J. Plant Biol. 41 (3) : 227~232 (1998)

SHORT COMMUNICATION

Cloning of a Polyubiquitin Gene from *Nicotiana tabacum* and Comparison to Other Polyubiquitin Genes

Kim, Jin Hyun[1,2], Soo Min Park[2], Ho Rim Lee[3], Jae-Ho Pyee[3], Choo Bong Hong[1,2,*] and Kwang Woong Lee[1]

[1]*Department of Biology, Seoul National University, Seoul 151-742, Korea*
[2]*Institute for Molecular Biology and Genetics, Seoul National University, Seoul 151-742, Korea*
[3]*Department of Molecular Biology, Dankook University, Seoul 140-714, Korea*

Using a tobacco cDNA clone as a probe, a genomic clone named TUQG-4, coding for a tobacco polyubiquitin protein with the five head-to-tail repeats of ubiquitin monomer was isolated. The five ubiquitin units were completely conserved except for the extra phenylalanine at the carboxy terminus of the last ubiquitin monomer. The putative open reading frame identified from the nucleotide sequence showed two possible intron sequences in the coding region for the first ubiquitin monomer. When the amino acid sequence deduced from the nucleotide sequence of TUQG-4 was compared to the amino acid sequences coded by other polyubiquitin genes of tobacco, there were three or four amino acid differences in the sequence. When the nucleotide sequences coding for the ubiquitin monomers were compared for various species origins, the degree of identity was at the highest between the ubiquitin monomers in one polyubiquitin and did not reflect the distance of the phylogenetic relationship.

Keywords: genomic clone, *Nicotiana tabacum*, nucleotide sequence comparison, polyubiquitin

Ubiquitin is one of the most highly conserved proteins known, and has been proposed to be involved in diverse processes such as DNA repair, cell cycle control, circadian rhythm, stress response, apoptosis, and other important cellular functions in eukaryotic cells. The ubiquitin-dependent protein degradation pathway appears to be the major mechanism to degrade abnormal and short-lived proteins in eukaryotic cells. Alternatively, ubiquitin can reversibly join to an acceptor protein in the nucleus and can modulate protein function without destabilizing the acceptor protein. The combined functions of ubiquitin thus play major roles in controlling the physiology and development of a eukaryotic cell (Ciechanover, 1994; Hochstrasser, 1995; Drexler, 1997).

In *Nicotiana tabacum*, there seem to be at least five ubiquitin genes present, and there have been three different heat-shock responses of the ubiquitin gene family reported (Seo *et al.*, 1996; Park *et al.*, 1998). For the identification of differential functions of ubiquitins coded by different genes, it would be essential to obtain and characterize the genomic clones.

*Corresponding author: Fax +82-2-874-1206

MATERIALS AND METHODS

Screening of the Genomic Library

Tobacco genomic library in EMBL3 arms were plated out on an *Escherichia coli* lawn (strain KW 251) as described by Sambrook *et al.* (1989). About 1×10^5 plaques were formed on a 23 cm × 23 cm LB bottom agar plate and blotted on a Hybond-N membrane for the primary screening. The membrane was prehybridized and hybridized in 50% formamid, 5X SSPE, 5X Denhardt's solution, 1% SDS, and 100 g/ml denatured salmon sperm DNA at 42°C. The ubiquitin cDNA clone, pTUQC-1 (Seo *et al.*, 1996), was labeled with ^{32}P-dCTP using Prime-a-Gene system (Promega, USA). After washing in 0.5X SSPE and 0.1% SDS at 42°C, the membrane was exposed on an X-ray film (Kodak, USA) with two intensifying screens (Dupont, USA) at −70°C (Sambrook *et al.*, 1989). For the secondary

screening, plaques at the positive signal in the primary screening were picked and plated out to form well-separated plaques. Hybridization of the blot was carried out as in the primary screening.

DNA Sequencing and Analysis

Insert of the genomic clone was subcloned into pBluescript II SK(+), and a set of unidirectional deletion derivatives was obtained by exonuclease III using the deletion kit for kilo-sequencing (Takara, Japan). Sequencing reactions were performed using Sequenase version 2.0 (United States Biochemical Corportation, USA). The nucleotide and the deduced amino acid sequences were analysed using the current GenBank and Swiss-Prot databases.

