNA polymerase I promoter. Nucl Acids Res 20: 279-285 (1992).

13. Christine K, Dinkl E, Grummt F: An intrinsically bent region upstream of the transcription start site of the rRNA gene of *Arabidopsis thaliana* interacts with an HMG-related protein. Plant Mol Biol 27: 705-713 (1995).

14. Clos J, Buttgereit D, Grummt I: A purified transcription factor (TIF-IB) binds to essential sequences of the mouse rDNA promoter. Proc Natl Acad Sci USA 83: 604-608 (1986).

15. Coens ES, Dover GA: Multiple Pol I initiation sequences of rDNA spacers of *Drosophila melanogaster*. Nucl Acids Res 10: 7017-7026 (1982).

16. Comai L, Tanese N, Tjian R: The TATA-binding protein and associated factors are integral components of the RNA polymerase I transcription factor, SL1. Cell 68: 965-976 (1992).

17. Comai L, Zomerdijk JCBM, Beckmann H, Zhou S, Admon A, Tjian R: Reconstitution of transcription factor SL1

- exclusive binding of TBP by SL1 or TFIID subunits. Science 266: 1966-1972 (1994).

18. Darocha PSCF, Bertrand H: Structure and comparative analysis of the rDNA intergenic spacer of *Brassica rapa* -implications for the function and evolution of the cruciferae spacer. Eur J Biochem 229: 550-557 (1995).

19. Delcasso-Tremousaygue D, Grellet F, Panabieres F, Ananiev ED, Delseny M: Structure and transcriptional characterization of the external spacer of a ribosomal RNA nuclear gene from a higher plant. Eur J Biochem 172: 767-776 (1988).

20. De Winter RFJ, Moss T: Spacer promoters are essential for efficient enhancement of *X. laevis* ribosomal transcription. Cell 44:313-318 (1986).

21. Didier DK, Klee HJ: Identification of an Arabidopsis DNA-binding protein with homology to nucleolin. Plant Mol Biol 18: 977-979 (1992)

22. Doelling JH, Gaudino RJ, Pikaard CS: Functional analysis of *Arabidopsis thaliana* rRNA gene and spacer

promoters *in vivo* and by transient expression. Proc Natl Acad Sci USA 90: 7528-7532 (1993).

23. Eberhard D, Tora L, Egly JM, Grummt I: A TBP-containing multiprotein complex (TIF-IB) mediates transcription specificity of murine RNA polymerase I. Nucl Acids Res 21:4180-4186 (1993).

24. Echeverria M, Delcasso-Tremousaygue D, Delseny M: A nuclear protein binding to dA/dT-rich sequences upstream from the radish rDNA promoter. Plant J 2: 211-219 (1992).

25. Echeverria M, Penon P, Delseny M: Plant ribosomal DNA external spacer binding factors: a novel protein binds specifically to a sequence close to the primary pre-rRNA processing site. Mol Gen Genet 243: 442-452 (1994).

26. Fan H, Sugiura M: A plant basal *in vitro* system supporting accurate transcription of both RNA polymerase II- and III-dependent genes: supplement of green leaf component(s) drives accurate transcription of a light-responsive rbsS gene. EMBO J 14: 1024-1031 (1995).

27. Gerster J, Schiebel K, Waldburg G, Hemleben V: Complex organization of the length heterogeneous 5' external spacer of mung bean (*Vigna radiata*) ribosomal DNA. Genome 30: 723-733 (1988).

28. Glibetic M, Taylor L, Larson D, HannanR, Sells B, Rothblum L: The RNA polymerase I transcription factor UBF is the product of a primary response gene. J Biol Chem 270: 4209-4212 (1995).

29. Grimaldi G, Di Nocera PP: Multiple repeated units in *Drosophila melanogaster* ribosomal DNA spacer stimulate rRNA precursor transcription. Proc Natl Acad Sci USA 85: 5502-5505 (1988).

30. Gruendler P, Unfried I, Pascher K, Schweizer D: rDNA intergenic region from *Arabidopsis thaliana* structural analysis, Interspecific variation and functional Implications. J Mol Biol 221: 1209-1222 (1991).

31. Grummt I: Specific transcription of mouse ribosomal DNA in a cell-free system that mimics control *in vivo*. Proc Natl Acad Sci USA 78: 727-731 (1981).

32. Grummt I, Roth E, Paule MR: Ribosomal RNA transcription *in vitro* is species specific. Nature 296: 173-174 (1982).

33. Haltiner MM, Smale ST, Tjian R: Two distinct promoter elements in the human rRNA gene identified by linker scanning mutagenesis. Mol Cell Biol 6: 227-235 (1986).

34. Haltiner JM, Learned RM, Tjian R: Analysis of clustered point mutations in the human ribosomal RNA gene promoter by transient expression *in vivo.* Proc Natl Acad Sci USA 85: 669-673 (1988).

35. Hisatake K, Nishimura T, Maeda Y, Hanada K, Song C-Z, Muramatsu M: Cloning and structural analysis of cDNA and the gene for mouse transcription factor UBF. Nucl Acids Res 19: 4631-4637 (1991).

36. Huynh TV, Young RA, Davis RW: Constructing and screening cDNA libraries in lgt10 and lgt11. In DNA Cloning-A Practical Approach, Volume 1, DM Glover, ed. (Oxford: IRL Press), 49-78 (1985).

37. Iida CT, Kownin P, Paule MR: Ribosomal RNA transcription: Proteins and DNA sequences involved in

preinitiation complex formation. Proc Natl Acad Sci USA 82: 1668-1672 (1985).

38. Iida CT, Paule MR: Purification of components required for accurate transcription of ribosomal RNA from *Acanthamoeba castellanii.* Nucl Acids Res 20: 3211-3221 (1992).

39. Jackson SD, Flavell RB: Protein-binding to reiterated motifs within the wheat rRNA gene promoter and upstream repeats. Plant Mol Biol 20: 911-919 (1992).

40. Jackson SP, MacDonald JJ, Lees-Miller S, Tjian R: GC box binding induces phosphorylation of Sp1 by a DNA-dependent protein kinase. Cell 63: 155-165 (1990).

41. Jacob ST: Regulation of ribosomal gene transcription. Biochemical J 306: 617-626 (1995).

42. Jantzen H-M, Admon A, Bell SP, Tjian R: Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. Nature 344: 830-836 (1990).

43. Jhones MH, Learned RM, Tjian R: Analysis of clustered point mutations in the human ribosomal RNA gene promoter

by transient expression *in vivo*. Proc Natl Acad Sci USA 85: 669-673 (1988).

44. Kato A, Nakajima T, Yamashita J, Yakura K, Tanifuji S: The structure of the large spacer region of the rDNA in *Vicia faba* and *Pisum sativum*. Plant Mol Biol 14: 983-993 (1990).

45. Kato A, Yakura K, Tanifuji S: Organization of ribosomal DNA in the carrot. Plant Cell Physiol 23: 151-154 (1982).

46. Kato A, Yakura K, Tanifuji S: Repeated DNA sequences found in the large spacer of *Vicia faba* rDNA. Biochem Biophys Acta 825: 411-415 (1985).

47. Kownin P, Iida CT, Brown-Shimer S, Paule MR: The ribosomal RNA promoter of *Acanthamoeba castellanii* determined by transcription in a cell-free system. Nucl Acids Res 13: 6237-6247 (1985).

48. Labhart P, Reeder RH: Enhancer-like properties of the 60/81 bp elements in the ribosomal gene spacer of *Xenopus laevis*. Cell 37: 285-289 (1984).

49. **Lakshmikumaran M, Nagi MS:** Structural analysis of two length variants of the rDNA intergenic spacer from *Eruca sativa.* Plant Mol Biol 24: 915-927 (1994).

50. **Learned RM, Learned TK, Haltiner MM, Tjian R:** Human rRNA transcription is modulated by the coordinate binding of two factor to an upstream control element. Cell 45: 847-857 (1986).

51. **Li Y, Sugiura M:** Three distinct ribonucleoprotein from tobacco chloroplasts: each contains a unique amino terminal acidic domain and two ribonucleoprotein consensus motifs. EMBO J 9: 3059-3066 (1990).

52. **MacDonald R, Swift GH, Przybyla AE, Chirgwin M:** Isolation of RNA using guanidium salts. In: Berger SL, Kimmel AR (eds) Methods in Enzymology, vol 152: 219-227. Academic Press New York (1987).

53. **Manley JL, Fire A, Cano A, Sharp PA, Gefter ML:** DNA-dependent transcription of adenovirus genes in a soluble whole-cell extract. Proc Natl Acad Sci USA 77: 3855-3859 (1980).

54. Maxam AM, Gilbert W: Sequencing end-labeled DNA with base specific chemical cleavages. In: Grossman L, Moldave K (eds) Methods in Enzymology vol 65: 499-560. Academic Press New York (1980).

55. McMullen MD, Hunter B, Phillips RL, Rubenstein I: The structure of the maize ribosomal DNA spacer region. Nucl Acids Res 14: 4953-4968 (1986).

56. McStay B, Hu C-H, Pikaard CS, Reeder RH: xUBF and Rib1 are both required for formation of a stable polymerase I promoter complex in *X. laevis*. EMBO J 10: 2297-2303 (1991).

57. Mikami K, Sakamoto A, Iwabuchi M: The HBP-1 family of wheat basic/leucine zipper proteins interacts with overlapping *cis*-acting hexamer motifs of plant histone genes. J Biol Chem 269: 9974-9985 (1994).

58. Miller KG, Tower J, Sollner-Webb B: A complex control region of the mouse rRNA gene directs accurate initiation by RNA polymerase I. Mol Cell Biol 5: 554-562 (1985).

59. Morrow BE, Johnson SP, Warner JR: Protein that bind to the yeast rDNA enhancer. J Biol Chem 264: 9061-9068 (1989).

60. Nagamine M, Kishimoto T, Aono J, Kato H, Kominami R, Muramatsu M: Sequestration analysis for RNA polymerase I transcription factors with various deletion and point mutations reveals different functional regions of the mouse rRNA gene promoter. Mol Cell Biol 7: 1486-1495 (1987).

61. Nakajima T, Suzuki A, Tanifuji S, Kato A: Characterization of nucleotide sequences that interact with a nuclear protein fraction in rRNA gene of *Vicia faba*. Plant Mol Biol 20: 939-949 (1992).

62. Olson MOJ, Kirstein MN, Wallace MO: Limited proteolysis as a probe of the conformation and nucleic acid binding regions of nucleolin. Biochemistry 29: 5682-5686 (1990).

63. O'Mahony DJ, Rothblum LI: Identification of two forms of the RNA polymerase I transcription factor UBF. Proc Natl Acad Sci USA 88: 3180-3184 (1991).

64. **O'Mahony DJ, Xie W, Smith SD, Singer HA, Rothblum LI**: Differential phosphorylation and localization of the transcription factor UBF *in vivo* in response to serum deprivation. J Biol Chem 267: 35-38 (1992).

65. **Perry KL, Palukaitis P**: Transcription of tomato ribosomal DNA and the organization of the intergenic spacer. Mol Gen Genet 221: 102-112 (1990).

66. **Pikaard CS, McStay B, Schultz MC, Bell SP, Reeder RH**: The *Xenopus* ribosomal gene enhancers bind an essential polymerase I transcription factor xUBF. Genes Devel 3: 1779-1788 (1989).

67. **Pikaard CS, Pape LK, Henderson SL, Ryan K, Paalman MH, Lopata MA, Reeder RH, Sollner-Webb B**: Enhancers for RNA polymerase I in mouse ribosomal DNA. Mol Cell Biol 10: 4816-4825 (1990).

68. **Polanco C, Vega PDL**: The structure of the rDNA intergenic spacer of *Avena sativa* L.: a comparative study. Plant Mol Biol 25: 751-756 (1994).

69. **Putnam CD, Pikaard CS**: Cooperative binding of the *Xenopus* RNA polymerase I transcription factor xUBF to

repetitive ribosomal gene enhancers. Mol Cell Biol 12: 4970-4980 (1992).

70. Raina SN, Ogihara Y: Ribosomal DNA repeat unit polymorphism in 49 Vicia species. Theo Appl Genet 90:477-486 (1995).

71. Reeder RH: Enhancers and ribosomal gene spacers. Cell 38: 349-351 (1984).

72. Reeder RH, Pennock D, MacStay B, Roan J, Tolentino E, Walker P: Linker scanner mutagenesis of the *Xenopus laevis* ribosomal gene promoter. Nucl. Acids Res 15: 7429-7441 (1987).

73. Reeder RH, Roan JG, Dunaway M: Spacer regulation of Xenopus ribosomal gene transcription: Competition in oocytes. Cell 35: 449-456 (1983).

74. Roberts MW, Okita TW: Accurate *in vitro* transcription of plant promoters with nuclear extracts prepared from cultured plant cells. Plant Mol Biol 16: 771-786 (1991).

75. Rudloff U, Eberhard D, Tora L, Stunnenberg H, Grummt I: TBP-associated factors interact with DNA and

govern species specificity of RNA polymerase I transcription. EMBO J 13: 2611-2616 (1994).

76. Rudloff U, Stunnenberg HG, Keaveney M, Grummt I: Yeast TBP can replace its human homologue in the RNA polymerase I-specific multisubunit factor SL1. J Mol Biol 243: 840-845 (1994).

77. Safrany G, Tanaka N, Kishimoto T, Ishikawa Y, Kato H, Kominami R, Muramatsu M: Structural determinant of the species-specific transcription of the mouse rRNA gene promoter. Mol Cell Biol 9: 349-353 (1989).

78. Sakonju S, Bogenhagen DF, Brown DD: A control region in the center of the 5S RNA gene directs specific initiation of transcription: I. The 5' border of the region. Cell 19: 13-25 (1980).

79. Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

80. Sanger F: Determination of nucleotide sequences in DNA. Science 214: 1205-1210 (1981).

81. Schiebel K, von Waldburg G, Gerstner J, Hemleven V: Termination of transcription of ribosomal RNA genes of mung bean occur within a 175 bp repetitive element of the spacer region. Mol Gen Genet 218: 302-307 (1989).

82. Schmitz ML, Maier UG, Brown JWS, Feix G: Specific binding of nuclear proteins to the promoter region of a maize nuclear rRNA gene unit. J Biol Chem 264: 1467-1472 (1989).

83. Schmidt-Puchta W, Gunter I, Sanger HL: Nucleotide sequence of the intergenic spacer (IGS) of the tomato ribosomal DNA. Plant Mol Biol 13: 251-253 (1989).

84. Schroth GP, Siino JS, Cooney CA, Th'ng JPH, Ho PS, Bradbury EM: Intrinsically bent DNA flanks both sides of an RNA polymerase I transcription start site. J Biol Chem 267: 9958-9964 (1992).

85. Singh H, Sen R, Baltimore D, Sharp PA: A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobulin gene. Nature 319: 154-158 (1986).

86. Smith SD, Oriahi E, Lowe D, Yang-Yen H-F, O'Mahony D, Rose K, Chen K, Rothblum LI:

Characterization of factors that direct transcription of rat ribosomal DNA. Mol Cell Biol 10: 3105-3116 (1990).

87. Sollner-Webb B, Tower J: Transcription of cloned eukaryotic ribosomal RNA genes. Annu Rev Biochem 55: 801-830 (1986).

88. Strauss F, Varshavsky A: A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37: 889-901 (1984).

89. Sugiura M: personal communication.

90. Suzuki A, Tanifuji S, Komeda Y, Kato A: Structural and functional characterization of the IGS region of the rDNA in *Daucus carota.* (submitted)

91. Suzuki A, Tanifuji S, Komeda Y, Kato A: Three protein fractions that bind to a narrow region of the promoter of the ribosomal RNA gene in *Vicia faba.* Plant Cell Physiol (in press).

92. Taira T, Kato A, Tanifuji S: Difference between two major size classes of carrot rDNA repeating units is due to

reiteration of sequence of about 460 bp in the large spacer. Mol Gen Genet 213: 170-174 (1988).

93. Takase H, Tabata T, Mikami K, Iwabuchi M: Partial purification and characterization of two transcription factors, HBP-1a and HBP-1b, specific for a *cis*-acting element, ACGTCA, of wheat histone genes. Plant Cell Physiol 32: 1195-1203 (1991).

94. Toloczyki C, Feix G: Occurrence of 9 homologous repeat units in the external spacer region of a nuclear maize rDNA gene unit. Nucl Acids Res 14:4969-4986 (1986).

95. Vincentz M, Flavell RB: Mapping of ribosomal RNA transcripts in wheat. Plant Cell 1: 579-589 (1989).

96. Vinson CR, LaMarco KL, Johnson PF, Landschulz WH, McKnight SL: In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Devel 2: 801-806 (1988).

97. Voit R, Schnapp A, Kuhn A, Rosenbauer H, Hirschmann P, Stunnenberg HG, Grummt I: The nucleolar transcription factor mUBF is phosphorylated by casein kinase

II in the C-terminal hyperacidic tail which is essential for transactivation. EMBO J 11: 2211-2218 (1992).

98. Weinberger J, Baltimore D, Sharp PA: Distinct factors bind to apparently homologous sequences in the immunoglobulin heavy-chain enhancer. Nature 322: 846-848 (1986).

99. Windle JJ, Sollner-Webb B: Two distinct and precisely positioned domains promote transcription of *Xenopus laevis* rRNA Genes: Analysis with linker-scanning mutants. Mol Cell Biol 6: 4585-4593 (1986).

100. Yakura K, Kato A, Tanifuji S: Length heterogeneity of the large spacer of Vicia faba rDNA is due to the differing number of a 325 bp repetitive sequence element. Mol Gen Genet 193: 400-405 (1984).

101. Yamamoto KK, Gonzalez GA, Biggs III WH, Montminy MR: Phosphorylation-induced binding and transcriptional efficacy of nuclear factor CREB. Nature 334: 494-498 (1988).

102. Yamashita J, Nakajima T, Tanifuji S, Kato A: Accurate transcription initiation of *Vicia faba* rDNA in a whole cell extract from embryonic axes. Plant J 3: 187-190 (1993).

103. Yamazaki K, Imamoto F: Selective and accurate initiation of transcription at the T-DNA promoter in a soluble chromatin extract from wheat germ. Mol Gen Genet 209: 445-452 (1987).

104. Yamazaki K, Katagiri F, Imaseki H, Chua N-H: TGA1a, a tobacco DNA-binding protein, increases the rate of initiation in a plant *in vitro* transcription system. Proc Natl Acad Sci USA 87: 7035-7039 (1990).

105. Yang T-H, Tsai W-H, Lee Y-M, Lei H-Y, Lai M-Y, Chen D-S, Yeh N-N, Lee S-C: Purification and characterization of nucleolin and its identification as a transcription repressor. Mol Cell Biol 14: 6068-6074 (1994).

106. Yanisch-Perron C, Vieria J, Messing J: Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119 (1985).

107. Zaitseva GN, Kleshchenko EV: Eukaryotic ribosomal RNA gene transcription. Molecular Biology 28: 621-629.

108. Zentgraf, U. and Hemleben V: Complex formation of nuclear proteins 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 (1992).

109. Zomerdijk JCBM, Beckmann H, Comai L, Tjian R: Assembly of transcriptionally active RNA polymerase I initiation factor SL1 from recombinant subunits. Science 266: 2015-2018 (1994).

Figure legends

Figure 1. Plasmid constructs of the IGS region of carrot rRNA gene.

Plasmid DER1 [92] was used for recloning of IGS region. *Eco*R I-*Pst* I fragment (for plasmid DEP1) and *Pst* I-*Eco*R I fragment (for plasmid DEP2) were recloned into *Sma* I site of plasmid pIBI31. *Hin*d III-*Hin*d III fragment (for plasmid DHH1) was recloned into *Hin*d III site of plasmid pIBI31. *Eco*R I, *Pst* I and *Hin*d III indicate restriction enzyme sites.

Figure 2. Plasmid constructs of the IGS region of *Vicia faba* rRNA gene.

Plasmid VER1 [100] was used for recloning of IGS region. *Mlu* I-*Mlu* I fragment (for pVrMM), *Mlu* I-*Hin*d III fragment (for pVrMH) and *Dra* I-*Hin*d III fragment (for pVrDH) were recloned into *Sma* I site of plasmid pIBI31. M, D and H indicate restriction enzyme sites for *Mlu* I, *Dra* I and *Hin*d III, respectively.

Figure 3. Synthetic oligonucleotides for DNAs with point mutation.

Several 30 bp nucleotides from -35 to -64, which contain one or a few mutated nucleotides and 15 bp nucleotide, designated RP, 5'-GCCACCAAATTGGTC-3', which corresponds to the bottom strand from -44 to -30 were synthesized. These synthetic oligonucleotides were used for preparation of DNAs with point mutation. Numbering is relative to the initiation site at +1.

Figure 4. Strategy of the fractionation of DNA binding activities.

For the first step, the crude nuclear extract was applied to a phosphocellulose column. And then each fraction was applied to a heparin column. Each conditions were described in materials and methods.

Figure 5. Nucleotide sequence of the IGS of *Daucus carota* rDNA.

Eight families of repeated sequences are indicated by underlines or boxes (A, B, C, D, E, F, G and H). Bold A residues at 990 and 3027 are putative transcription initiation sites. A; nucleotide sequence from 1 to 2900. B; nucleotide sequence from 2901 to 5775.

Figure 6. Diagrammatic representation of the *Daucus* rDNA IGS region.

The members of each repeated sequence are boxed and marked with alphabetical letters followed by numerals.

Figure 7. Alignment of repeated members of each family of the repeated sequences shown in Figure 5 and the sequence homology.

Hyphens indicate the same nucleotides, and plus signs indicate deletions of the nucleotide. Two members of the Repeat 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 sequences. X marks represent undetermined nucleotides in the consensus sequences.

Figure 8. Identification of the transcription initiation site of the gene promoter and the spacer promoter by S1 nuclease mapping.

A) The gene promoter probe was the *Hin*d III-*Eae* I fragment (*Hin*d III: -126, *Eae* I: +163) labeled at the 5' end of the coding strand. The spacer promoter probe was the *Hind* III-*Eco*T14 I fragment (*Hind* III: -126, *Eco*T14 I: +110) labeled at the 5' end of the coding strand. Numbering is relative to the each putative initiation site at +1. Probe DNA was hybridized with 75 μg carrot RNA (lanes 6, 7, 8 and 9) or yeast tRNA (lane 5), and treated with 10 units (lane 6), 40 units (lanes 5 and 7), 160 units (lane 8) and 640 units (lane 9) of S1 nuclease at 37 °C for 30 min. After denaturation, protected fragments were electrophoresed with the DNA sequence marker (lanes 1 and 10: G, lanes 2 and 11: G+A, lanes 3 and 12: C+T, lanes 4 and 13: C). The arrow head in panel B indicates the endpoint of the clear protected band and the arrow heads in panel C indicate the endpoints of the faint protected bands.

Figure 9. *In vitro* transcription using whole cell extract and exogenous rDNA fragments of *Vicia faba*.

A) Schematic representation of the DNA fragments as template. pVrMH contained *Vicia faba* rDNA fragments (from -363 to +115) in pIBI31. Position numbers show nucleotide distances from the transcription start site (+1, indicated by arrowheads). P, H and B indicate the restriction enzyme sites of *Pvu* II, *Hind* III and *Bgl* I, respectively. Broken lines with arrows indicate expected transcripts. Numerals indicate the template size (bp) and transcript size (b) of each template (i, ii and iii). B) Run-off transcription products using several templates. Lanes 1-3 show the addition of variable length DNA fragments as templates (lane 1: i, lane 2: ii and lane 3: iii). Lane 4 shows no addition of exogenous DNA fragments. Numerals indicate transcript size.

