Volume 225, number 1,2, 243-246 FEB 05383 December 1987

Nucleotide sequence of the PR-1 gene of *Nicotiana tabacum*

Masahiro Ohshima, Makoto Matsuoka, Naoki Yamamoto*, Yoshiyuki Tanaka+, Yuriko Kano-Murakami°, Yoshihiro Ozeki†, Akira Kato, Naoki Harada and Yuko Ohashi

*National Institute of Agrobiological Resources, *Forestry and Forest Product Research Institute, +National Institute of Agro-Environmental Science, °Fruit Tree Research Station, Tsukuba Science City, Ibaraki 305* and *†College of Arts and Sciences, Tokyo University, Meguro-ku, Tokyo 153, Japan*

Received 28 October 1987

A gene encoding one of the pathogenesis-related proteins, PR1a, and two related pseudogenes were isolated from *Nicotiana tabacum*. The cloned PR1a gene (pPR-γ) and one of the pseudogenes (pPR-α) were sequenced and found to have similar structures. The sequence of pPR-γ was quite similar to that of the cDNA clone of PR1a. The plasmid pPR-γ did not contain an intron and had a typical promoter sequence in the 5'-flanking region.

Pathogenesis-related protein; PR1a gene; Genomic cloning; Pseudogene; *(Nicotiana tabacum)*

1. INTRODUCTION

Some plants produce a set of new proteins in response to viral infection [1]. These proteins, called pathogenesis-related (PR) proteins, consist of 10 different proteins in *Nicotiana tabacum* [2], and are thought to be associated with 'acquired systemic resistance' [3]. PR-1 proteins constitute a well-characterized subgroup of PR proteins of low molecular mass and consist of 3 molecular species PR1a, PR1b and PR1c, which have similar biochemical and serological properties [4–6]. Recently, we and others have isolated and sequenced cDNA clones for PR-1 proteins [7–9]. However, the genomic sequences of PR proteins have not yet been reported. Here, we describe the nucleotide sequences of a PR1a gene and PR-1 related pseudogenes. Furthermore, we discuss the putative regulatory sequence in the 5'-flanking region.

Correspondence address: M. Ohshima, National Institute of Agrobiological Resources, Tsukuba Science City, Ibaraki 305, Japan

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00707

2. MATERIALS AND METHODS

Plants (*N. tabacum* cv. Samsun NN) were grown in a greenhouse at 25°C. DNA was prepared as follows. Mature leaves were homogenized in 50 mM Tris-HCl buffer (pH 7.5), containing 0.3 M sucrose, 5 mM $MgCl_2$ and 0.1 M *N,N*-diethyldithiocarbamate sodium salt. Nuclei were collected by centrifugation (12000 × *g*) for 20 min at 0°C, then lysed with 50 mM Tris-HCl buffer (pH 8.0) containing 20 mM EDTA and 1.5% sarkosyl. High-M_r DNA was purified by CsCl density gradient centrifugation, followed by complete digestion with *Eco*RI, and fractionation using 10–40% linear sucrose density gradient centrifugation.

DNA fragments were ligated into the phage vector λgt11 to clone PR-1 genes. The cDNA clone for the PR1a peptide, pPR1183 [9], was used as a

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

hybridization probe. The DNA sequence was determined using the dideoxynucleotide chain-termination method [10]. Southern blot analysis was performed as in [11]. Filters were washed with 0.2 × SSC, containing 0.5% SDS at 42°C for 3 h.

3. RESULTS AND DISCUSSION

Seven bands were detected on Southern blot analysis of total genomic DNA digested with *Eco*RI using the insert of pBR1183 as a probe. We termed each band α–η in ascending order of fragment size. The copy number of each fragment was estimated in comparison with an internal marker, based on a haploid genome size of 2×10^9 bp for *N. tabacum* [12]. The copy numbers of α, β, γ and ϵ, which gave strong signals, were very low, perhaps one or two. The results suggest that there are at least four kinds of PR-1-related genes having different *Eco*RI sites. The other fragments gave weak signals. The results suggest that the fragments were less homologous with the probe.

We cloned three fragments, α, γ and δ, into λgt11, then recloned into pSK(+) (Stratagene) for further analysis. Fig.1 shows the physical maps of α and γ. The synthetic 17-nucleotide oligomer, which corresponds to the antisense strand of the 5′-non-coding region just upstream of the initiation codon of PR1b (GACTATAGGAGAAATGT), hybridized to the region indicated by the broken line. The insert of pPR1183 hybridized to the region denoted by underlining. These results show that only one gene exists in these fragments. The hybridization signal of the cDNA insert against pPR-δ was very weak. We therefore infer that pPR-δ is a pseudogene of PR-1 or a gene of another subgroup of PR proteins of similar structure [13]. pPR-α and pPR-γ were further characterized by DNA sequence analysis.