RESULTS AND DISCUSSION

From the screening of a tobacco (*Nicotiana tabacum* cv. Xanthi) genomic library in EMBL3 with the ^{32}P-labeled ubiquitin cDNA clone, pTUQC-1 (Seo *et al.*, 1996), a putative polyubiquitin genomic clone designated as TUQG-4 was isolated. The genomic clone had an insert size of 15-kb at the *Sal*I site of EMBL3 and had several *Eco*RI and *Sal*I sites flanking the putative ubiquitin coding region (Fig. 1). The 3.1-kb *Eco*RI/*Sal*I restriction fragment hybridizing to the ^{32}P-labeled pTUQC-1 was subcloned into pBluescriptII SK(+) for nucleotide sequencing.

A nucleotide sequence determined from the deletion series of the subcloned *Eco*RI/*Sal*I fragment of TUQG-4 identified an open reading frame for a polyubiquitin protein. A sequence of 541 nucleotides was determined in the 5'-upstream of the open reading frame which contained putative CAAT and TATA boxes. The consensus transcription initiation sequence, TCATCA (Joshi, 1987), could not be located, and the putative translation initiation codon was not flanked with the consensus sequence

```
                                    AAATCCTTACATTAGAATCAGAACCAAAAAGATTTTGAAGAAGCGAAATC   50
CTTTTAGCTTATAATTAAATTCCCTGCATACGAAGACGAAATATTATACGAACTCACCAGAGAATGATG  119
AGCCGCTGGCGATATTCGATCAAACAATCAAACCGAATTCGAGACAGAGAAACCAGTACCACCTAAGCC  188
TTGCCTTCTAAACAAAAACGAAAACAAGTGGAAGTATTTAGCAATAACTACTTTTGGTTGTCGACGTGC  257
CCTTGCTGTTATTAACGCTTACGTCTTTTTGAGTTTCGAAACGCGTTTATGTACATAGTTTTTGTAGAT  326
TAGAAGAGTTTGACAATGCTATTTGGAATATATTCTTGTTTCAAATACTGTTGAATAAAATGTAAAGTA  395
CAGAAGTTCTAAAATATTGATCTTGTGTGTTTCTATGACCCCAGTTAATGGAAAGTAAGGGTCTTGGTG  464
ACATTGCCGGAGATTACTTGAACACGTAGCATTTACGCAGAAAAAATCATTCTTATTGTAAATTTAAAC  533
AACACAAGATGCAAATTTTCGTTAAAACTTTAACTGGAAAGACTATCACCCTCGAG GTTTGTTTAATC  601
          M  Q  I  F  V  K  T  L  T  G  K  T  I  T  L  E                16
GGATGAGCGCATTAAAGGCTAATTAATTTTAG GTTGAGAGCTCTGATACCATTGAAAATGTAAAACAA  669
                                    V  E  S  S  D  T  I  E  N  V  K  Q    28
AAAATCCAAGACAAAGAAGGAATCCCTCCAGATCAACAAAGACTTATCTTTGCCGGAAAGCAATTAGAA  738
 K  I  Q  D  K  E  G  I  P  P  D  Q  Q  R  L  I  F  A  G  K  Q  L  E    51
GACGGAAGAACTCTTGCTGACTATAACATTCAAAAGGAGTCTACTCTCCACTTGGTTCTTAG GTAATT  806
 D  G  R  T  L  A  D  Y  N  I  Q  K  E  S  T  L  H  L  V  L  R          72
TCTCTGTCGCAAAATCTTTCAATCCGCTCTTTAG A TTAAGAGGAGGAATGCAAATTTTCGTTAAG  873