Figure 10. Determination of *cis* elements required for rRNA transcription.

A) Schematic representation of the DNA fragments as template. Plasmids pVrMH and several 5'-deletion mutant plasmids (5'Δ-121, 5'Δ-91, 5'Δ-62, 5'Δ-54, 5'Δ-45 and 5'Δ-41) were used for preparation of DNA templates. Template 8) is a polylinker fragment of pIBI31. Position numbers show nucleotide distances from the transcription start site. B) Run-off transcription products using 5'-deletion templates. Lanes 1-8 show the addition of several 5'-deletion fragments (0.5 μg) as templates (lane 1: template 1, lane 2: template 2, lane 3: template 3, lane 4: template 4, lane 5: template 5, lane 6: template 6, lane 7: template 7 and lane 8: template 8). Template C. (0.1 μg) was added in each run-off transcription reaction mixture as a positive control. Lane M: size marker. Numerals indicate transcripts size.

Figure 11. Gel-retardation assay using gene promoter and spacer promoter in *Daucus carota.*

(A) Schematic diagram illustrating the probe DNA fragment for gel-retardation assay. *Hin*d III, *Eae* I and *Rsa* I indicate each restriction enzyme site. Five fragments (gp1, gp2, gp3, sp1 and sp2) were used as probes. Numbering is relative to the each putative initiation site at +1. (B) Each DNA fragment was used as a probe. Lanes 1-3, gp1 fragment (-368 to -285); lanes 4-6, gp2 fragment (-285 to -126); lanes 7-9, gp3 fragment (-126 to +165); lanes 10-12, sp1 fragment (-290 to -126) and lanes 13-15, sp2 fragment (-126 to +90). Reactions contain 0 μg (lanes 1, 4, 7, 10 and 13), 5 μg (lanes 2, 5, 8, 11 and 14), 10 μg (lanes 3, 6, 9, 12 and 15) of nuclear proteins.

Figure 12. Gel-retardation assay using gp2 and sp1 fragments as probes with various competitors.

The *Hin*d III(at -285)-*Hin*d III(at -126) fragment, named gp2 fragment, was used as a probe in the left panel and the *Rsa* I(at -290)-*Hin*d III(at -126) fragment, named sp1 fragment, was used as a probe in the right panel. Numbering is relative to the each putative initiation site at +1. Lane 1, no nuclear extracts; lane 2, no competitor; lanes 3 and 4, a competitor is identical to the probe DNA; lanes 5 and 6, a competitor is the sp1 and gp2 in the left panel and the right panel, respectively; lane 7, the polylinker segment of pIBI31 was used as a competitor. Each competitor contained 10-fold (lanes 3 and 5) and 100-fold (lanes 4, 6 and 7) that of the probe.

Figure 13. Schematic diagram of the intergenic spacer (IGS) of the *Vicia faba* rRNA gene.

Open triangles indicate Repeat C, and closed triangles indicate other repeat sequences [46]. The arrowhead indicates the transcription initiation site. Mlu (*Mlu* I), Dra (*Dra* I), EcoT (*Eco*T 14I), EcoO (*Eco*O 19I), EcoR I (*Eco*R I), Hinf (*Hin*f I) and Sal (*Sal* I) indicate restriction enzyme sites. Sequences conserved among various plant rRNA genes around the initiation site are shown [44]. Numbering is relative to the initiation site at +1.

Figure 14. Detection of a DNA binding protein of *Vicia faba* with gel-retardation assay.

Assay of three probes near the initiation site. Lanes 1-4, *Mlu* I-*Dra* I (-363 to -123); lanes 5-8, *Dra* I-*Eco*T 14I (-122 to +96); lanes 9-12, *Eco*T 14I-*Sal* I (+97 to +273). Reactions contain 0 μg (lanes 1, 5 and 9), 10 μg (lanes 2, 6 and 10), 20 μg (lanes 3, 7 and 11) and 40 μg (lanes 4, 8 and 12) of nuclear proteins.

Figure 15. Gel-retardation assay with various competitors.

A DNA fragment from -122 to +10 was used for the probe. Lane 1, no nuclear extract; lane 2, no competitor; lanes 3-5, the same fragment as the probe; lanes 6-8, -363 to -123 fragment; lanes 9-11 sonicated *E. coli* DNA; lanes 12-14, repeat C (shown in Fig. 5). Each competitor contained 1-fold (lanes 3, 6, 9 and 12), 10-fold (lanes 4, 7, 10 and 14) that of the probe. The arrowhead indicates bands of the DNA-protein(s) complex.

Figure 16. Gel-retardation assay with deletion mutant probe.

Lanes 1-3, 3'Δ-34; lanes 4-6, 5'Δ-62; lanes 7-9, 5'Δ-54; lanes 10-12, 3'Δ-59. Reaction contain 0 μg (lanes 1, 4, 7 and 10), 10 μg (lanes 2, 5, 8 and 11) and 40 μg (lanes 3, 6, 9 and 12) of nuclear proteins. Junction sequences of rDNA are shown. Thick lines indicate rDNA sequences and thin lines indicate vector sequences. Arrowheads indicate the DNA-protein(s) complex.

Figure 17. Gel-retardation assay using phosphocellulose column fraction.

Gel-retardation assays with fractions from a phosphocellulose column were carried out. In lanes 1-3; P100, In lanes 4-6; P200, in lanes 7-9; P300, in lanes 10-12; P400, in lanes 13-15; P500, in lanes 16-18; P600, in lanes 19-21; P700, in lanes 22-

24; P800, in lanes 25-27; P900 and in lanes 10-12; P1000 was added to the binding reaction. Reaction contain about 200 ng (lanes 1, 4, 7, 10, 13, 16, 19, 22, 25 and 28), about 400 ng (lanes 2, 5, 8, 11, 14, 17, 20, 23, 26 and 29) and about 800 ng (lanes 3, 6, 9, 12, 15, 18, 21, 24, 27 and 30) of nuclear proteins.

Figure 18. Competition analysis with fractions from a phosphocellulose column.

An *Nsp* V-*Mlu* I (WT) fragment that contained the -94 to -34 region was used as the probe. Lanes 1-5; P100, lanes 6-10; P200, lanes 11-15; P300, lanes 16-20; P400, lanes 21-25; P500. Lane 1, 6, 11, 16 and 21; no competitor. In lanes 2, 7, 12, 17, 22 (x10) and 3, 8, 13, 18, 23 (x100), the WT fragment was added to the binding reaction as a competitor. In lanes 4, 9, 14, 19, 24 (x10) and 5, 10, 15, 20, 25 (x100), a polylinker segment of pIBI31 was added.

Figure 19. Fractionation of P300 plus P400 and P500 on a heparin column, and subsequent gel-retardation assays.

Gel-retardation assays with fractions from a heparin column were carried out. P300 plus P400 (panel A) and P500 (panel B) were fractionated on a heparin column. In lanes 2 and 3, The flow-through fraction was added to the binding reaction. In lanes 4 and 5, H100; in lanes 6 and 7, H200; in lanes 8 and 9, H300; and in lanes 10 and 11, H400 was added to the binding reaction. Odd-numbered lanes contained 0.001 μg of poly(dI-dC)•poly(dI-dC). Even-numbered lanes contained 0.01 μg of poly(dI-dC)•poly(dI-dC).

Figure 20. Determination of the binding sites of RBP-1, RBP-2 and RBP-3 by DNase I footprinting experiments.

The WT fragment was used as the probe. The labeled ends of the DNA fragments used for the binding studies were the 3' end of the top strand (panel A) and the 3' end of the bottom strand (panel B). Lane F shows the pattern after DNase I digestion of the free DNA probe. Lanes B, D and C indicate the patterns after DNase I digestion of the DNA probe with bound RBP-1, RBP-2 and RBP-3, respectively. Bars show the protected regions. The ladders on the left are sequencing ladders. Numerals show positions relative to the start site of initiation of transcription.

Figure 21. Methylation interference experiments.

The WT fragment was used as the probe. The labeled ends of the DNA fragments used for binding studies were the 3' end of the top strand (panel A) and the 3' end of the bottom strand (panel B). The [^{32}P]-labeled probe was partially methylated and then incubated with nuclear proteins. Free (lane F) and protein-bound (lanes B, C and D, corresponding to bands B, C and D, respectively) fragments of DNA were separated on 5% polyacrylamide gels and eluted. After piperidine cleavage of the eluted DNA fragments, the resultant products were separated on a 6% sequencing gel, in parallel with the products of sequencing reactions of the same DNA fragment. Closed circles (B, C and D) indicate a G residue that, when methylated, inhibited binding to protein(s). Numerals indicate positions relative to the site of initiation of transcription.

Figure 22. Gel-retardation assays with various mutated DNAs.

Mutated nucleotides in each synthetic DNA are indicated in panel A and lines indicate the same nucleotides as those in the wild-type (WT) DNA. As shown in panels B and C, gel-retardation assays were carried out using fragments 1-13 as probes. P300 plus P400 (panel B) and P500 (panel C) were used as nuclear proteins. In panels D and E, gel-retardation assays using fragments 1-13 as competitors and the WT fragment as a probe were carried out. Each competitor was present at 50 times the level of the probe. P300 plus P400 (panel D) and P500 (panel E) were used as nuclear proteins.

Figure 23. Stability of binding of RBP-1, RBP-2 and RBP-3 to DNA.

The experimental procedure is described in panel A. The competitor DNA that was identical to the probe DNA was added at the indicated times (A-G) during the incubation. Gel-retardation assays using P300 plus P400 (panel B) and P500 (panel C) were carried out. Lanes (-), no competitor (positive control).

Figure 24. Analysis of the affinity of nuclear protein and rDNA promoter sequence in heterogeneous system.

I carried out gel-retardation assay in heterogeneous system. When WT fragment (which contained the region from -94 to -34 of *Vicia faba*) and 300 mM KCl heparin column fraction of *Daucus carota* were added in the binding reaction mixture, I was unable to detect any retarded bands (lanes 1-4). Crude nuclear extract of *Vicia faba* and gp2 fragment (lanes 5-9) or sp1 fragment (lanes 10-14) were added in the binding reaction mixture. Lanes 6, 7, 11 and 12; competitor was identical to the probe DNA. Lanes 8, 9, 13 and 14; competitor was *Hin*d III-*Eco*R I fragment from pIBI31.

Figure 25. Gel-retardation assay using crude nuclear extract from several stages during germination.

Gel-retardation assay using crude nuclear extract from several stages during germination was carried out. Lanes 2 and 3; dormant, lanes 4 and 5; 20 hrs after soaking, lanes 7 and 8; 24 hrs after soaking, lanes 9 and 10; 28 hrs after soaking, lanes 12 and 13; 32 hrs after soaking, lanes 14 and 15; 36 hrs after soaking. Reaction contain 0 μg (lanes 1, 6, and 11), 15 μg (lanes 2, 4, 7, 9, 12 and 14) and 40 μg (lanes 3, 5, 8, 10, 13 and 15) of nuclear proteins.

Figure 26. Blot analyses of the proteins extracted from the induced lysogens of λgt11-NO.2 and λgt11.

Bacterial lysogens were prepared, and each lysogen extract was separated by SDS-PAGE on a 10% gel and transferred to PVDF filters. The filters were probed with the ^{32}P-labeled DNA fragment. Lanes 1, 3, 5, 7 and 9, lysogen of λgt11 (negative control); lanes 2, 4, 6, 8 and 10, lysogen of λgt11-NO.2. Lanes 1 and 2, no competitor; lanes 3 and 4, rDNA promoter-1 (*Nsp* V-*Mlu*-I fragment of plasmid 3'Δ-34) was added in the binding mixture as a competitor; lanes 5 and 6, rDNA promoter-2 (*Nsp* V-*Hin*d III fragment of plasmid 3'Δ-34); lanes 7 and 8, M4-RV

PCR product from plasmid pIBI31; lanes 9 and 10, *Eco*R I-*Hin*d III fragment of plasmid pIBI31.

Figure 27. DNA sequence of the λgt11-NO.2 cDNA and the deduced amino acid sequence.

Numbering of nucleotides and amino acids begins with the first in-frame AUG codon. Amino acid sequence is shown using single letters. Asterisk indicate the stop codon. Deduced nucleic acid binding domain are underlined.

Figure 28. Diagrammatic representation of the *Daucus carota* and *Vicia faba*.

The members of each repeated sequence are boxed and marked with alphabetical letters. In *Daucus carota* (upper), there are eight kinds of repeated sequences and in *Vicia faba* (lower), there are five kinds of repeated sequences.

Figure 29. Comparison of the promoter sequences of rRNA genes.

Higher plant rRNA genes (A) and rRNA genes of other species (B). Larger letters in bold indicate similar sequences to TATG(N)xCAGG. Under lines indicate the purine-rich regions.

Numbering is relative to the transcription initiation site at +1; core, core promoter of human [8] or yeast [12]; UCE, upstream control element [7]; initiation, gene promoter; spacer, spacer promoter; T3, terminator region closed gene promoter [55].

Figure 30. Schematic representation of the binding properties of RBP-1, RBP-2 and RBP-3.

Brackets correspond to the regions protected by DNase I digestion. Nucleotides with arrowheads in the sequence indicate essential G residues for binding of the three protein fractions determined by methylation interference experiments. Nucleotides with white lettering show the essential residues for binding of RBP-1, RBP-2 and RBP-3, and the closed circles in the sequence indicate the residues that have strong effects on the stability of DNA-protein(s) interactions. Numerals indicate positions relative to the site of initiation of transcription.

Figure 31. Comparison of the protein(s) binding sites of *Vicia faba* and maize rRNA genes.

Conserved sequences are boxed. Brackets correspond to the regions protected in DNase I footprinting experiments. Numbering is relative to the transcription initiation site at +1.

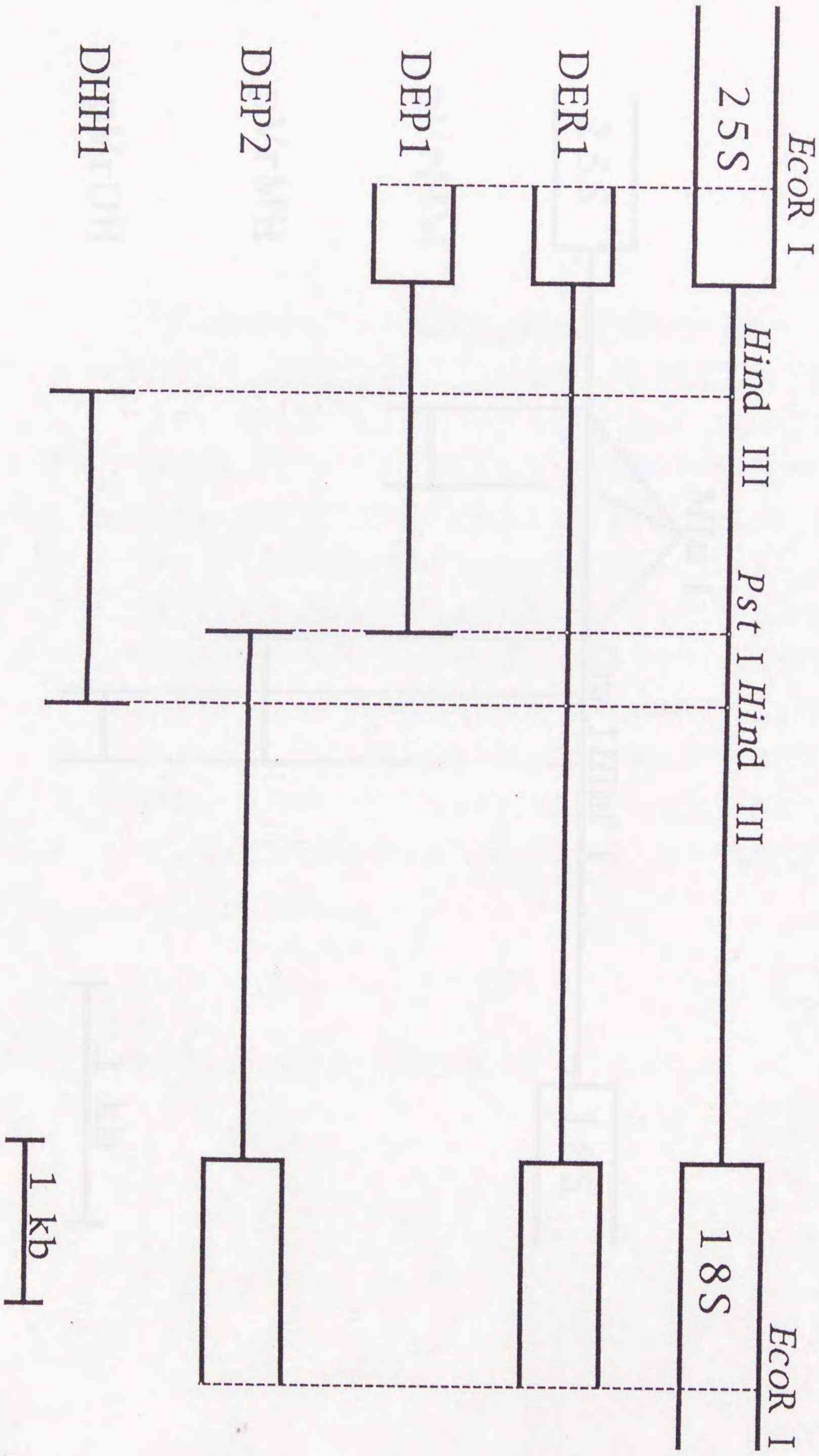

EcoR I
25S
Hind III
Pst I
Hind III
18S
EcoR I
DER1
DEP1
DEP2
DHH1
1 kb

Figure 1

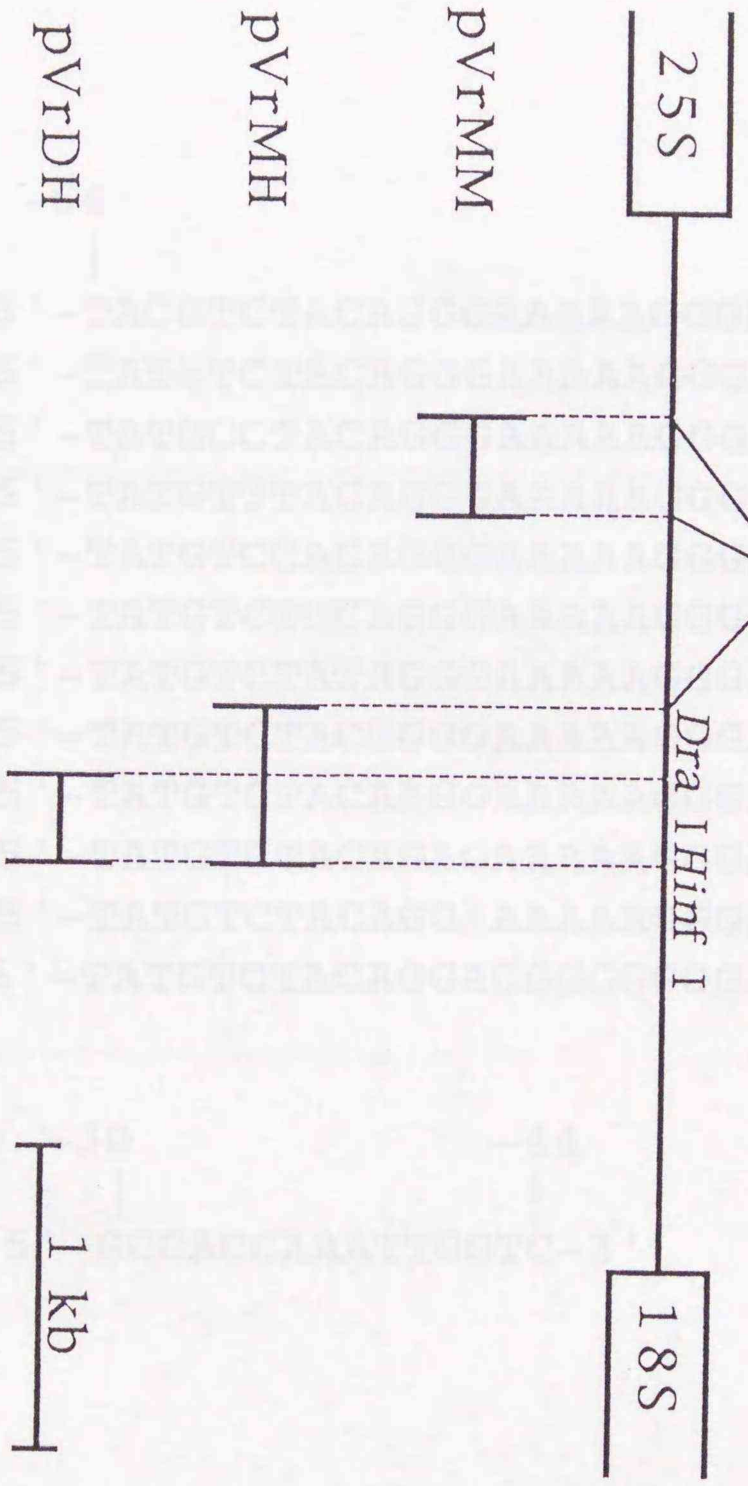


Figure 2

```
              -64                           -35
               |                             |
M-62       5'-TACGTCTACAGGGAAAAAGGGACCAATTTG-3'
M-61       5'-TATATCTACAGGGAAAAAGGGACCAATTTG-3'
M-60       5'-TATGCCTACAGGGAAAAAGGGACCAATTTG-3'
M-59       5'-TATGTTTACAGGGAAAAAGGGACCAATTTG-3'
M-58       5'-TATGTCCACAGGGAAAAAGGGACCAATTTG-3'
M-57       5'-TATGTCTGCAGGGAAAAAGGGACCAATTTG-3'
M-56       5'-TATGTCTATAGGGAAAAAGGGACCAATTTG-3'
M-55       5'-TATGTCTACGGGGAAAAAGGGACCAATTTG-3'
M-54       5'-TATGTCTACAAGGAAAAAGGGACCAATTTG-3'
M-53       5'-TATGTCTACAGAGAAAAAGGGACCAATTTG-3'
M-52       5'-TATGTCTACAGGAAAAAAGGGACCAATTTG-3'
M-51/-47   5'-TATGTCTACAGGAGGGGGGGGACCAATTTG-3'


               -30         -44
                |           |
RP          5'-GCCACCAAATTGGTC-3'
```