The nucleotide sequences of pPR-α and pPR-γ are shown in fig.2. The sequence of pPR-γ between nucleotides 25 and 776 was identical with that of pPR1183. This result shows that pPR-γ is the genomic clone for PR1a. This gene did not have an intron. We have also sequenced 318 bp of the 3′-non-coding region and 436 bp of the 5′-upstream region. The sequence of the 5′-upstream region was very AT-rich. The transcription starting site was determined to be A

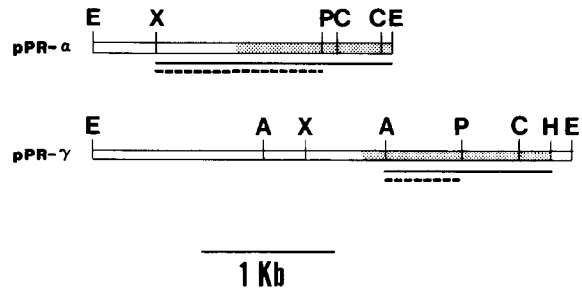

Fig.1. Physical maps of pPRα and γ. The regions which hybridize with the insert of pPR1183 and the synthetic oligomer just upstream of the coding region are indicated by the underlining and the broken line, respectively. The sequenced region is shaded. E, *Eco*RI; X, *Xho*I; P, *Pst*I; C, *Sca*I; A, *Acc*I; H, *Hin*dIII.

at 29 bp upstream from the first ATG by primer extension assay (not shown). This initiation point lies within a sequence homologous with the mammalian cap site consensus sequence, YCATTCR (Y, pyrimidine nucleoside; R, purine nucleoside [14]). The sequence TATAAATA is found at position −34 corresponding to the TATA box. A perfect dyad symmetrical structure was detected just upstream of this sequence (indicated by arrows in fig.2). It is notable that, overlapping this region, a sequence similar to the heat shock element of soybean was detected [15]. Actually, the sequence GTGAAATCTTCAAG designated by the shaded area in fig.2 shares 64% homology with the heat shock element 5′ upstream of the soybean hsp70 gene (CTGGAACATACAAG). This sequence is also similar to the consensus sequence of *Drosophila* (CTnGAAnnTTCnAG [16]).

Fig.2 also shows the nucleotide sequence of pPR-α. This sequence is quite similar to pPR-γ (PR1a gene) and the cDNA sequence of other PR1 groups (PR1b, PR1c [8]). However, there was an insertion of two bases in the coding region at position 145 in addition to some base exchanges. Moreover, the sequence corresponding to the initiation codon of PR1a was replaced by ATC in pPR-α. These base exchanges give rise to new, potential translation start sites (ATG) at some positions. However, they are immediately followed by the termination codon. These results show that pPR-α is an inactive pseudogene. This gene may be derived from a common ancestor of the PR-1 subgroup.