                                    L  R  G  G  M  Q  I  F  V  K          82
ACTTTAACTGGAAAGACCATTACCTTAGAAGTCGAATCCTCTGACACCATCGAAAATGTCAAACAAAAG  942
 T  L  T  G  K  T  I  T  L  E  V  E  S  S  D  T  I  E  N  V  K  Q  K    105
ATTCAAGATAAGGAAGGAATTCCTCCAGATCAACAAAGACTTATCTTTGCTGGAAAGCAACTTGAAGAC  1011
 I  Q  D  K  E  G  I  P  P  D  Q  Q  R  L  I  F  A  G  K  Q  L  E  D    128
GGCCGCACCTTAGCCGACTACAATATTCAAAAGGAATCCACTCTCCATTTGGTCCTTAGACTCCGTGGT  1080
 G  R  T  L  A  D  Y  N  I  Q  K  E  S  T  L  H  L  V  L  R  L  R  G    151
GGAATGCAAATCTTTGTTAAGACTCTCACTGGGAAGACTATTACTTTAGAAGTCGAATCATCCGATACC  1149
 G  M  Q  I  F  V  K  T  L  T  G  K  T  I  T  L  E  V  E  S  S  D  T    174
ATTGAAAATGTAAAGCAAAAGATCCAAGACAAGGAAGGAATCCCTCCAGATCAACAACGTTTGATTTTC  1218
 I  E  N  V  K  Q  K  I  Q  D  K  E  G  I  P  P  D  Q  Q  R  L  I  F    197
GCCGGAAAGCAATTAGAAGATGGCCGCACATTAGCTGATTACAACATTCAAAAGGAATCCACTCTTCAC  1287
 A  G  K  Q  L  E  D  G  R  T  L  A  D  Y  N  I  Q  K  E  S  T  L  H    220
TTGGTTCTTAGACTTCGTGGTGGTATGCAAATCTTTGTCAAGACCCTTACTGGCAAAACTATTACTCTT  1356
 L  V  L  R  L  R  G  G  M  Q  I  F  V  K  T  L  T  G  K  T  I  T  L    243
GAAGTAGAAAGCTCTGATACTATTGAAAACGTCAAGCAAAAGATCCAAGACAAGGAAGGAATCCCTCCA  1425
 E  V  E  S  S  D  T  I  E  N  V  K  Q  K  I  Q  D  K  E  G  I  P  P    266
GATCAACAACGTCTCATCTTTGCCGGAAAACAATTGGAAGATGGAAGAACTTTGGCTGACTATAACATT  1494
 D  Q  Q  R  L  I  F  A  G  K  Q  L  E  D  G  R  T  L  A  D  Y  N  I    289
CAAAAGGAATCCACTCTTCACTTGGTTCTTAGATTAAGAGGAGGAATGCAAATCTTTGTCAAAACTTTG  1563
 Q  K  E  S  T  L  H  L  V  L  R  L  R  G  G  M  Q  I  F  V  K  T  L    312
ACTGGAAAGACAATTACTTTGGAAGTCGAATCTAGCGACACCATTGAAAATGTCAAGCAAAAGATTCAA  1632
 T  G  K  T  I  T  L  E  V  E  S  S  D  T  I  E  N  V  K  Q  K  I  Q    335
GATAAGGAAGGAATCCCTCCAGATCAACAAAGACTTATCTTTGCTGGAAAACAATTAGAAGATGGTCGT  1701
 D  K  E  G  I  P  P  D  Q  Q  R  L  I  F  A  G  K  Q  L  E  D  G  R    358
ACCTTAGCGGACTATAACATTCAAAAGGAATCTACTCTTCACTTGGTTCTTAGATTGAGAGGAGGTTTC  1770
 T  L  A  D  Y  N  I  Q  K  E  S  T  L  H  L  V  L  R  L  R  G  G  F    381
TAA ACATGTGCTTTGCAATGTTGTCTGATTGAATAAATTGTTGTCAATCCTTATATAATGTCTTATAT  1837
*
AATGTCTTATATTTTTTACCTTTGTTGGGATAAATGGTGATGCCGCGGATTGTAAGTAACATCAATTTC  1906
AATTTAGCCGGAAGTACCATTTTGAATTTGCTTGTACGATGTTCAAAGGGATTAAACTTTTGCTTCTCA  1975
GGTTTGATTACATTGCATACCAAAGGCATATTGAAATCTGATATTCCAATACAGATCGATTTAAAGTTC  2044
GAGATCCAAGACAAGGAAGGAATCCCTCCAGATCAACAACGTCTCATCTT                     2094
```