Figure 3

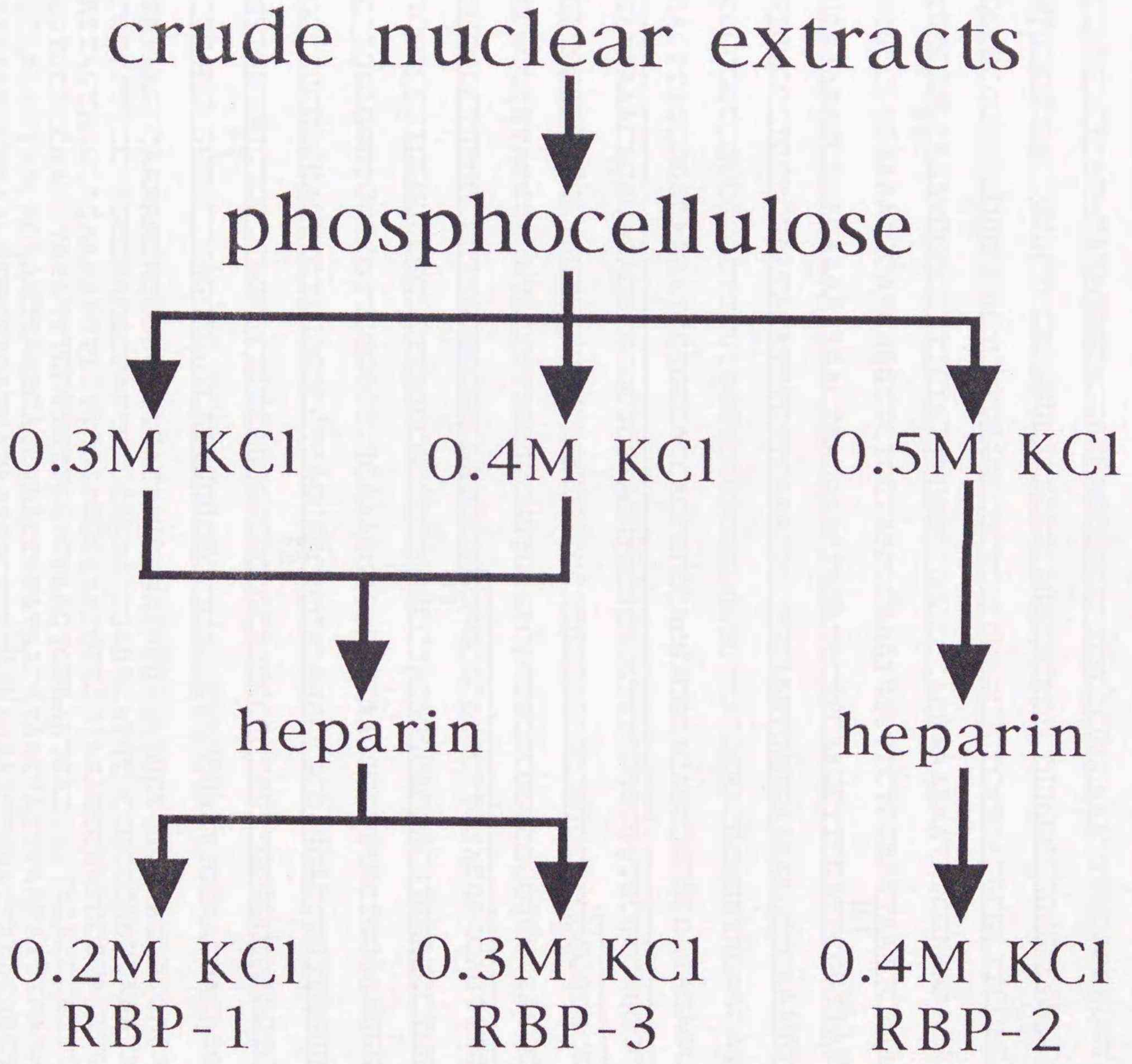
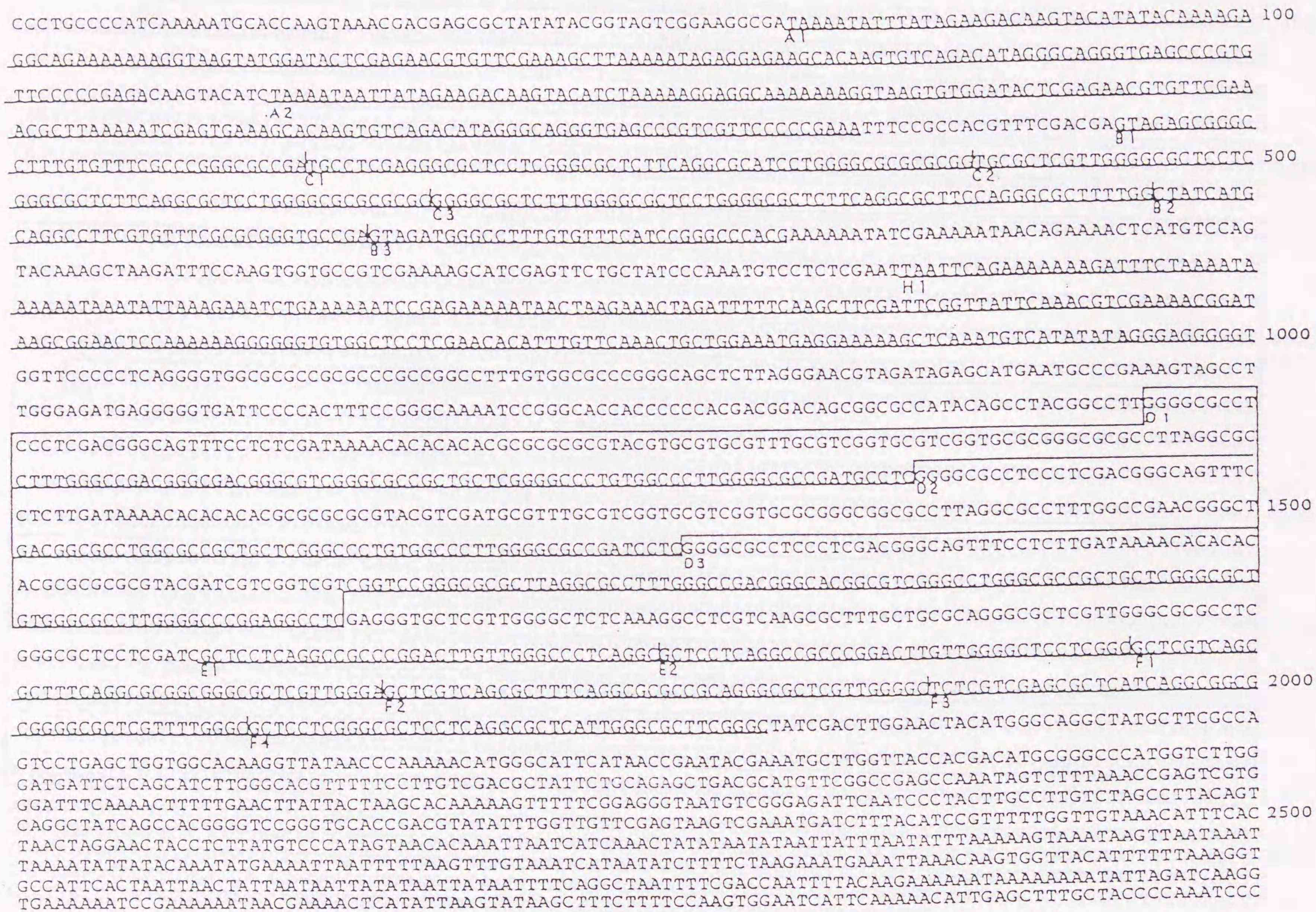

crude nuclear extracts
phosphocellulose
0.3M KCl
0.4M KCl
0.5M KCl
heparin
heparin
0.2M KCl
RBP-1
0.3M KCl
RBP-3
0.4M KCl
RBP-2

Figure 4



```
CCCTGCCCCATCAAAAATGCACCAAGTAAACGACGAGCGCTATATACGGTAGTCGGAAGGCGATAAAATATTTATAGAAGACAAGTACATATACAAAAGA 100
                                                               A1
GGCAGAAAAAAAAGGTAAGTATGGATACTCGAGAACGTGTTCGAAAGCTTAAAAATAGAGGAGAAGCACAAGTGTCAGACATAGGGCAGGGTGAGCCCGTG
TTCCCCCGAGACAAGTACATCTAAAATAATTATAGAAGACAAGTACATCTAAAAAGGAGGCAAAAAAAAGGTAAGTGTGGATACTCGAGAACGTGTTCGAA
                    .A2
ACGCTTAAAAATCGAGTGAAAGCACAAGTGTCAGACATAGGGCAGGGTGAGCCCGTCGTTCCCCCGAAATTTCCGCCACGTTTCGACGAGTAGAGCGGGC
                                                                                         B1
CTTTGTGTTTCGCCCGGGCGCCGATGCCTCGAGGGCGCTCCTCGGGCGCTCTTCAGGCGCATCCTGGGGCGCGCGCGGTGCGCTCGTTGGGGCGCTCCTC 500
                        C1                                                     C2
GGGCGCTCTTCAGGCGCTCCTGGGGCGCGCGCGCGGGGCGCTCTTTGGGGCGCTCCTGGGGCGCTCTTCAGGCGCTTCCAGGGCGCTTTTGGCTATCATG
                                  C3                                                          B2
CAGGCCTTCGTGTTTCGCGCGGGTGCCGAGTAGATGGGCCTTTGTGTTTCATCCGGGCCCACGAAAAAATATCGAAAAATAACAGAAAACTCATGTCCAG
                             B3
TACAAAGCTAAGATTTCCAAGTGGTGCCGTCGAAAAGCATCGAGTTCTGCTATCCCAAATGTCCTCTCGAATTAATTCAGAAAAAAAGATTTCTAAAATA
                                                                        H1
AAAAATAAATATTAAAGAAATCTGAAAAAATCCGAGAAAAAATAACTAAGAAACTAGATTTTTCAAGCTTCGATTCGGTTATTCAAACGTCGAAAACGGAT
AAGCGGAACTCCAAAAAAGGGGGTGTGGCTCCTCGAACACATTTGTTCAAACTGCTGGAAATGAGGAAAAAGCTCAAATGTCATATATAGGGAGGGGGT 1000
GGTTCCCCCTCAGGGGTGGCGCGCCGCGCCCGCCGGCCTTTGTGGCGCCCGGGCAGCTCTTAGGGAACGTAGATAGAGCATGAATGCCCGAAAGTAGCCT
TGGGAGATGAGGGGGTGATTCCCCACTTTCCGGGCAAAATCCGGGCACCACCCCCCACGACGGACAGCGGCGCCATACAGCCTACGGCCTTGGGGCGCCT
                                                                                            D1
CCCTCGACGGGCAGTTTCCTCTCGATAAAACACACACACGCGCGCGCGTACGTGCGTGCGTTTGCGTCGGTGCGTCGGTGCGCGGGCGCGCCTTAGGCGC
CTTTGGGCCGACGGGCGACGGGCGTCGGGCGCCGCTGCTCGGGGCCCTGTGGCCCTTGGGGCGCCGATGCCTCGGGCGCCTCCCTCGACGGGCAGTTTC
                                                                         D2
CTCTTGATAAAACACACACACGCGCGCGCGTACGTCGATGCGTTTGCGTCGGTGCGTCGGTGCGCGGGCGGCGCCTTAGGCGCCTTTGGCCGAACGGGCT 1500
GACGGCGCCTGGCGCCGCTGCTCGGGCCCTGTGGCCCTTGGGGCGCCGATCCTCGGGGCGCCTCCCTCGACGGGCAGTTTCCTCTTGATAAAACACACAC
                                                       D3
ACGCGCGCGCGTACGATCGTCGGTCGTCGGTCCGGGCGCGCTTAGGCGCCTTTGGGCCGACGGGCACGGCGTCGGGCCTGGGCGCCGCTGCTCGGGCGCT
GTGGGCGCCTTGGGGCCCGGAGGCCTCGAGGGTGCTCGTTGGGGCTCTCAAAGGCCTCGTCAAGCGCTTTGCTGCGCAGGGCGCTCGTTGGGCGCGCCTC
GGGCGCTCCTCGATCGCTCCTCAGGCCGCCCGGACTTGTTGGGCCCTCAGGCGCTCCTCAGGCCGCCCGGACTTGTTGGGGCTCCTCGGCGCTCGTCAGC
               E1                                    E2                                     F1
GCTTTCAGGCGCGGCGGGCGCTCGTTGGGAGCTCGTCAGCGCTTTCAGGCGCGCCGCAGGGCGCTCGTTGGGGCTCTCGTCGAGCGCTCATCAGGCGGCG 2000
                               F2                                           F3
CGGGGCGCTCGTTTTGGGCGCTCCTCGGGCGCTCCTCAGGCGCTCATTGGGCGCTTCGGGCTATCGACTTGGAACTACATGGGCAGGCTATGCTTCGCCA
                   F4
GTCCTGAGCTGGTGGCACAAGGTTATAACCCAAAAACATGGGCATTCATAACCGAATACGAAATGCTTGGTTACCACGACATGCGGGCCCATGGTCTTGG
GATGATTCTCAGCATCTTGGGCACGTTTTTCCTTGTCGACGCTATTCTGCAGAGCTGACGCATGTTCGGCCGAGCCAAATAGTCTTTAAACCGAGTCGTG
GGATTTCAAAACTTTTTGAACTTATTACTAAGCACAAAAAGTTTTTCGGAGGGTAATGTCGGGAGATTCAATCCCTACTTGCCTTGTCTAGCCTTACAGT
CAGGCTATCAGCCACGGGGTCCGGGTGCACCGACGTATATTTGGTTGTTCGAGTAAGTCGAAATGATCTTTACATCCGTTTTTGGTTGTAAACATTTCAC 2500
TAACTAGGAACTACCTCTTATGTCCCATAGTAACACAAATTAATCATCAAACTATATAATATAATTATTTAATATTTAAAAAGTAAATAAGTTAATAAAT
TAAAATATTATACAAATATGAAAATTAATTTTTAAGTTTGTAAATCATAATATCTTTTCTAAGAAATGAAATTAAACAAGTGGTTACATTTTTTAAAGGT
GCCATTCACTAATTAACTATTAATAATTATATAATTATAATTTTCAGGCTAATTTTCGACCAATTTTACAAGAAAAAATAAAAAAAATATTAGATCAAGG
TGAAAAAATCCGAAAAAATAACGAAAACTCATATTAAGTATAAGCTTTCTTTTCCAAGTGGAATCATTCAAAAACATTGAGCTTTGCTACCCCAAATCCC
```

Figure 5A

```
CTCTCCCAATAATTCAAAAAAAAAGATTTATAAAATACAAAATAAATATTAAAAAAATCTGAAAAAAATCCGAAAAAAAATAACTAAGAAACTAGATTTTTC 3000
          H2
AAGCTTCGATTCGGTTATTCAAACGTCGAAAACGGATAAGCGGAACTCCAAAAAAGGGGGTGTGGCTCCTCGAACACATTTGTTCAAACTGCTGGAAAT
GAGGAAAAAGCTCAAATGTCATATATAGGGAGGGGGTTGTCGCCCCTCCCTGGGCAGTGGGGTTGGGCGCTCGGTAGGCGCGCGCGGGGCCGTCGTTGGG
CGCTCGATGGGGCGCTCTTCGGGCGCTCTTCGGGCGCTTCTCGGGCGCTTTGGGCTATCATCCTGGAACTATGCCGGCATTCTATGCTTGGCCAGTCCTG
AGTTGGTGGCATAAGGTTATAACCCAAAAACTGCTCTCCTGGAGGCGGGCCTGTCGACTCTGGAGTCCTTAGATGACACCAGTGCTCGGCGCCGTCTTGG
TAGCTTTGAGTCTTTGATCTCATTGCTTGGGGAAAATTGACAGCGTTGGTGAGGATCCAGGTCACTTCACCAACGGGACCCTAGGGAACGGATTAGAAGG 3500
GAGGGGTTCTCCACCCTCCCTGTGGGCAGACGGTATTCGGGCGCCTCGGGCTCTCTTCGGGATCCTGGTTGAGCACTCCACTTGGATGCTTGGCCAGTCT
TGAAGGGGGCATGTTCCAACTAAAAACTGCCCTCCCCCATGGCTGAGGAGCCCTTGGGCTCTTTTGCTTGTGGGGAACAACGTTGGTGTGATTAGGATCA
                                       G1
GACGCACCAACGGTGACCTGGTAGGAGTGATCCGTTTGTTGGAGGGGTCGGGCTCCCCCCTCCAAAAACAACGTTGGTGTAAAATTGGCCATAGCACCAA
CGGAGACCGTCCCGGAGCGGCCAAGCGGGAGGGGAAACTTCCCCCCTCCCGCTGGGAAGATTTTTAAGGGACCAGATGGTTGTGTTCCCAAAGACATTCA
TGTGGGGTGCTTGGCGTAGACTTGATGGTATTTTGCCCAAAGACACTCCTATGGGGTGCTCGGCCTAGACCTGATGGTTCGTGCTAGGACATTCTGCGGG 4000
GTGCTCGGAATAGACCTGAAGAGAACTTGAAATCTTCCCTCCCCCCTGGCGGGGCTGCCAACTCTGGTGTCCTCGGATGATGCCGGTGCTTGGCGCTGCT
AGGGTCGCTGAGGAGTCCTCGGGCTCTTTGGCGTTTGGGAAAAACGTTGGCGTGAATAGGGTCAGACACGTCGACGTTGACCGTGAAGGAGAAGTGATCC
    G2
ATTCGTATGGAGGGGACCCCGGCTCCCCCCTCCCGAAAACAACGTTGGTGTAAATGGCCATAGCACCAACGGAGACCGTCCCGGAGCGGCCAAGTGGGAG
GGAAACTTCCCCCCTCCCGCTGGGAAGATTTTTAAGAGACCTGATGGTTGTGTTCCCAAAGATATTCATGTGGGGTGCCTGGCGTAGACTTGATGGTATT
TTGCCCAAAGACACTCCTATGGGGTGCTCGGCCTAGACCTGATGGTTCGTGCTTGGACATTCTGCGGGGTGCTTGGAATAGACCTGATGAGAACTTGAAA 4500
TCTTCCCTCCCCTGGCGGGGCTGCCAACTCTGGTGTCCTCGGAGTACGCCGGTGCTTGGCGCTGCTAGGGCGCTTAGGAGTCCTCGGGCTCTTTGGCGT
                                                                       G3
TTGGGAAAAACGTTGGCGTGAATAGGGTCAGACACGTGGACGTTGACCGTGAAGGAGAAGTGATCCATTCGTATGGAGGGGACCCCGGCTCCCCCCTCCC
GAAAACAACGTTGGTGTAAATGGCCATAGCACCAACGGAGACCGTCCCGGAGCGGCCAAGTGGGAGGGGAAACTTCCGCCCTCCGCTGGAAGATTTTTAA
GAGACCTGATGGTTGTGTTCCCAAAGATATTCATGTGGGGTGCCTGGCGTAGACTTGATGGTATTTTGCCCAAAGACACTCCTATGGGGTGCTCGGCCT
AGACCTGATGGCTCGTGCTTGGACATTCTGCGGGGTGCTTGGAATAGACCTGATGAGAACTTGAAATCTTCCCTCCCCTGGCGGGGCTGCCAACTCTGGT 5000
GTCCTCGGATGACGCCGGTGCTTGGCGCTGCTAGGGTCGCTGAGGAGTCCTCGGGCTCTTTGGCGTTTGGGAAAAACGTTGGCGTGAATAGGGTCAGACA
                                    G4
CGTTGACGTTGACCGTGAAGGAGAAGTGATCCATTCGTATGGAGGGGACCCCGGCTCCCCCCTCCCGAAAACAACGTTGGTGTAAATGGCCATAGCACCA
ACGGAGACCGTCCCGGAGTGGCCAAATTTGCGCATAGCGTTCACGTTGCTTACACTACTGTGGGTTGATTTCTTTTAGACCTTTGGGTTTTGAAGGGTTC
AGCCTAATGTATGTGGGTTTCGTGTTGTACGTGGAGCGCATGGGTTGTGATTTGTATTTCCCGGTTGGTTGGATCCCTGTCTCGGCAGCGACAATCGCTT
GGTCTTACATGTTCATAAGTCAGAGAGTTGTGAAAGATGCCCTCAAGATGAGATTGGTGGACCCTGCGTAATATACCTATCGTGATGGAATTTCGATGTC 5500
CCATAGTCCCTAATTCGTCTTGTGCATTTTATTTTGTGCAAGACTAACAATGAATCTAACATGGTCTTCCGCCCTTTCATTTCCTTTTGGTCTTGTAAGT
GTCGGGAGTCCATGTGTTGAGGGAAATCTTGAATGCGGAACTGGTAATCGATTATGGACTTGGTCCATACTTTTCCGAGATATATGTCCCCTACGGAAGT
CCGTTGCGCTTTGTGTTGAGTAATGGGTGCTAATGTATCTTTTGCTCATGCTTAGCCTTCGGCCTGAAGGAATGC
```