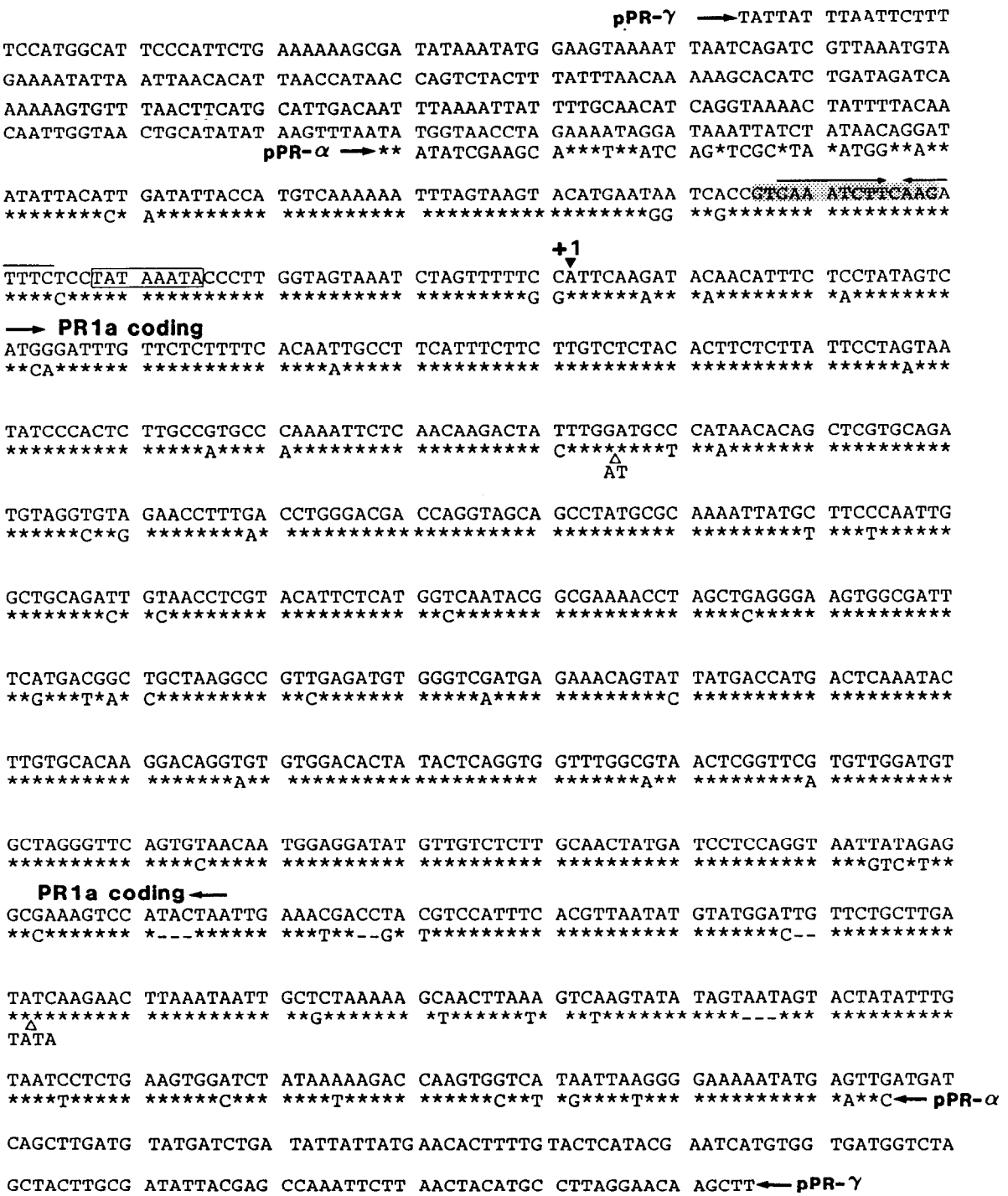

Fig.2. Nucleotide sequences of pPR-α and pPR-γ. The sequence of pPR-α is shown only where differences occur with pPR-γ. Identical bases are indicated by asterisks. A bar and open triangle signify a deletion and insertion, respectively. The position +1 is denoted the putative transcription starting site. The potential TATA sequence is boxed. The dyad symmetrical structure is indicated by arrows. Transcription starting point is indicated by a closed triangle.

ACKNOWLEDGEMENT

This work was supported by a project grant from the Ministry of Agriculture, Forestry and Fisheries of Japan.

REFERENCES

[1] Van Loon, L.C. (1985) Plant Mol. Biol. 4, 111–116.
[2] Van Loon, L.C. (1982) in: Active Defence Mechanisms in Plants (Wood, R.K.S. ed.) pp.247–274, Plenum, New York.
[3] Van Loon, L.C. (1985) Virology 67, 566–575.
[4] Antoniw, J.F., White, R.F., Barbara, D.J., Junes, P. and Longley, A. (1985) Plant Mol. Biol. 4, 55–60.
[5] Matsuoka, M. and Ohashi, Y. (1984) J. Gen. Virol. 65, 2209–2215.
[6] Hooft van Huijsduijnen, R.A.M., Cornelissen, B.J.C., Van Loon, L.C., Van Boom, J.H., Trump, M. and Bol, J.F. (1985) EMBO J. 4, 2167–2171.
[7] Cornelissen, B.J.C., Hooft van Huijsduijnen, R.A.M., Van Loon, L.C. and Bol, J.F. (1986) EMBO J. 1, 37–40.
[8] Pfitzner, U.M. and Goodman, H.M. (1987) Nucleic Acids Res. 15, 4449–4465.
[9] Matsuoka, M., Yamamoto, N., Kano-Murakami, Y., Tanaka, Y., Ozeki, Y., Hirano, H., Kagawa, H., Ohshima, M. and Ohashi, Y. (1987) Plant Physiol., in press.
[10] Sanger, F., Coulson, A.R., Barrell, B.J., Smith, A.J.H. and Roe, B.A. (1980) J. Mol. Biol. 143, 161–178.
[11] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, NY.
[12] Bennett, M.D. and Smith, J.B. (1976) Phil. Trans. R. Soc. Lond. Ser.B 274, 227–273.
[13] Hooft van Huijsduijnen, R.A.M., Van Loon, L.C. and Bol, J.F. (1986) EMBO J. 5, 2057–2061.
[14] Busslinger, M., Portmann, R., Irminger, J.-C. and Birnstiel, M.L. (1980) Nucleic Acids Res. 8, 957–977.
[15] Czarnecka, E., Gurley, W.B., Nagao, R.T., Mosquera, L.A. and Key, J.L. (1985) Proc. Natl. Acad. Sci. USA 82, 3726–3730.
[16] Pelham, H.R.B. (1982) Cell 30, 517–528.