Figure 5B

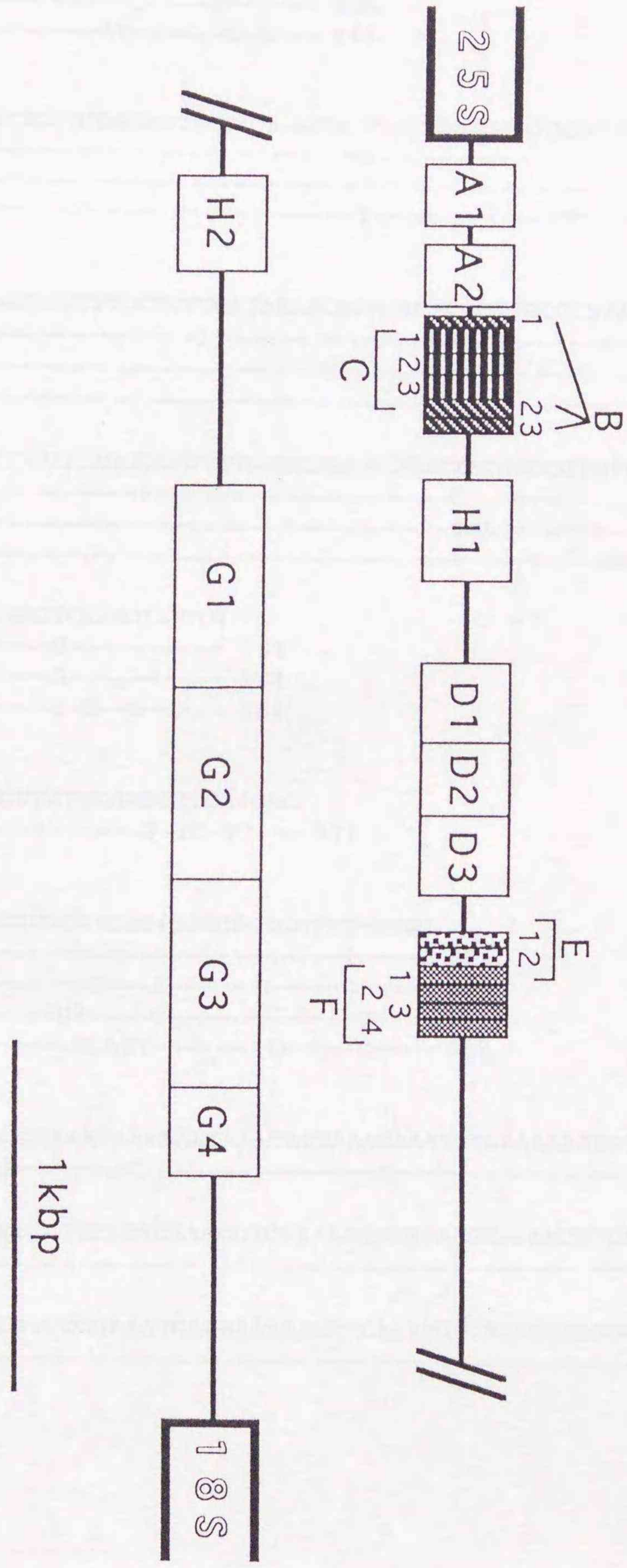


Figure 6

```
Repeat-A
A1   TAAAATATTTATAGAAGACAAGTACATATACAAAAGAGGCAGAAAAAAAAGGTAAGTATGGATACTCGAGAACGTGTTCGA
A2   -------A-------------------C--A---G-----++--------------++-----------------------

     ATAAGA+GCTTAAAAATAGAGGAGAAGCACAAGTGTCAGACATAGGGCAGGTGAGCCCGT+GTTCCCCCGAGA
     ------C----------C---TGA------------------------------------C------------ 91%

Repeat-B
C.S. GTAG+ATGCXGGCCTTTGTGTTTCGCCCGGGCGC+CGA
B1   -------G-G---------------------------- 95%
B2   C--TC----A------C----------G----T----- 82%
B3   --------++--------------AT------C-A--- 84%

Repeat-C
C.S. ++TGCGCTCGTTGGGGCGCTCCTCGGGCGCTCTTCAGGCGCXTCCTGGGGCGCGCGCGG
C1   -----+----++A----------------------------A----------------- 92%
C2   -----------------------------------------+----------------- 98%
C3   GGG------T-------------G-----------------T---A---+---TTTT-- 80%

Repeat-D
C.S. GGGGCGCCTCCCTCGACGGGCAGTTTCCTCTTGATAAAACACACACACGCGCGCGTACGTGXATGCGTTTGCGTCGG
D1   ------------------------------C------------------------------CG--------------
D2   ------------------------------------------------------------CG---------------
D3   -----------------------------------------------------------+++--+++++++------

     TGCGTCGGTGCGCGGGCG+CGCCTTAGGCGCCTTTGGGCCGA+CGGGCXACGGGCGTC++GCCTGGGCGCCGCTGCTCG
     ------------------------------------------------G-----------++++---------------
     ------------------G------------------+----A-----TGAC---+++--------+------------
     -+-------++C----------+-------------------------+----+----GG-------------------

     GGCCCTGTGGGC+CCTTGGGGCTCCGATGCCTC
     --G--CTGT-------------G---------- 93%
     ----------+-----------G-----+---- 92%
     ---G--------G---------C-G--G----- 88%

Repeat-E
E1   GCTCCTCAGGCCGCCCGGACTTGTTGGGCCCTCA+GGC
E2   ----------------------------G-TC-TC--- 87%

Repeat-F
C.S. GCTCGTCXAGCGCTX+TCAGGCGCG+CGC+GGGCGCTCGTTT+GGGC
F1   -------+------T-----------+--------------+----A 89%
F2   -------+------T----------C---A-----------G----- 89%
F3   T------G------CA-------G+----G------------T---- 83%
F4   ----C--GG-----CC--------TCATT-------++++--C---- 68%

Repeat-H
H1   TAATTCAGAAAAAAAGATTTCTAAAATAAAAAATAAATATTAAAGAAATCTGAAAAAATCCGAGAAAAATAACTAAGAAA
H2   -------A------------A-------C---------------A------------------A----------------

     CTAGATTTTTCAAGCTTCGATTCGGTTATTCAAACGTCGAAAACGGATAAGCGGAACTCCAAAAAAGGGGGGTGTGGCTC
     --------------------------------------------------------------------------------

     CTCGAACACATTTGTTCAAACTGCTGGAAATGAGGAAAAAGCTCAAATGTCATATATAGGGAGGGGGT
     -------------------------------------------------------------------- 98%
```

Figure 7

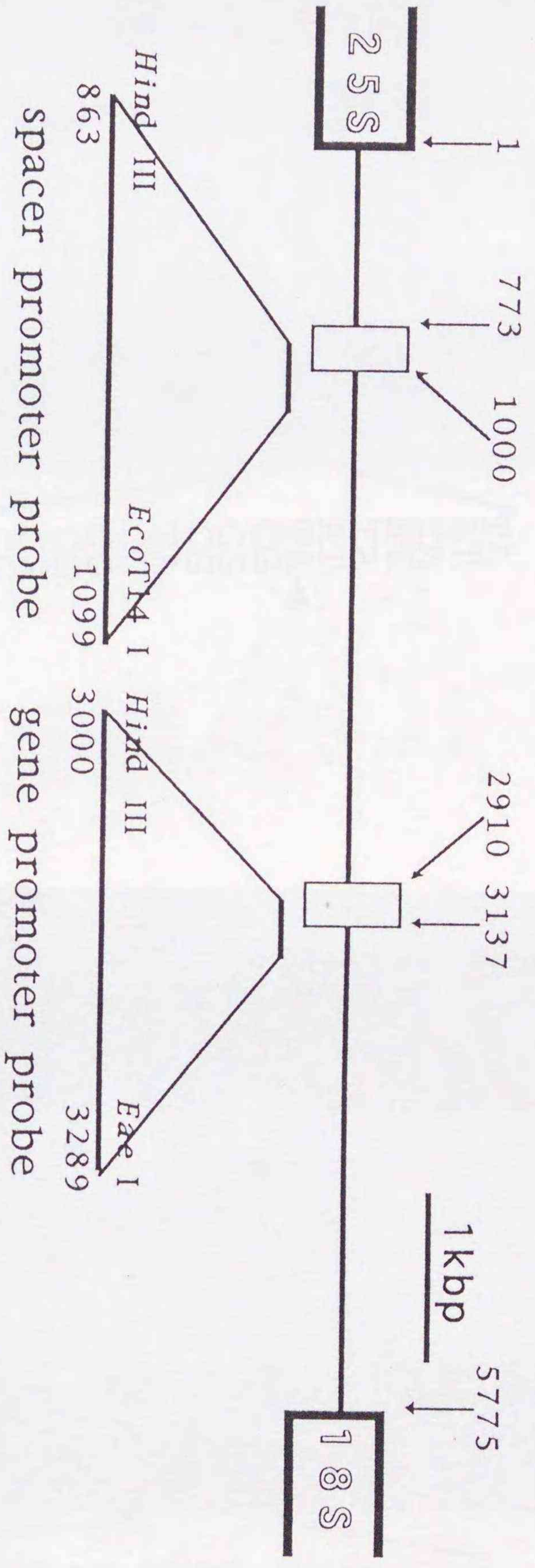
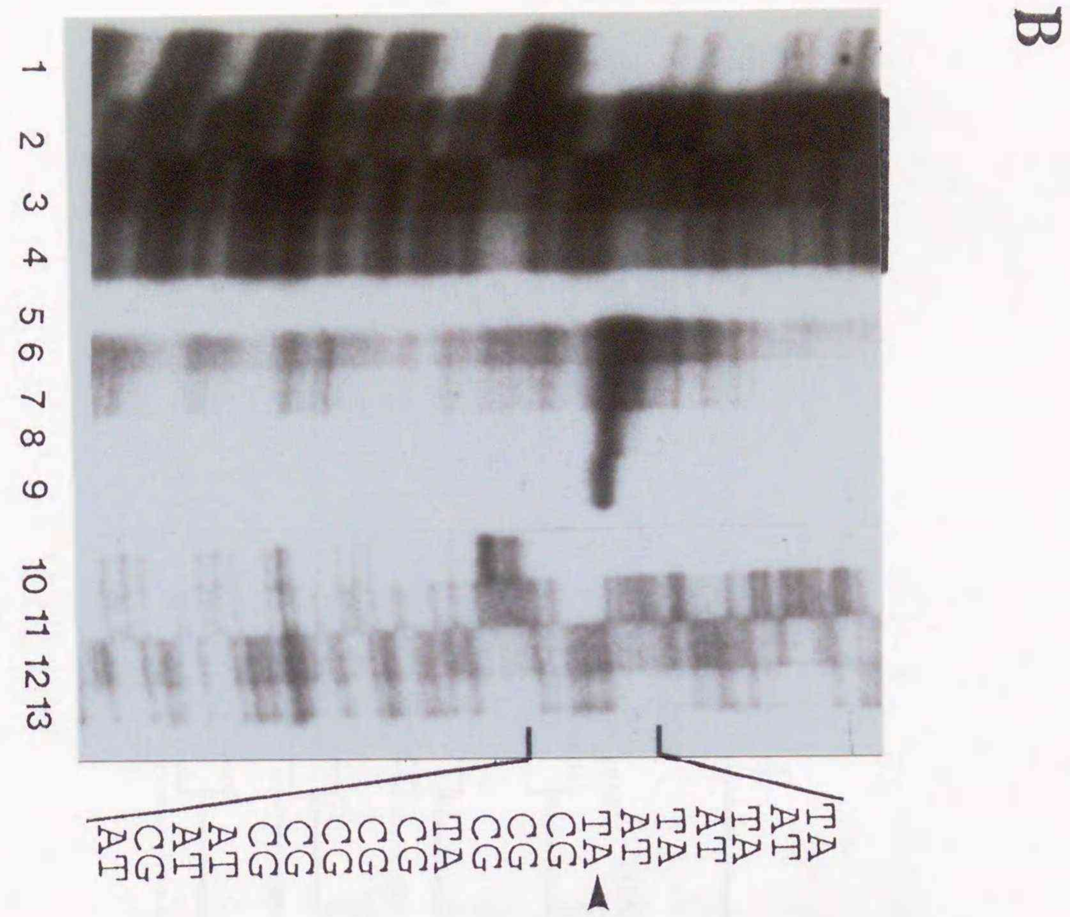

25S
1
773
1000
2910
3137
1kbp
5775
18S
Hind III
863
EcoT14 I
1099
Hind III
3000
Eae I
3289
spacer promoter probe
gene promoter probe

Figure 8A


B
1
2
3
4
5
6
7
8
9
10
11
12
13
TA
AT
TA
AT
TA
AT
TA
CG
CG
CG
TA
CG
CG
CG
CG
AT
AT
CG
AT

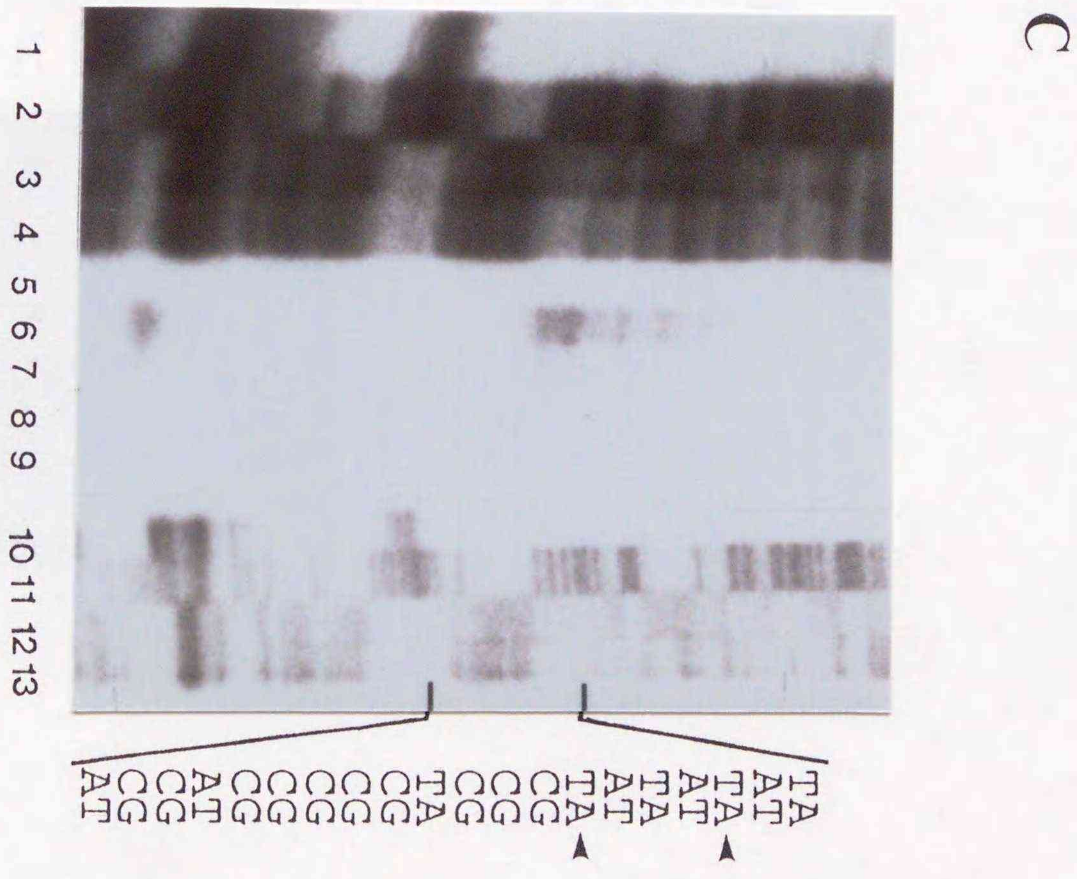

C
1
2
3
4
5
6
7
8
9
10
11
12
13
TA
AT
TA
AT
TA
AT
TA
CG
CG
CG
TA
CG
CG
CG
CG
AT
CG
CG
AT

Figure 8B, C

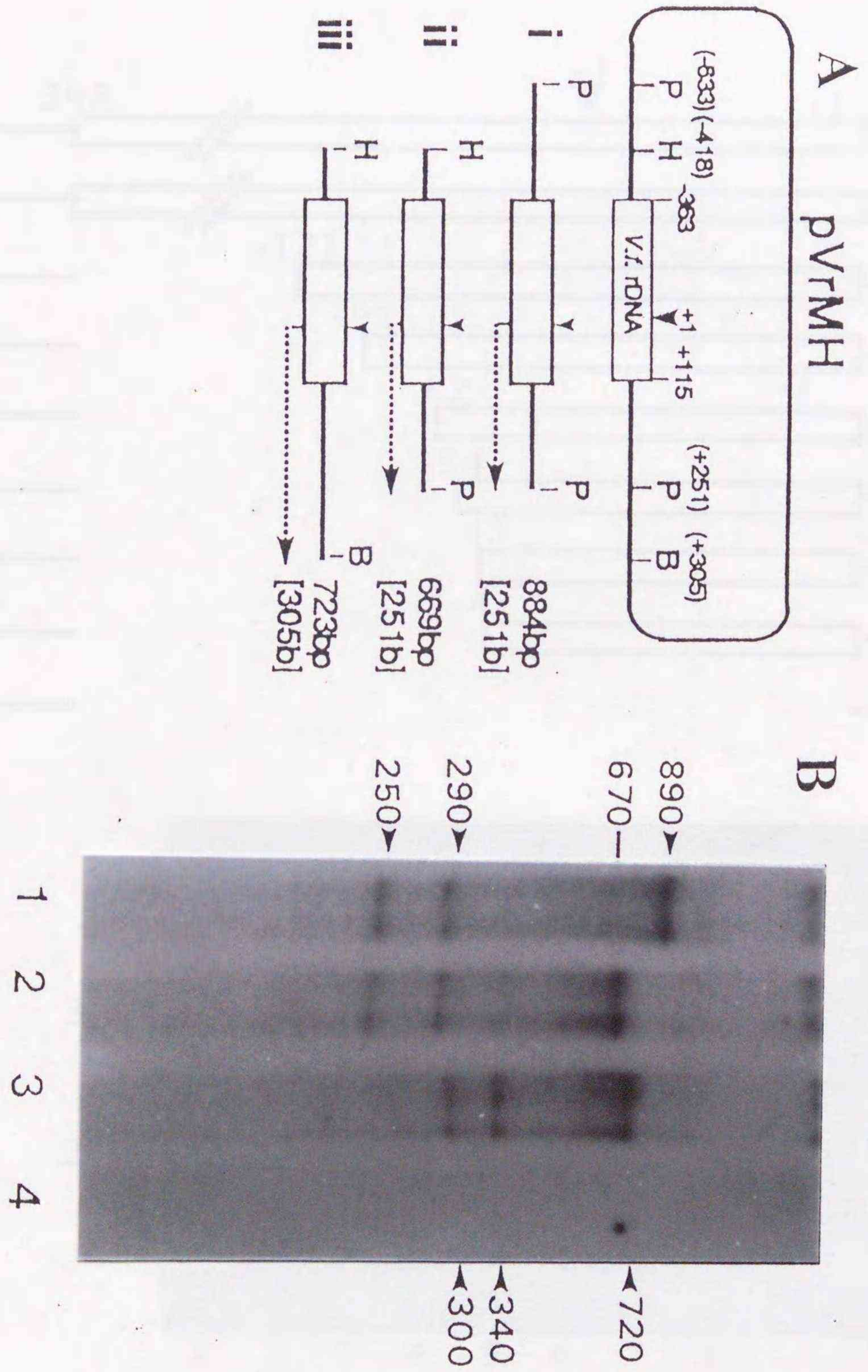

A
pVrMH
(-833)(-418)
P
H
-363
V.r. rDNA
+1
+115
(+251)
(+305)
P
B
i
ii
iii
P
H
H
P
P
B
884bp [251b]
669bp [251b]
723bp [305b]
B
890
670
290
250
1
2
3
4
720
340
300

Figure 9

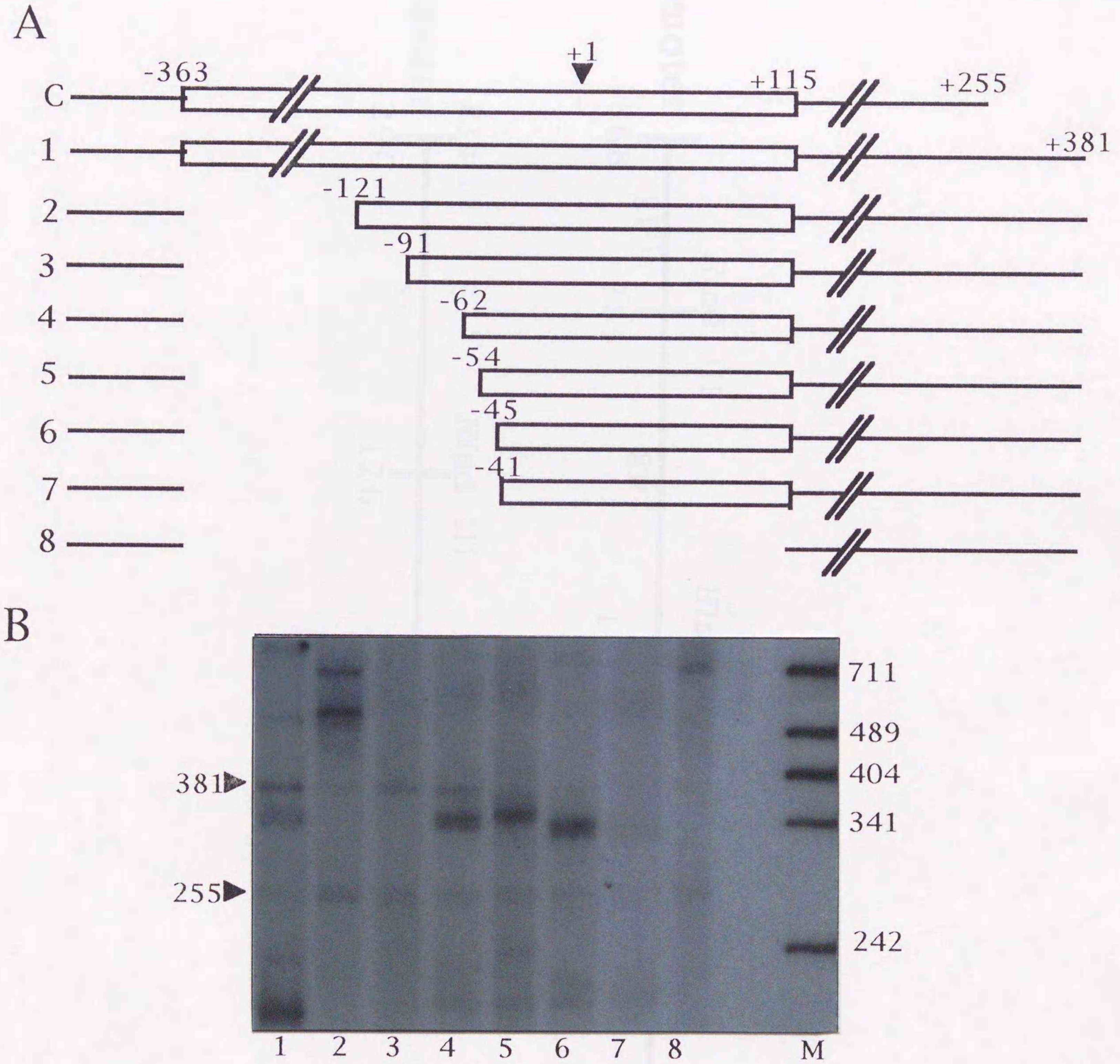

A
+1
-363
+115
+255
C
1
+381
2
-121
3
-91
4
-62
5
-54
6
-45
7
-41
8
B
711
489
381
404
341
255
242
1
2
3
4
5
6
7
8
M

Figure 10

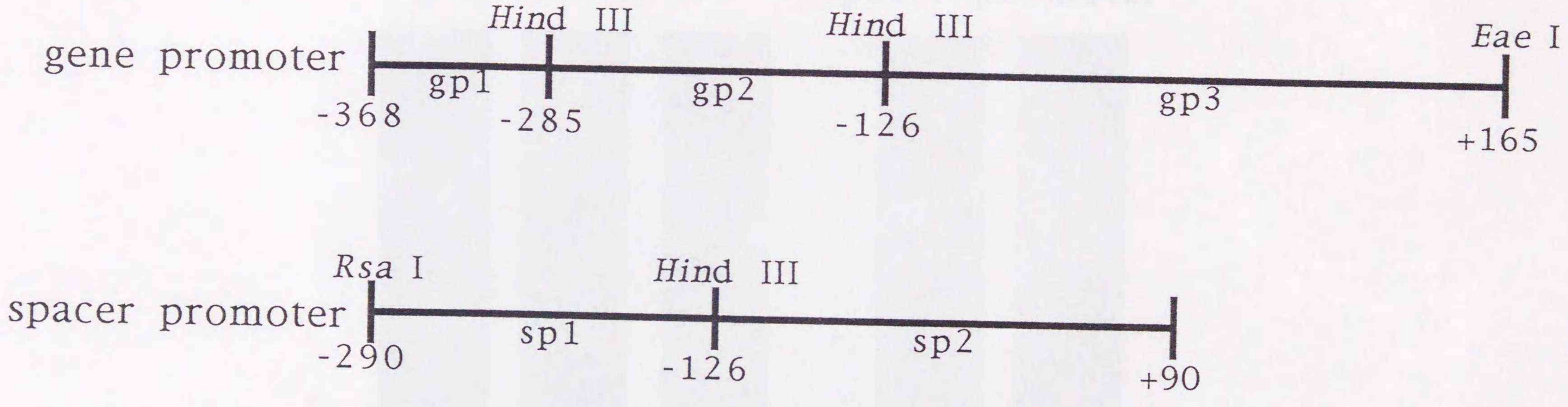

gene promoter
Hind III
Hind III
Eae I
gp1
gp2
gp3
-368
-285
-126
+165
spacer promoter
Rsa I
Hind III
sp1
sp2
-290
-126
+90

Figure 11A

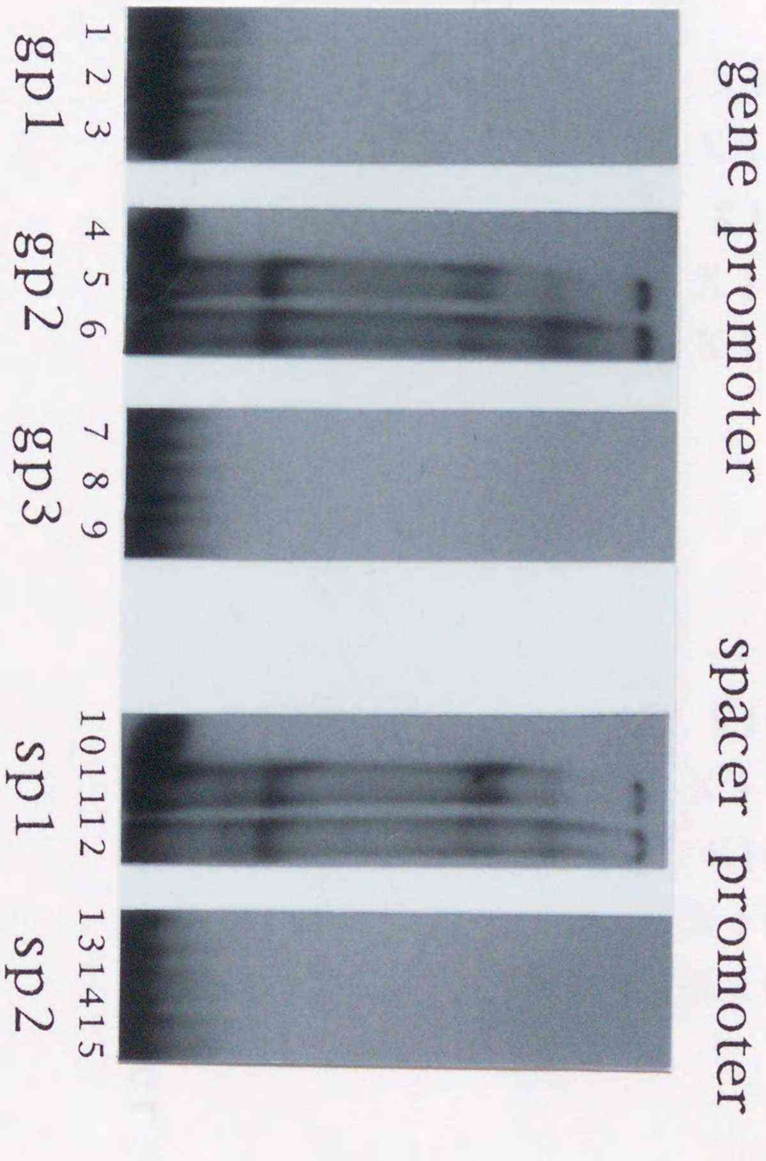

gene promoter
spacer promoter
1 2 3
4 5 6
7 8 9
10 11 12
13 14 15
gp1
gp2
gp3
sp1
sp2

Figure 11B

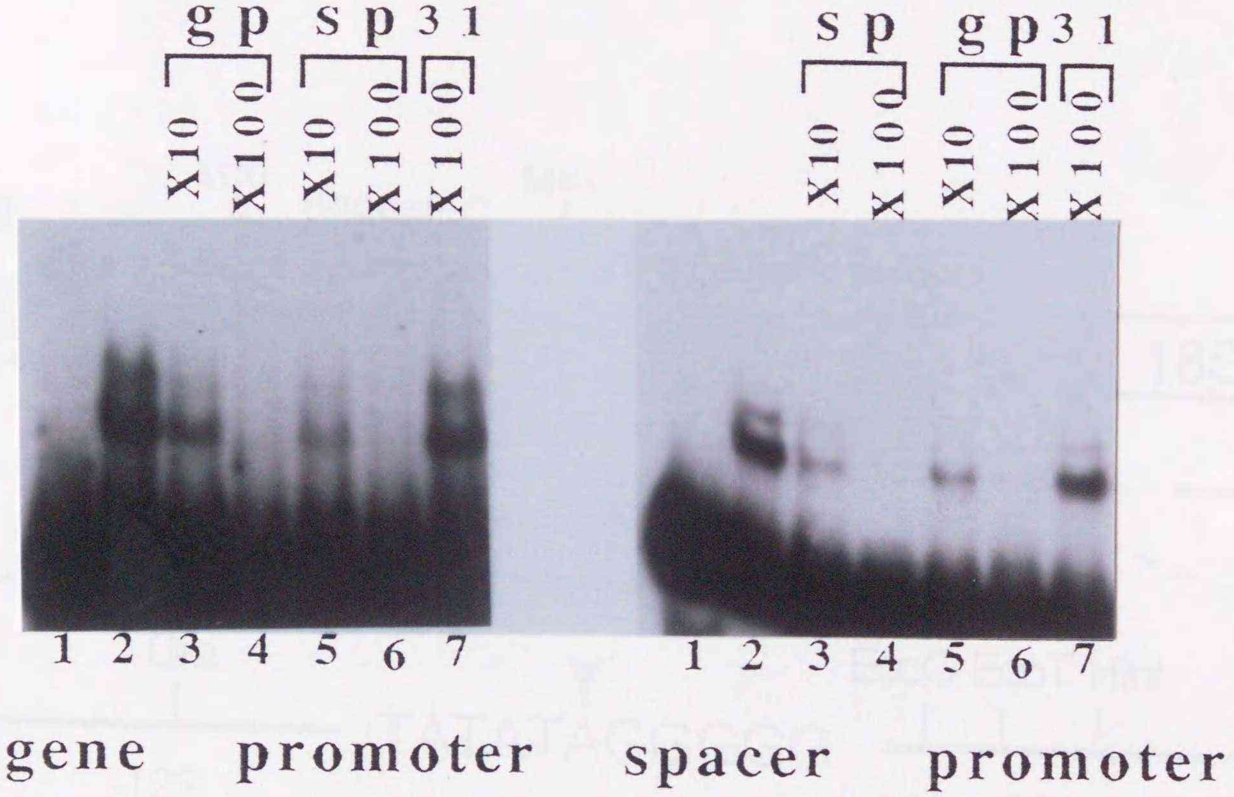

gp
X10
X100
sp
X10
X100
31
X100
1 2 3 4 5 6 7
gene promoter
sp
X10
X100
gp
X10
X100
31
X100
1 2 3 4 5 6 7
spacer promoter

Figure 12

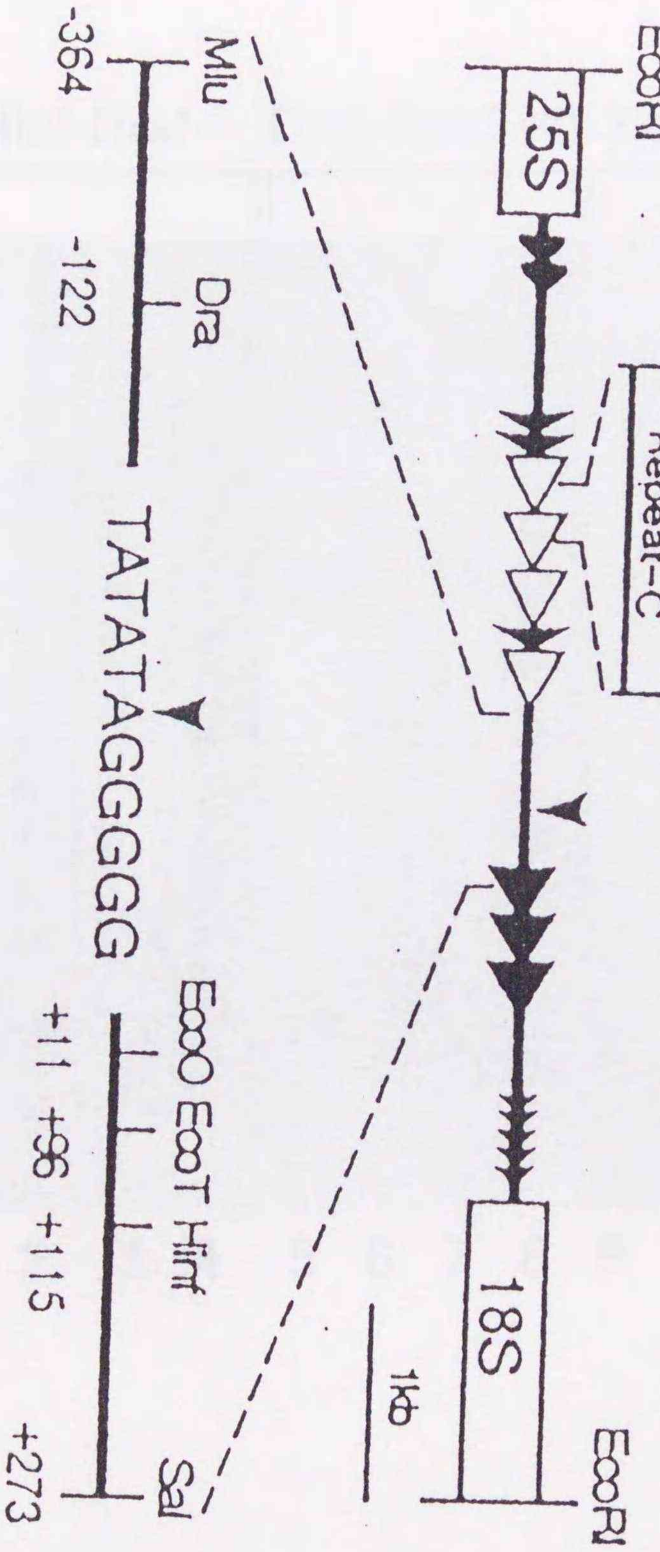

EcoRI
25S
Mlu
Repeat-C
Mlu
18S
EcoRI
1kb
Mlu
-364
Dra
-122
TATATAGGGG
EcoO
EcoT
Hinf
+11
+96
+115
Sal
+273

Figure 13

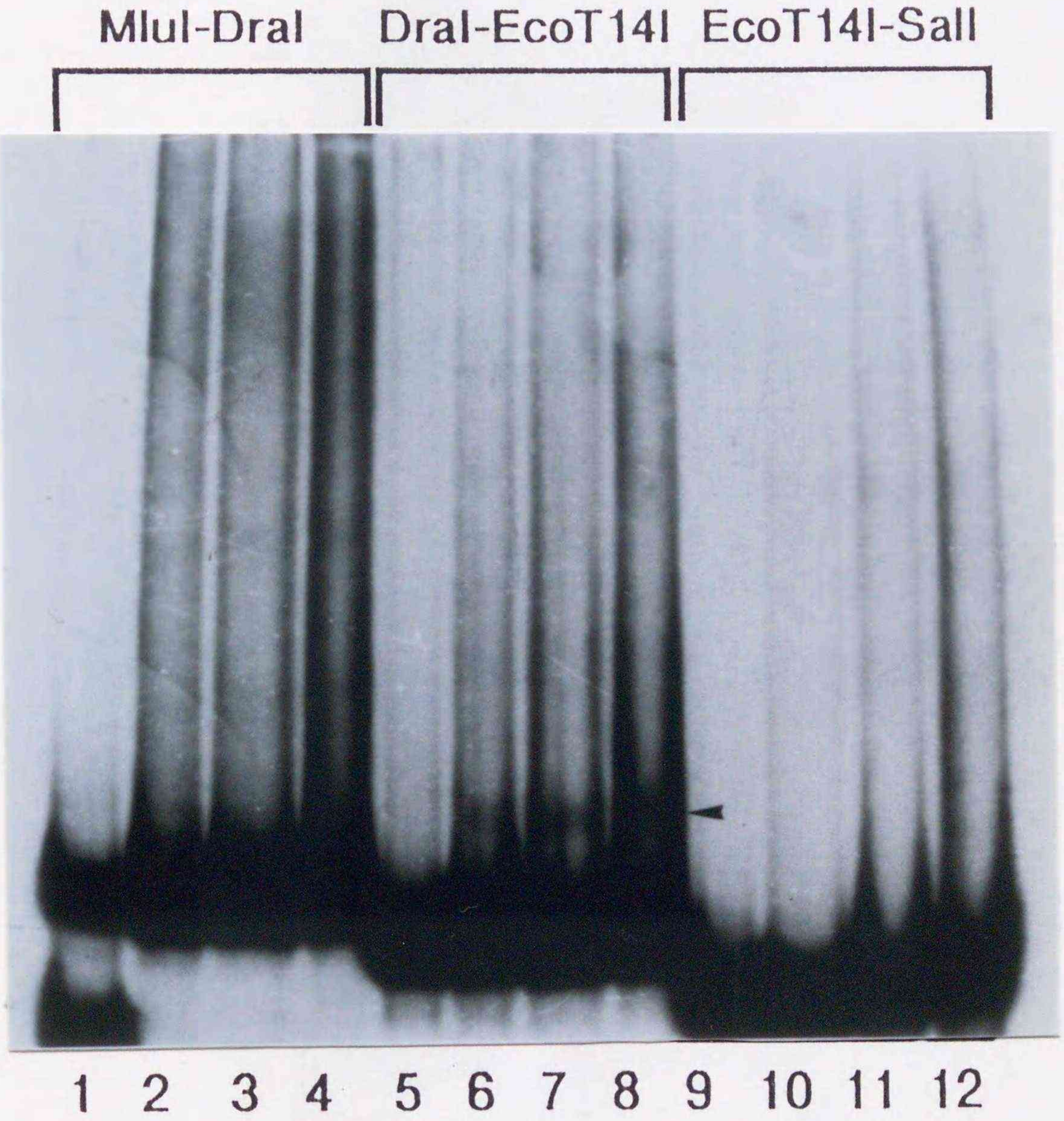


Figure 14

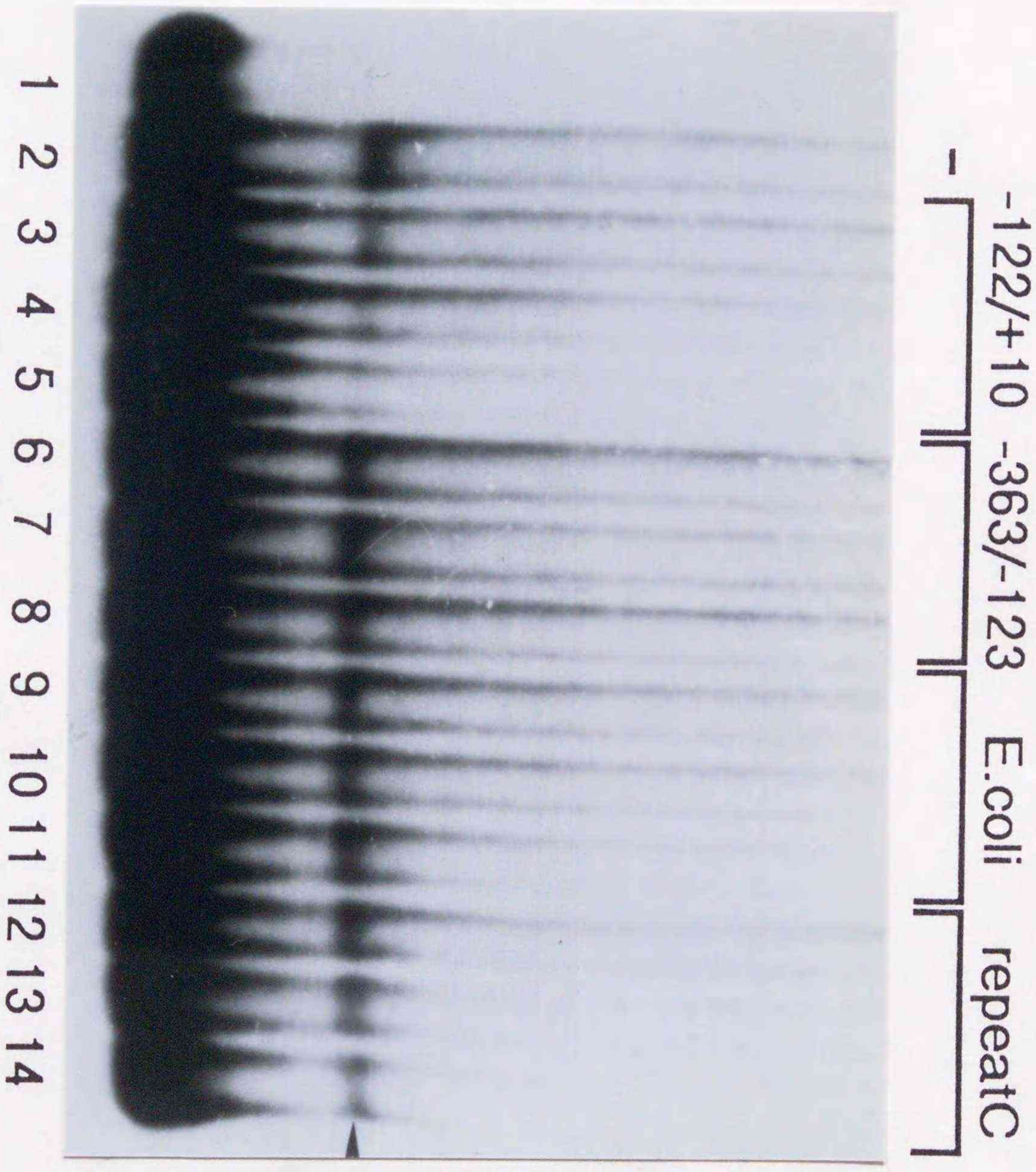


Figure 15

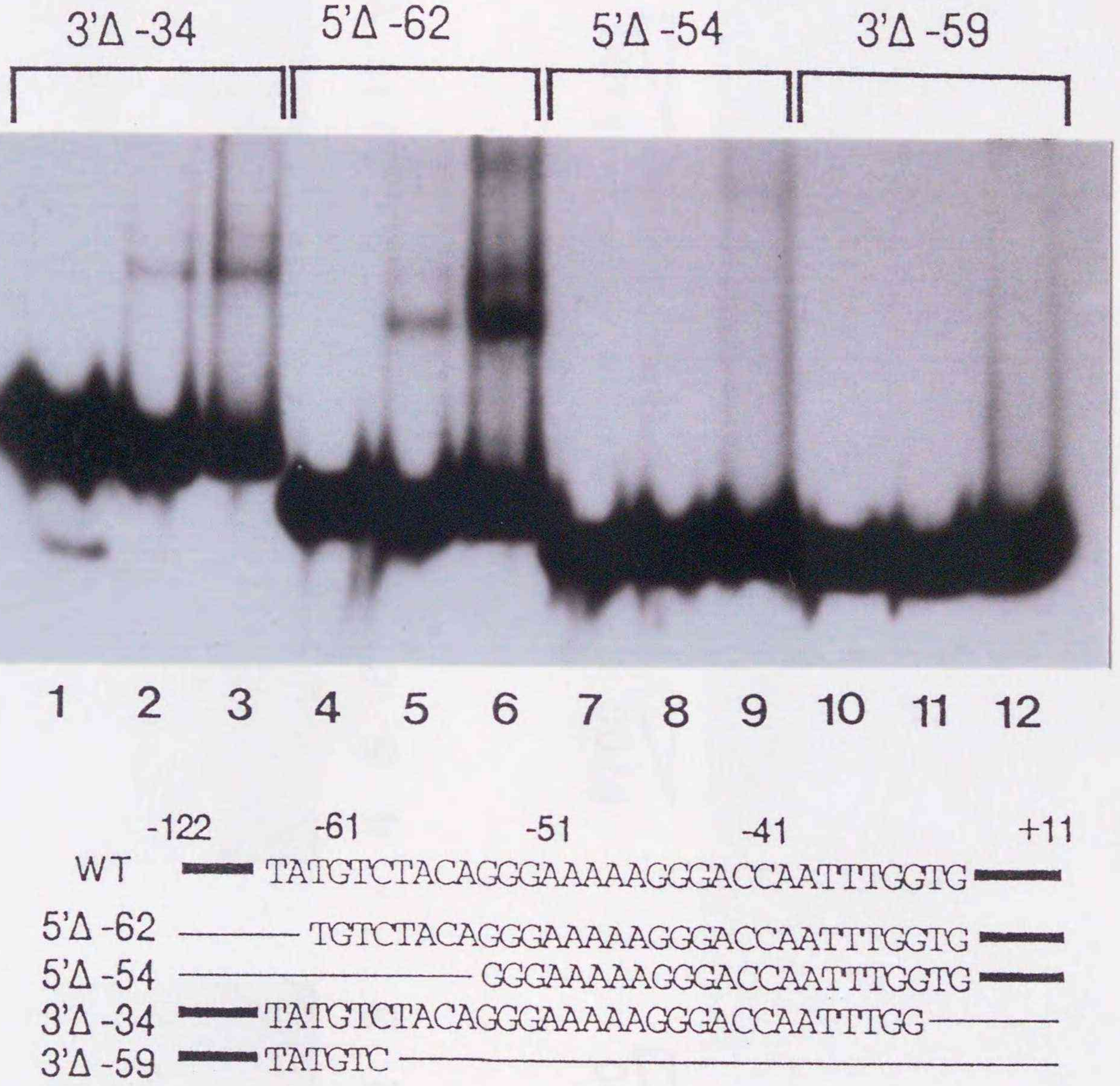

3'Δ -34
5'Δ -62
5'Δ -54
3'Δ -59
1 2 3 4 5 6 7 8 9 10 11 12
-122 -61 -51 -41 +11
WT TATGTCTACAGGGAAAAAGGGACCAATTTGGTG
5'Δ -62 TGTCTACAGGGAAAAAGGGACCAATTTGGTG
5'Δ -54 GGGAAAAAGGGACCAATTTGGTG
3'Δ -34 TATGTCTACAGGGAAAAAGGGACCAATTTGG
3'Δ -59 TATGTC

Figure 16

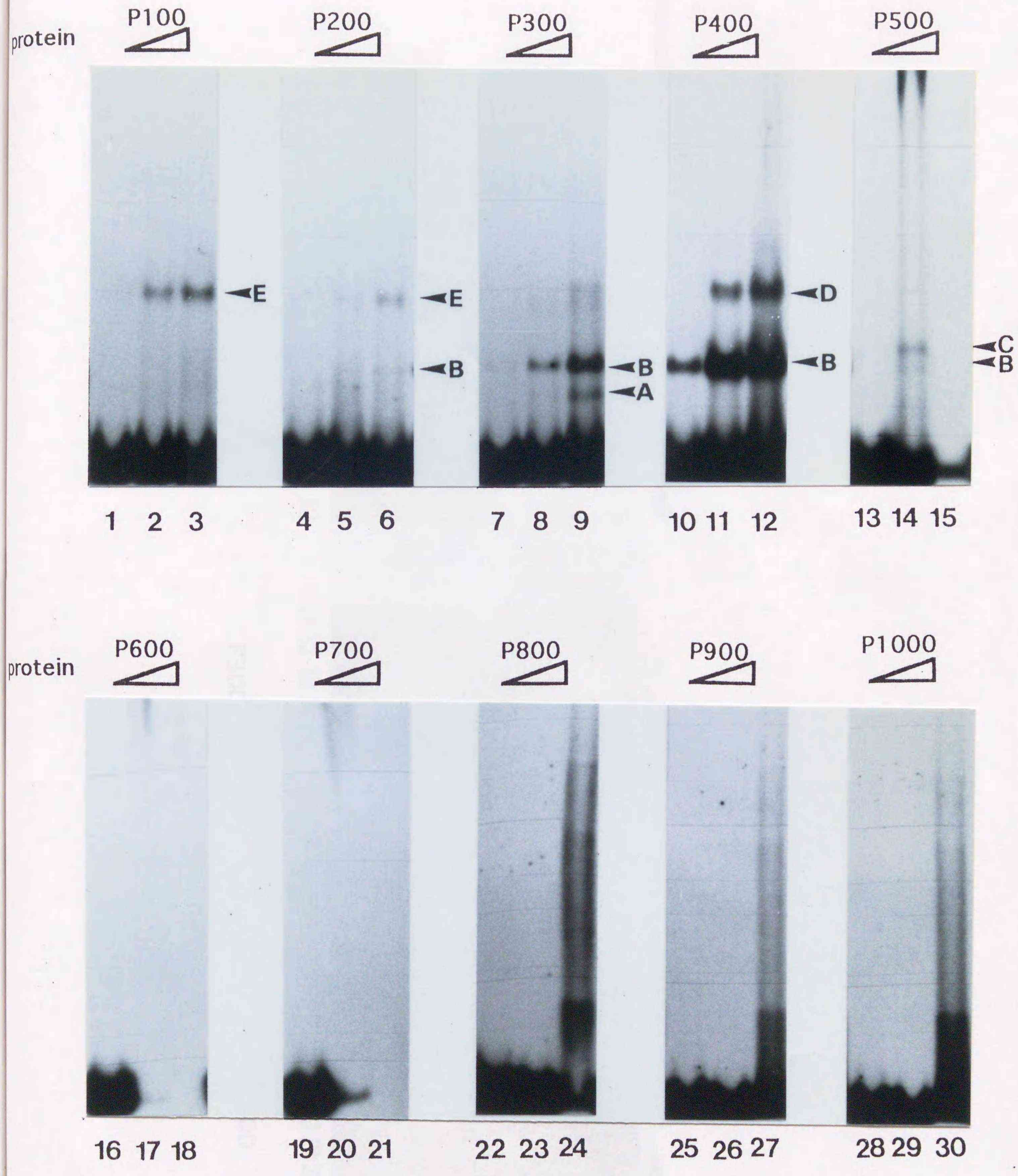

protein
P100
P200
P300
P400
P500
E
E
B
B
A
D
B
C
B
1 2 3
4 5 6
7 8 9
10 11 12
13 14 15
protein
P600
P700
P800
P900
P1000
16 17 18
19 20 21
22 23 24
25 26 27
28 29 30

Figure 17

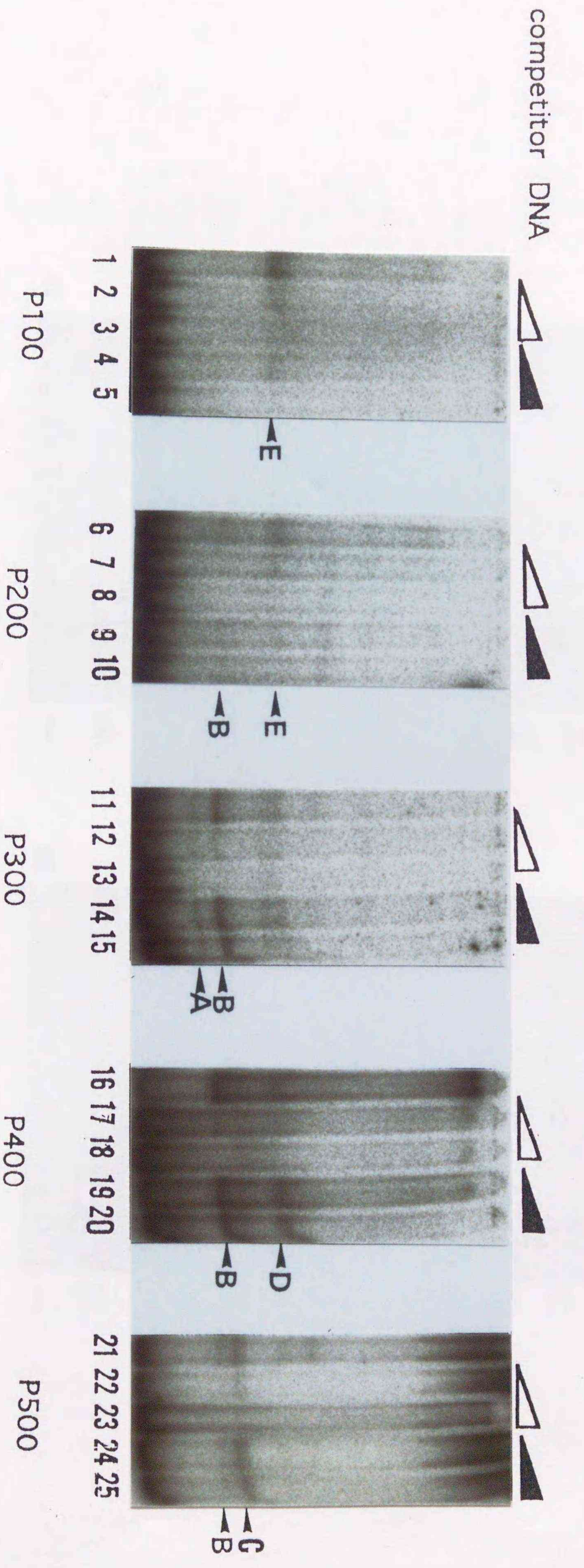

competitor DNA
1 2 3 4 5
P100
E
6 7 8 9 10
P200
B
E
11 12 13 14 15
P300
A
B
16 17 18 19 20
P400
B
D
21 22 23 24 25
P500
B
C

Figure 18

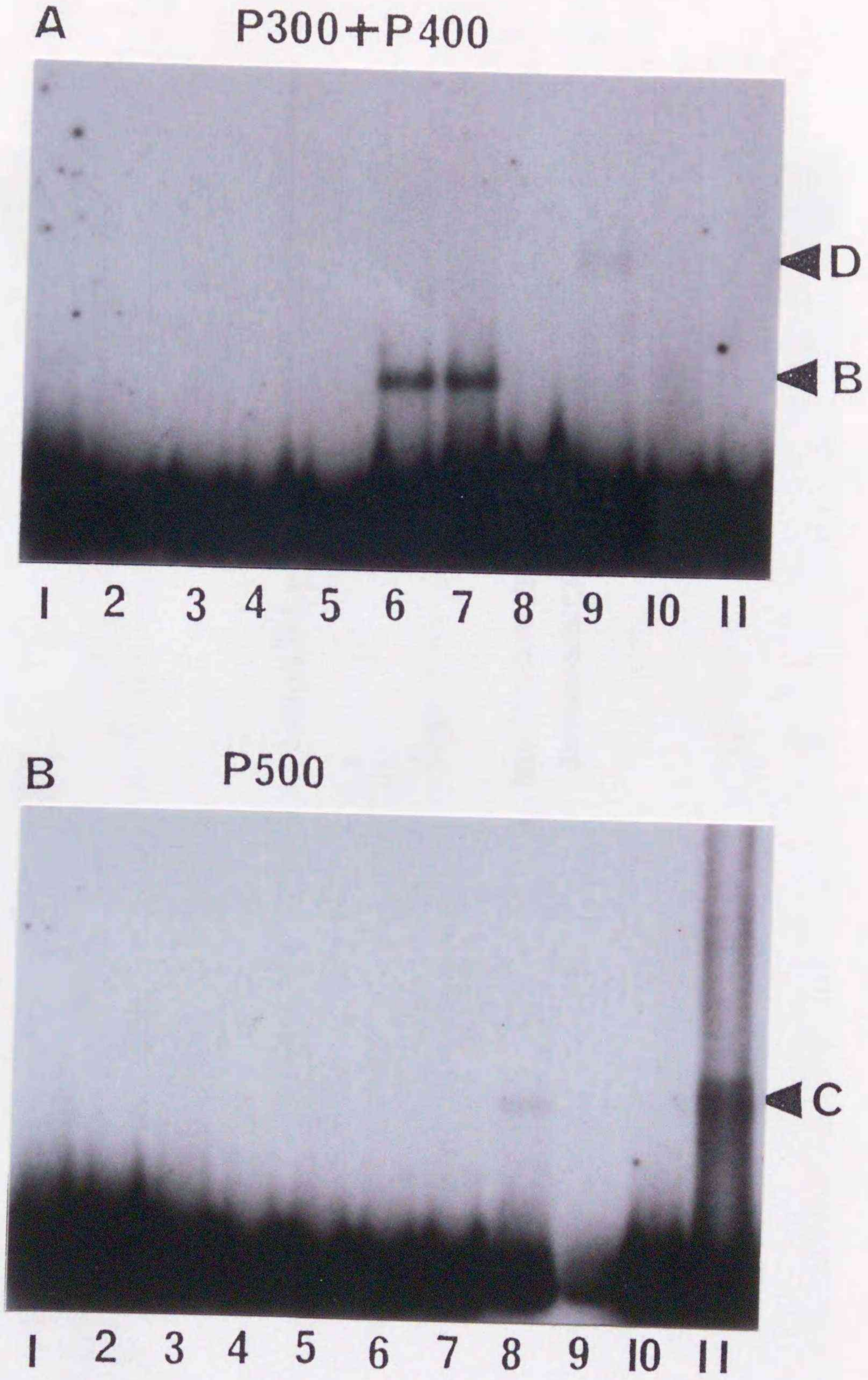


Figure 19

DNaseI Footprinting Assay (Top Strand)

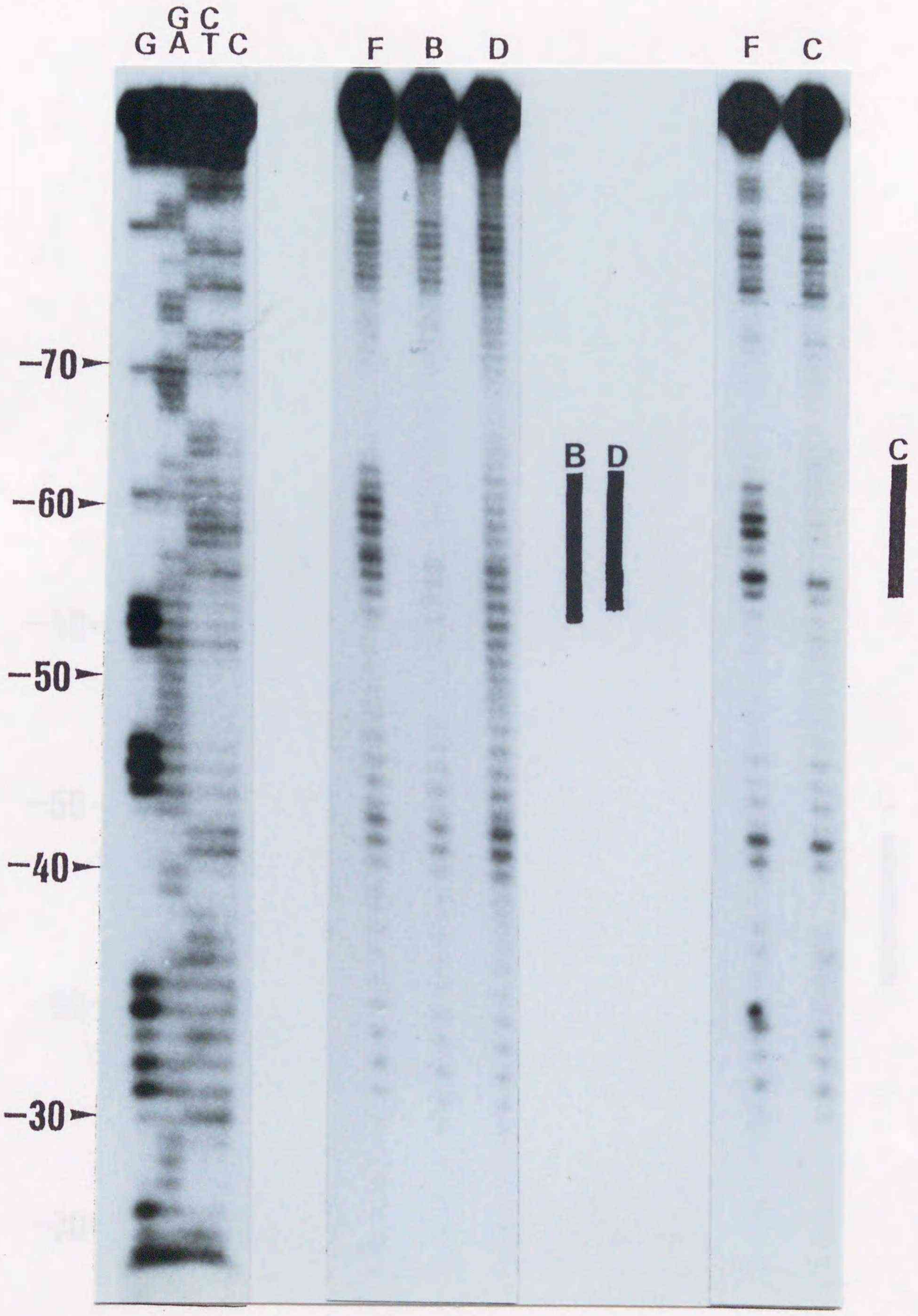


Figure 20A

DNaseI Footprinting Assay (Bottom Strand)

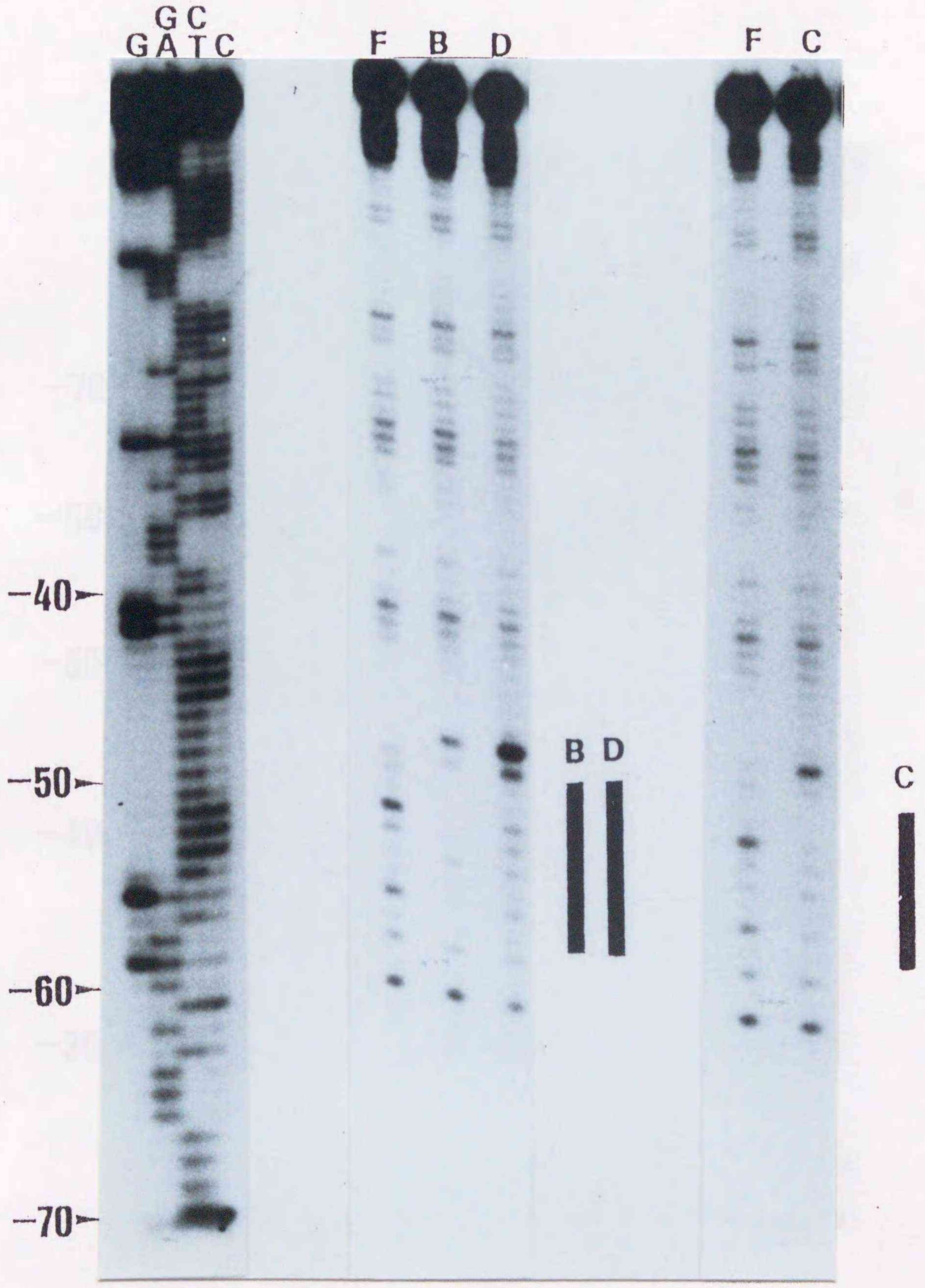


Figure 20B

A top strand

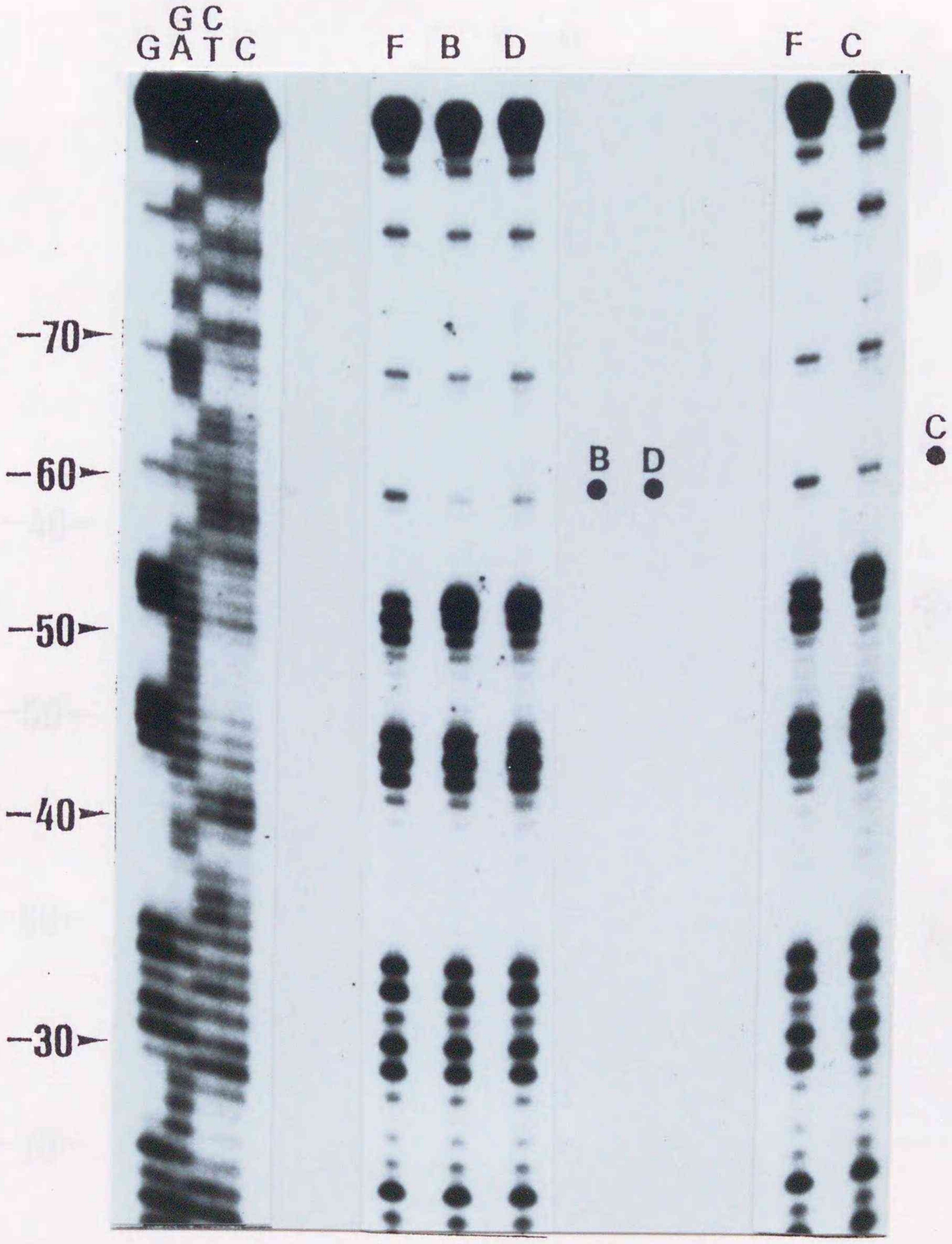


Figure 21A

B bottom strand

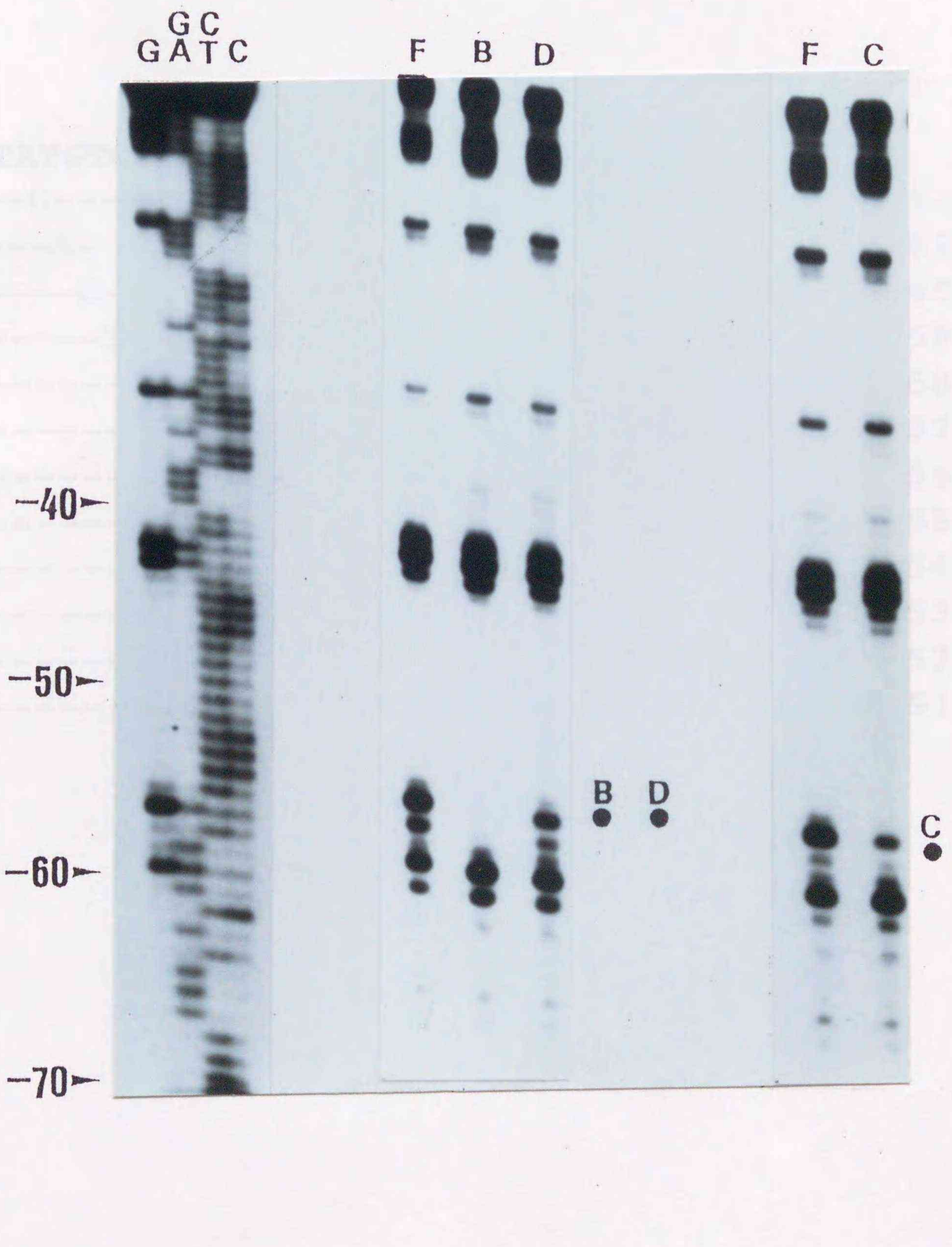


Figure 21B

A

```
1    TATGTCTACAGGGAAAAAGGGACCAATTTGGTGGC  WT
2    --C--------------------------------  M-62
3    ---A-------------------------------  M-61
4    ----C------------------------------  M-60
5    -----T-----------------------------  M-59
6    ------C----------------------------  M-58
7    -------G---------------------------  M-57
8    --------T--------------------------  M-56
9    ---------G-------------------------  M-55
10   ----------A------------------------  M-54
11   -----------A-----------------------  M-53
12   ------------A----------------------  M-52
13   -------------GGGGG-----------------  M-51/-47
```

Figure 22A

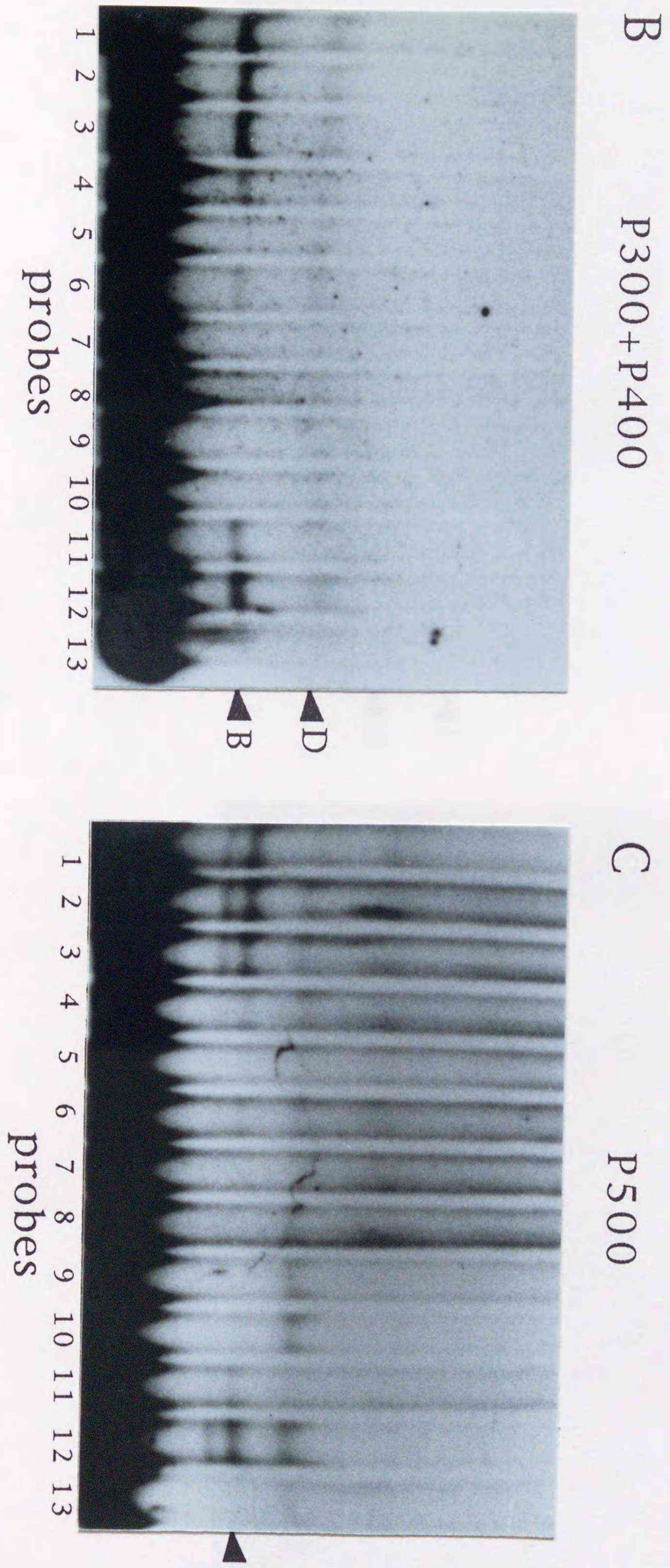

B
P300+P400
1 2 3 4 5 6 7 8 9 10 11 12 13
probes
B
D
C
P500
1 2 3 4 5 6 7 8 9 10 11 12 13
probes
C

Figure 22B, C

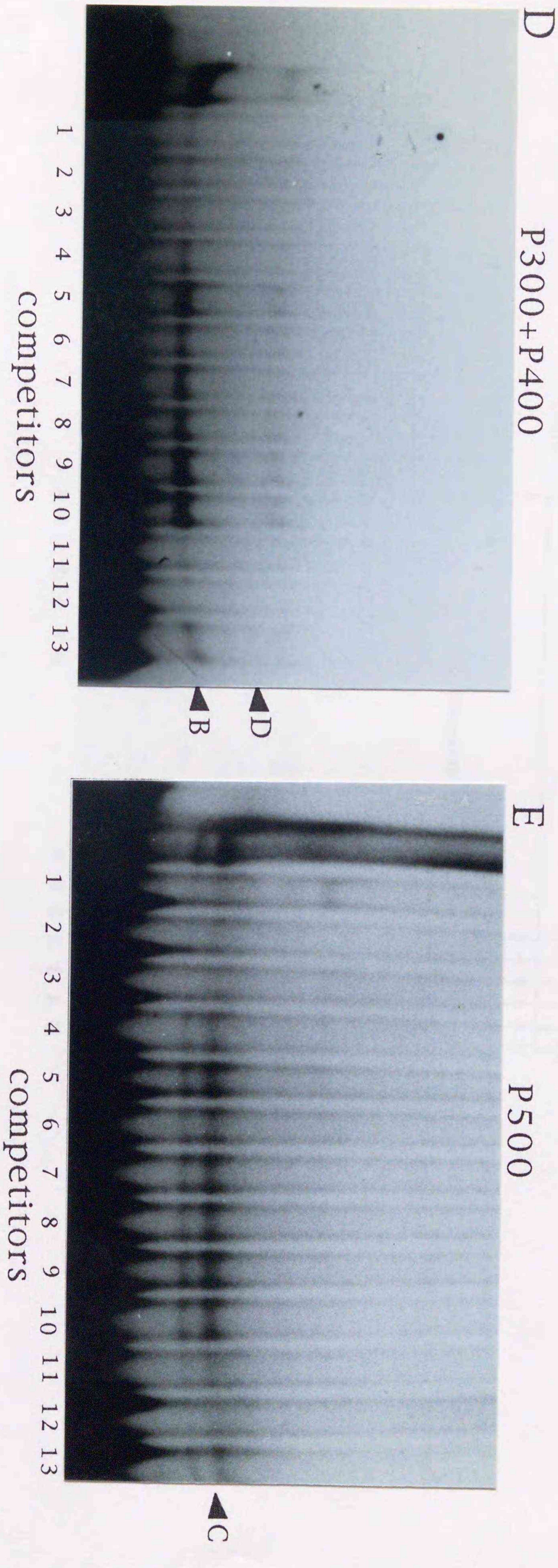

D
P300+P400
1 2 3 4 5 6 7 8 9 10 11 12 13
competitors
B
D
E
P500
1 2 3 4 5 6 7 8 9 10 11 12 13
competitors
C

Figure 22D, E

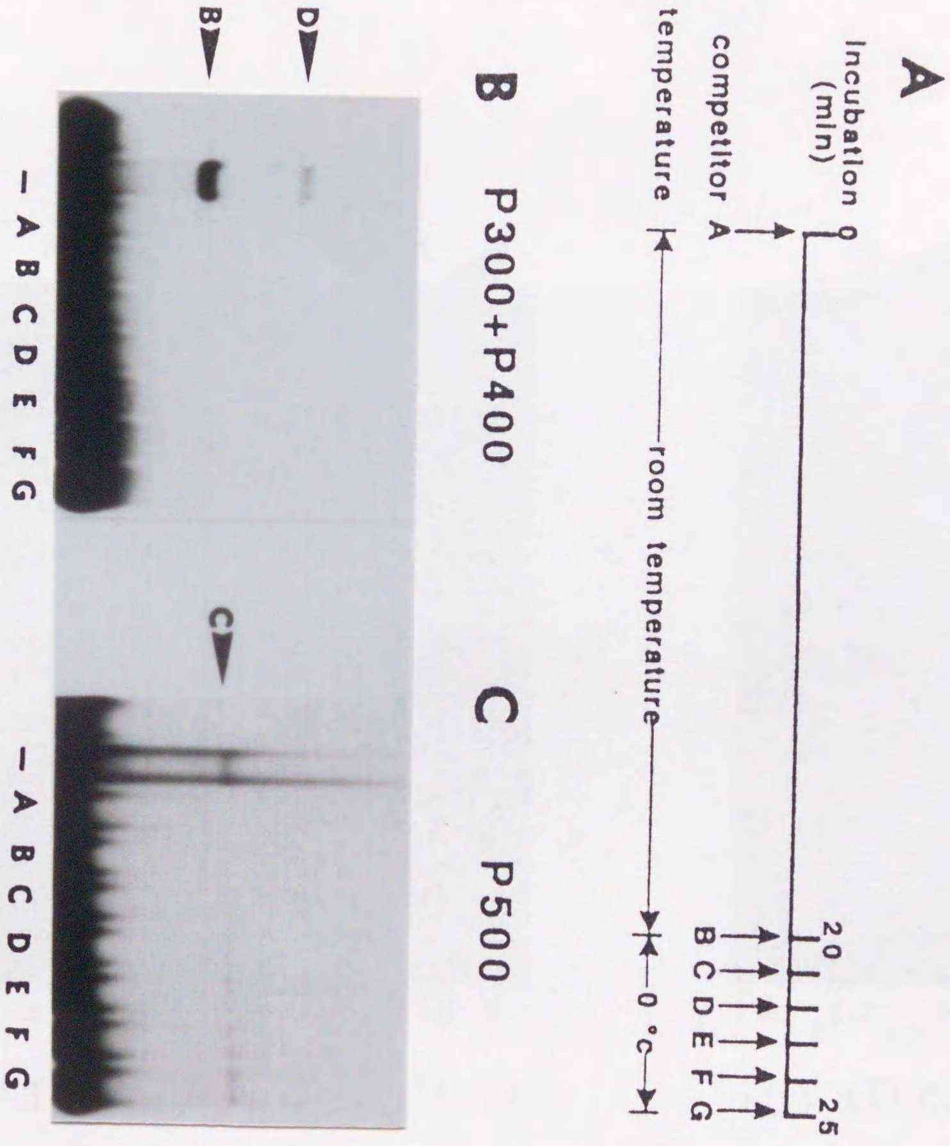

A
Incubation (min)
0
20
25
competitor A
B C D E F G
temperature
room temperature
0 °C
B
P300+P400
C
P500
D
B
C
– A B C D E F G
– A B C D E F G

Figure 23

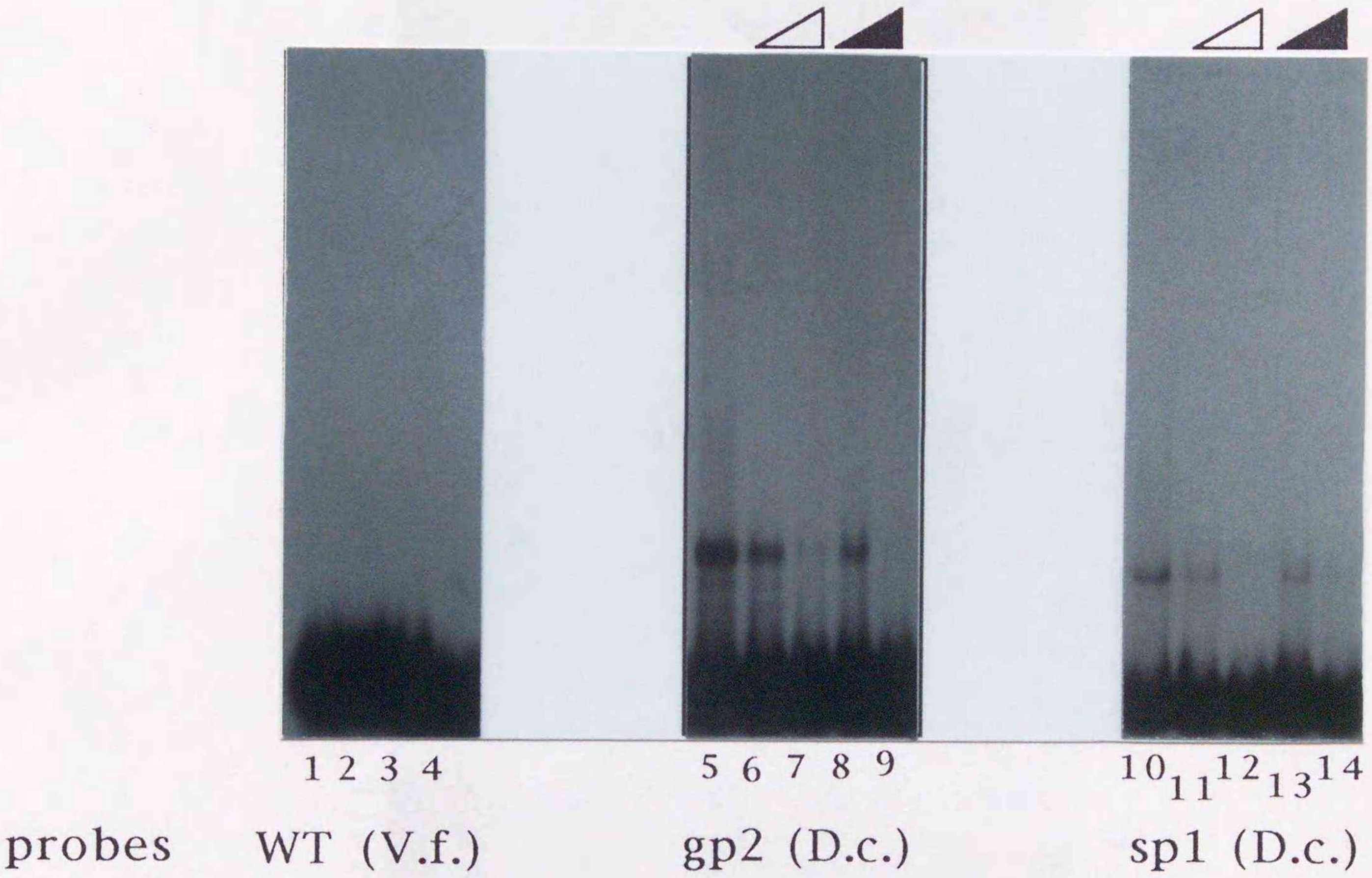


Figure 24

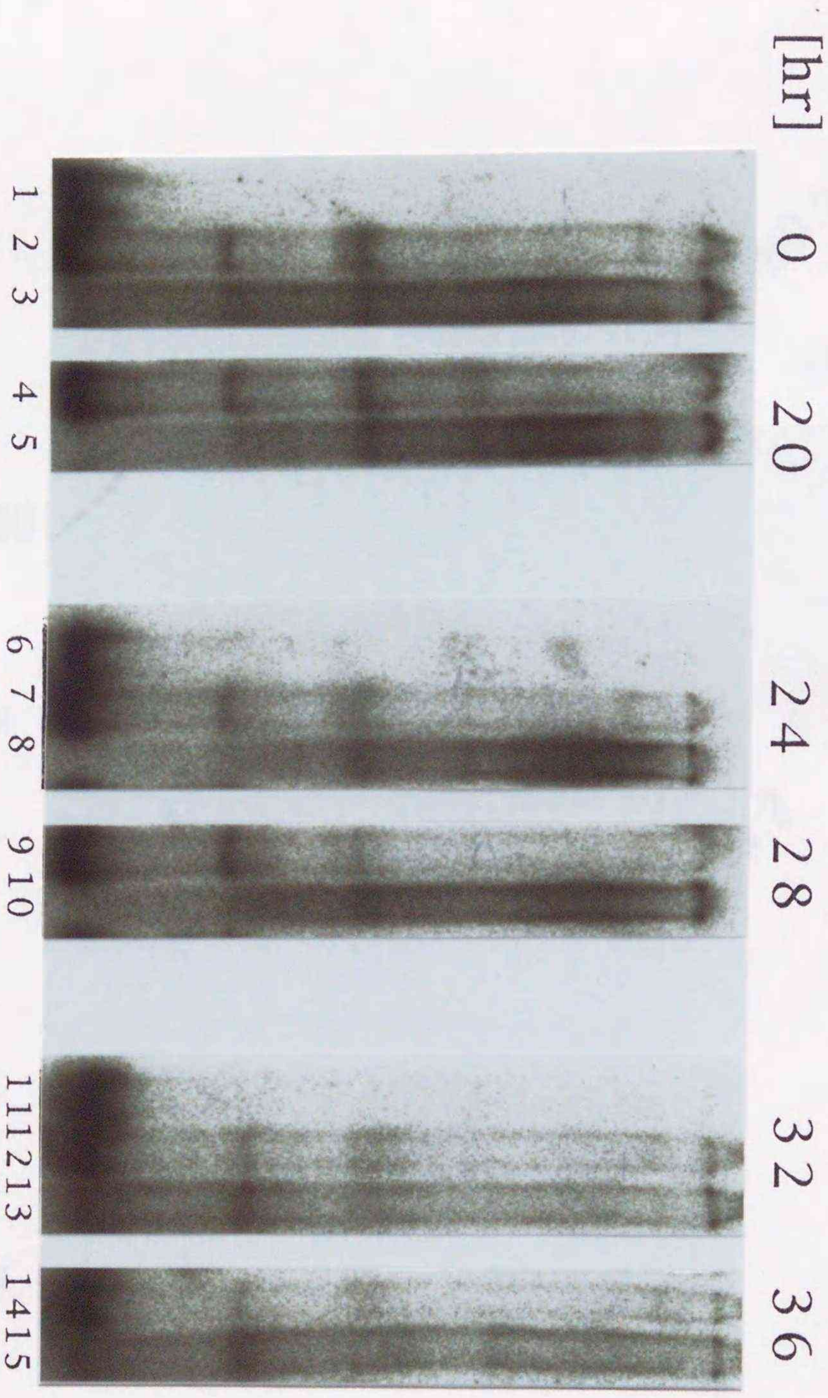
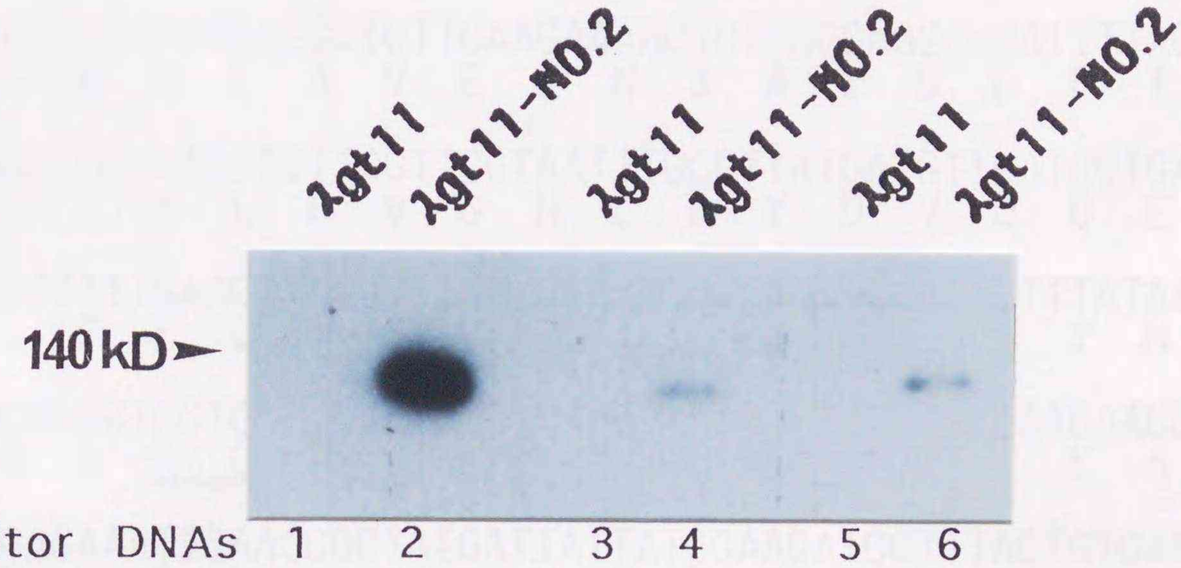


Figure 25


λgt11
λgt11-NO.2
λgt11
λgt11-NO.2
λgt11
λgt11-NO.2
140 kD
competitor DNAs 1 2 3 4 5 6

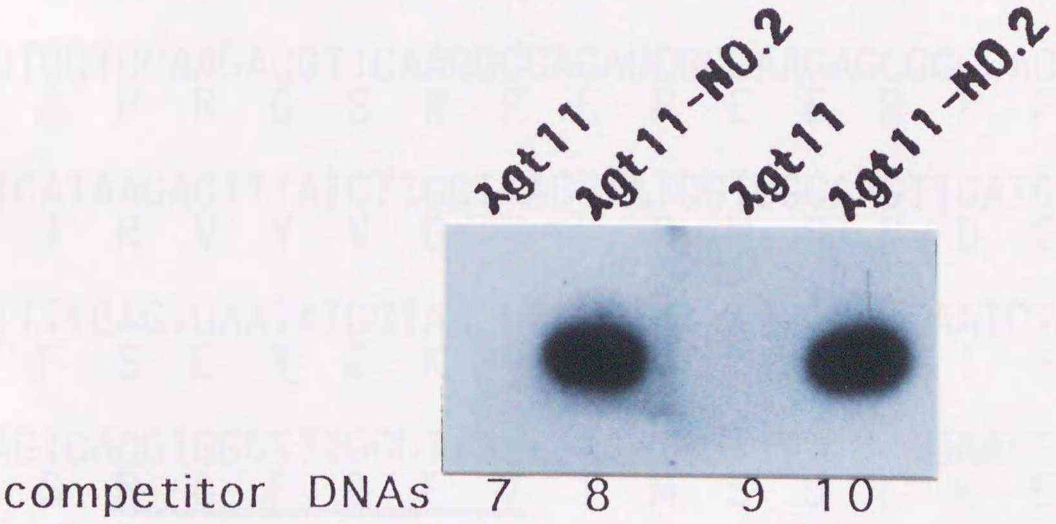

λgt11
λgt11-NO.2
λgt11
λgt11-NO.2
competitor DNAs 7 8 9 10

Figure 26

```
CAATTCCCCTCCACTTCAGTACCTTCTTCGTTCTATATCACCATGTCTGTGACTTCCACC    60
                                          M  S  V  T  S  T       6

ACTCCTCTCTTCAAAACCCTAACCATAGCCTCACTCTTCCTCACCAAAACCAAAACCACC   120
T  P  L  F  K  T  L  T  I  A  S  L  F  L  T  K  T  K  T  T      26

TCCTTATCACTCCCCTCCCAACCCTTCAAACTCCATCTCTCCCTCAATTCTTCACCCTCT   180
S  L  S  L  P  S  Q  P  F  K  L  H  L  S  L  N  S  S  P  S      46

CTCACTCTCTCACTAAAAACCAATCGCGCTTCACCTCTCTTCGCCGCGCAAGAAGGTGGC   240
L  T  L  S  L  K  T  N  R  A  S  P  L  F  A  A  Q  E  G  G      66

ACACTAACCACAGACGAACAAGGTGTCGATGAAACCGAAGGGGAACTTCTCGATTGGGAG   300
T  L  T  T  D  E  Q  G  V  D  E  T  E  G  E  L  L  D  W  E      86

GAAACCGACGCTGCTGAAAACGAGGCTGTTGAAGAAAACTCTGCGGAGGGAGATTTTACT   360
E  T  D  A  A  E  N  E  A  V  E  E  N  S  A  E  G  D  F  T     106

GAACCATCCGAAGATGCCAAATTGTTCGTTGGTAATTTGCCTTATGATGTTGATGGTGAG   420
E  P  S  E  D  A  K  L  F  V  G  N  L  P  Y  D  V  D  G  E     126

AAACTCGCCATGCTTTTTGAACAAGCTGGAACTGTAGAGATTGCAGAGGTTGTTTATAAC   480
K  L  A  M  L  F  E  Q  A  G  T  V  E  I  A  E  V  V  Y  N     146

AGGGAAACTGACCAGAGTCGTGGCTTTGGATTTGTAACCATGAGTACGGTTGAAGAAGCT   540
R  E  T  D  Q  S  R  G  F  G  F  V  T  M  S  T  V  E  E  A     166

GAAGCTGGCGCGGAGAAATTCAACCGCTATGATTATTATGGAAGATCCTTTACTGTGAAT   600
E  A  G  A  E  K  F  N  R  Y  D  Y  Y  G  R  S  F  T  V  N     186

AAGGCTGCTCCAAGAGGTTCAAGGCCAGAGCGGGAAGAGCGCCCACCTCGTCCCTATGTT   660
K  A  A  P  R  G  S  R  P  E  R  E  E  R  P  P  R  P  Y  V     206

CCTGTCATAAGAGTTTATGTTGGTAACTTATCGTGGCAAGTTGATGGTAGTAGACTTGAG   720
P  V  I  R  V  Y  V  G  N  L  S  W  Q  V  D  G  S  R  L  E     226

CAAGTTTTCAGTGAATATGGTAAGGTTGTGAATGCTCGAGTAATCTATGATAGAGAGACG   780
Q  V  F  S  E  Y  G  K  V  V  N  A  R  V  I  Y  D  R  E  T     246

GGTCAGTCACGTGGCTTTGGCTTTGTCACGATGTCTGATGAGAAAGAAATGAATGATGCT   840
G  Q  S  R  G  F  G  F  V  T  M  S  D  E  K  E  M  N  D  A     266

ATCGCTGCCCTTAATGGTCAGAGTTTGGAAGGGAGGACATTAAAAGTGAGTGTTGCTGAG   900
I  A  A  L  N  G  Q  S  L  E  G  R  T  L  K  V  S  V  A  E     286

GACAACAGGCCCAGTCGTGGCCCCTTTTGAGTGGTGCCGAGGATATAGTAGAAATTCTAC   960
D  N  R  P  S  R  G  P  F  *                                   295

TTACATAATTTGCTAAAGTTATTTCTCTATTTTAGTAGCCAAGGATATAGTAGAAATTCT  1020

ACTTACATAATTTTCTTCTCAGGCAGAAGGTTCCAAGTTTTCGCAATTCCTGTTGACTAT  1080

TTGCTTTTATGAAATGAGAAAAATGGGAGAGATGGAGTGTAGGTAGGAGTAGTTCTTACT  1140

GTCGCAAAAGCGTTTAGCAGGAAATTTTCTACCATTTTGTTTTACCGTACTTGAATGTAT  1200

TTGGATGAGTTATTTATTGTCAAAAAAAA                                 1221
```

Figure 27

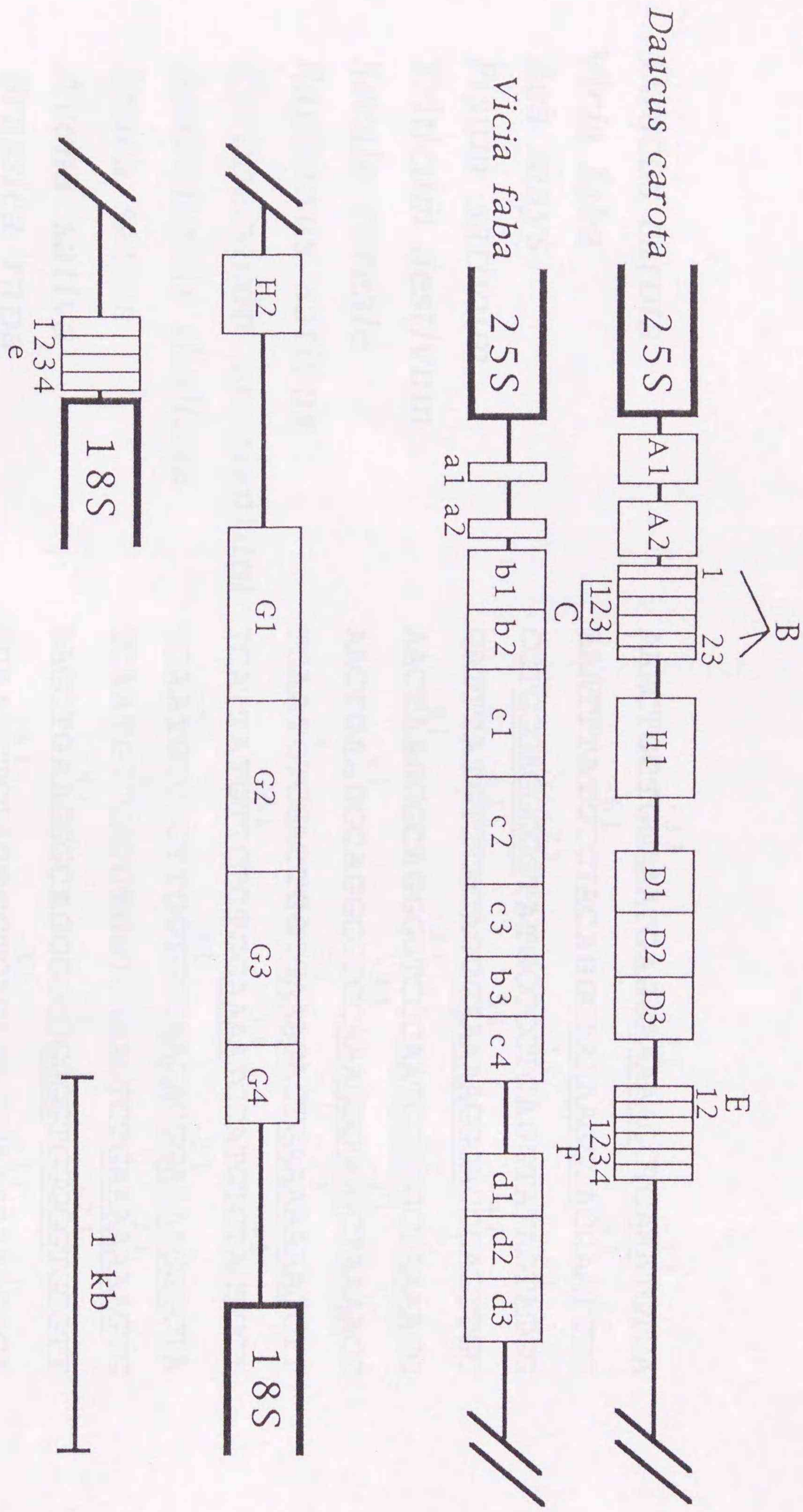

Daucus carota
Vicia faba
25S
A1
A2
B
1
23
C
123
H1
D1
D2
D3
E
12
1234
F
25S
a1
a2
b1
b2
c1
c2
c3
b3
c4
d1
d2
d3
H2
G1
G2
G3
G4
18S
1234
e
18S
1 kb

Figure 28

Daucus carota	-31 -21 -11 AAACT**GCTG**GAAAT**GAGG**AAAAAGCTCAAATGTCA
Vicia faba	-61 -41 AAATT**TATG**TCTA**CAGG**GAAAAAGGGACCAATTTG
Zea mays	-21 -11 -1 CCTGGGAGGGGG**TATG**CCCCT**CAGG**TATAGTAGGG
Pisum sativum	-61 -51 -41 GATT**TATG**TCTA**CAGG**GAAAAAGGGACCCAGTTGC
Triticum aestivum	-41 -31 -21 AAC**TAAG**GG**CAGG**GGTCCCAATGGGGGCTAAAACC
Secale cereale	-41 -31 -21 AAC**TGAAG**G**CAGG**GGTCCAAAGGGGGGCTAAAACC
Raphanus sativus	-41 -31 -21 TC**AATG**TCGA**CTGG**TGAAAGACTGAAAAAAAAGCTT
Lycopersicon esculentum	-61 -51 -41 TCAT**TATG**TTCT**CTGG**AAAATCCATGTCTACTCCT
Arabidopsis thaliana	-41 -31 -21 TC**AATG**TCTG**TTGG**TGCCAAGAGGGAAAAGGGCTA
Eruca sativa	-41 -31 -21 TC**AATG**TCTAT**CTGG**TGAAAGTCTGAAAAAAAAGTG
Avena sativa	-41 -31 -21 AAGC**TGAA**GGG**CAGG**GGTCGGGGTGGGGGTGCGTT
Brassica rapa	-41 -31 -21 CC**AATG**TCTA**CTGG**TGACAGACTGAAAAAAAAGCGT

Figure 29A

human (core)	-21 -11 -1	CGGT**TTTG**GGCCGC**CGGG**TTATTG
(UCE)	-121 -111	GGCGTT**TTTG**GGGA**CAGG**TGTCCGT
mouse	-21 -11 -1	TTCC**TATT**GGAC**CTGG**AGATAGGTA
rat	-21 -11 -1	GTCC**TATT**GTA**CTGG**AGATATATGC
Xenopus (initiation)	-21 -11 -1	TTTGG**CATG**TGCGGG**CAGG**AAGGTA
(spacer)		TTC**CAGG**AGCTCGGG**CAGG**GGGAGC
(T3)	-191 -181	ACTTGC**TCGG**CCGGG**CCGG**GCCGGG
yeast (upstream)	-120 -110	AAGAAA**TACG**TAGTT**AAGG**CCGAGCGA
(core)	-40 -30	TTTAGT**CATG**GAGTA**CAAG**TGTGAGGA
Acanthamoeba	-40 -30 -20	TTT**TCTG**GCACCTAAA**CTGG**TCGGACC
Physarum polycephalum	-30 -20 -10	C**TATG**CTTCTTAAAAAGAAACC**CAAG**A

Figure 29B

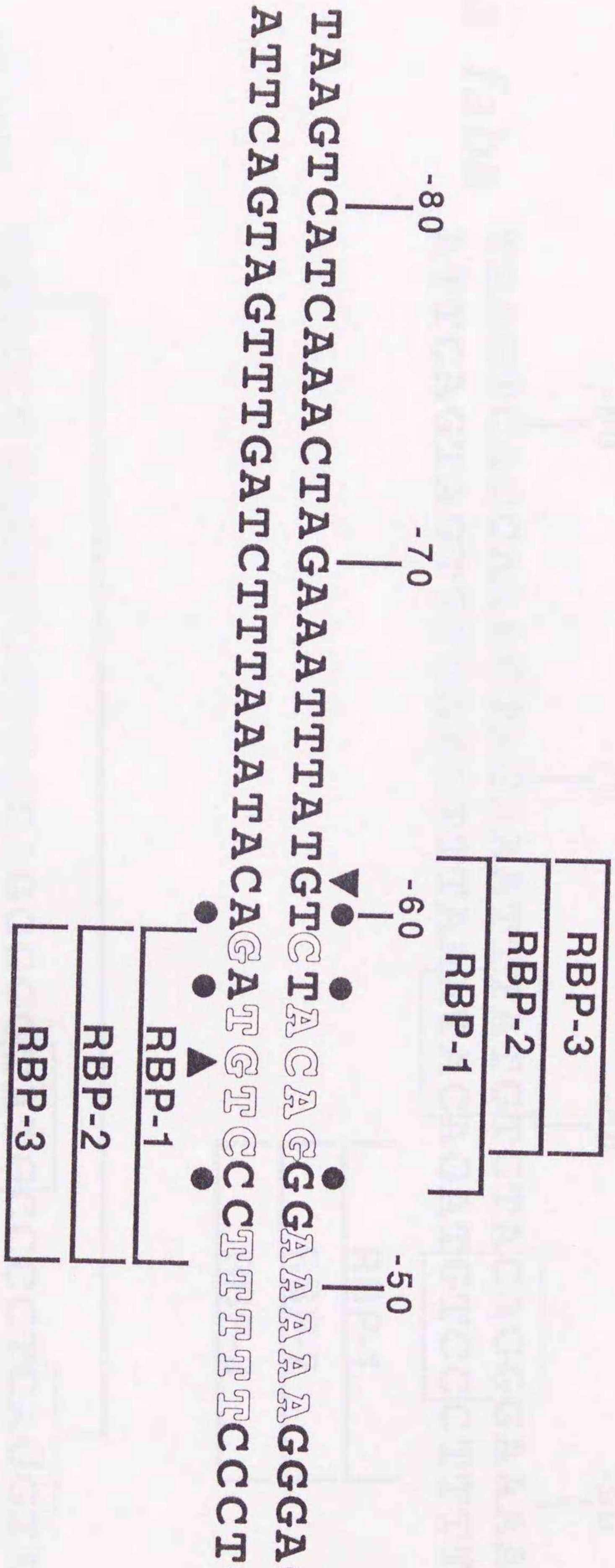

RBP-3
RBP-2
RBP-1
-80
-70
-60
-50
TAAGTCATCAAACTAGAAATTTATGTCTACACAGGGAAAAGGGACC
ATTCAGTAGTTTGATCTTTAAATACAGATGTGTCCCTTTTCCCTGG
RBP-1
RBP-2
RBP-3

Figure 30

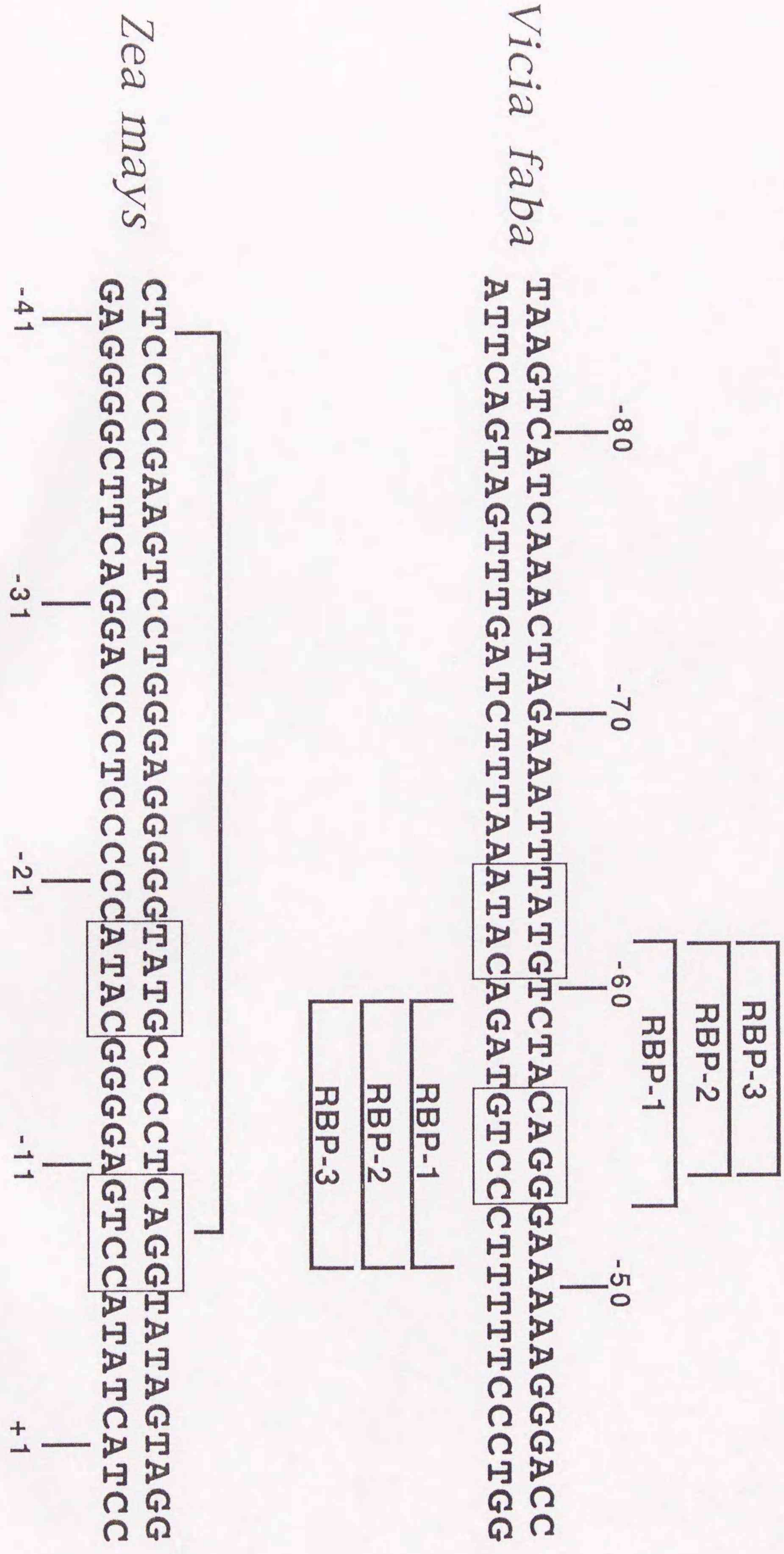
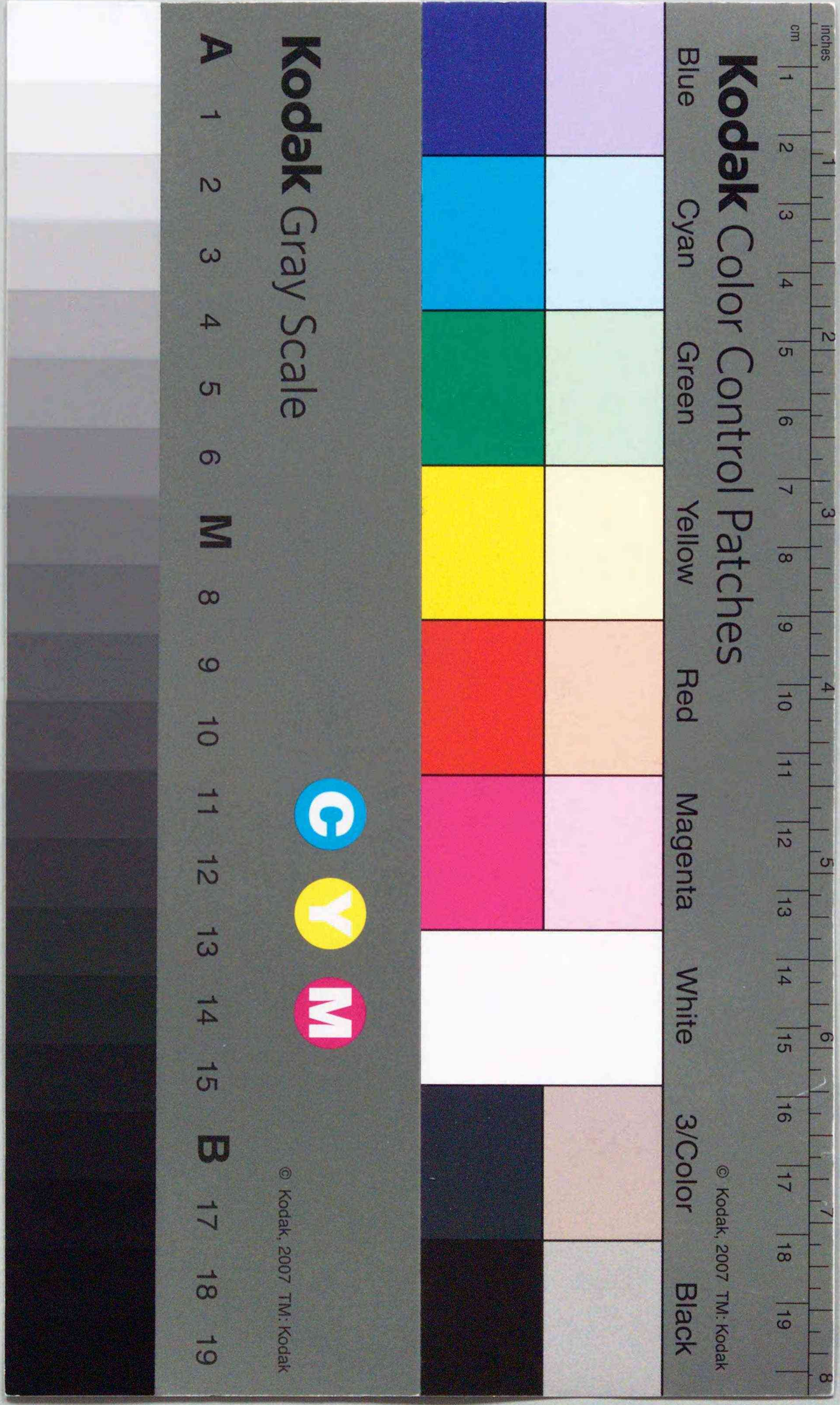

Vicia faba
-80
-70
-60
-50
TAAGTCATCAAACTAGAAATTTATGTCTACAGGAAAAAGGACC
ATTCAGTAGTTTGATCTTTAAATACAGATGTCCTTTTTCCTGG
RBP-3
RBP-2
RBP-1
RBP-1
RBP-2
RBP-3
Zea mays
CTCCCGAAGTCCTGGAGGGGTATGCCCCTCAGGTATAGTAGG
GAGGGCTTCAGGACCTCCCCATACGGGGAGTCCATATCATCC
-41
-31
-21
-11
+1

Figure 31


Kodak Color Control Patches
Blue
Cyan
Green
Yellow
Red
Magenta
White
3/Color
Black
© Kodak, 2007 TM: Kodak
cm
inches
Kodak Gray Scale
A
1
2
3
4
5
6
M
8
9
10
11
12
13
14
15
B
17
18
19
C
Y
M
© Kodak, 2007 TM: Kodak