Fig. 2. Nucleotide sequence and a deduced amino acid sequence for TUQG-4. A putative open reading frame was started with ATG and ended with TAA. Amino acid sequence is represented in one letter symbol. The methionine residues expected to be the starting amino acids for each ubiquitin monomer are shown in boldface. CAAT box-like and TATA box-like sequences are underlined. Putative intron sequences are underlined. Polyadenylation signal-like sequences are underlined.

frequently found in plants, AACAATGGC (Futterer and Hohn, 1996). Codon usage in the open reading frame showed heavy biases in arginine, glutamine, glutamic acid, glycine, and lysine to AGA, CAA, GAA, GGA, and AAG, respectively. There were two putative intron sequences interrupting the open reading frame of the first ubiquitin monomer. The

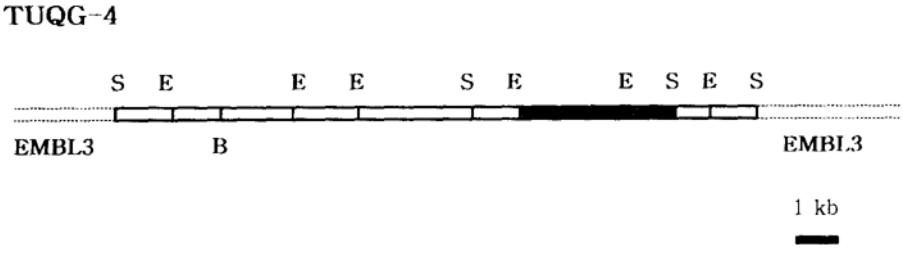

Fig. 1. Restriction map of TUQG-4. About a 15-kb DNA insert was in an EMBL3 vector at *Sal*I site. The shaded fragment hybridized specifically to the ^{32}P-labeled ubiquitin cDNA clone, TUQC-1 (Seo *et al.*, 1996). B, *Bam*HI; E, *Eco*RI; S, *Sal*I.

putative intron sequences were started with GT and ended with AG-the consensus intron-flanking sequence. In the 3'-downstream of the open reading frame, 321 nucleotides were determined, and polyadenylation signal-like sequences (Rothnie, 1996); far-upstream element, TGTTTGTC; and near-upstream element, AATAAA, were located (Fig. 2).

Amino acid sequences deduced from the open reading frame of TUQG-4 identified five direct repeats of the typical ubiquitin monomer. The polyprotein had four monomers of 76 amino acids and one monomer of 77 amino acids. Molecular mass could be estimated as 43 kD, and the isoelectric point was calculated as 7.48 (LaserGene, DNASTAR, USA). Comparison of the amino acid sequence between the ubiquitin monomers in TUQG-4 showed 100% identities except for the last repeat which had one additional amino acid, phenylalanine. Perfect identities between the ubiquitin monomers in a polyubiquitin and an additional amino acid at the last repeat followed the previously reported characteristics of ubiquitins (Christensen *et al.*, 1992; Genschik *et al.*, 1992; Hu and Henney, 1997).

The amino acid sequence of ubiquitin coded by TUQG-4 was compared to the ubiquitins from various sources. They were TUQG-2, a polyubiquitin gene from *N. tabacum* (Park *et al.*, 1998), TUQG-3, a ubiquitin-extension protein gene from *N. tabacum* (Park *et al.*, 1996), NTU4, a polyubiquitin gene from *N. tabacum* (Genschik *et al.*, 1992), ATUB4, a polyubiquitin gene from *Arabidopsis thaliana* (Burke *et al.*, 1988), YSTU4, a polyubiquitin gene from yeast (Ozkaynak *et al.*, 1987), and HMUB3, a polyubiquitin gene from humans (Baker and Board, 1987). For *A. thaliana*, yeast, and human, we chose the polyubiquitin genes showing the highest similarity to TUQG-4 from each organism. The comparison proved strong conservation of ubiquitin

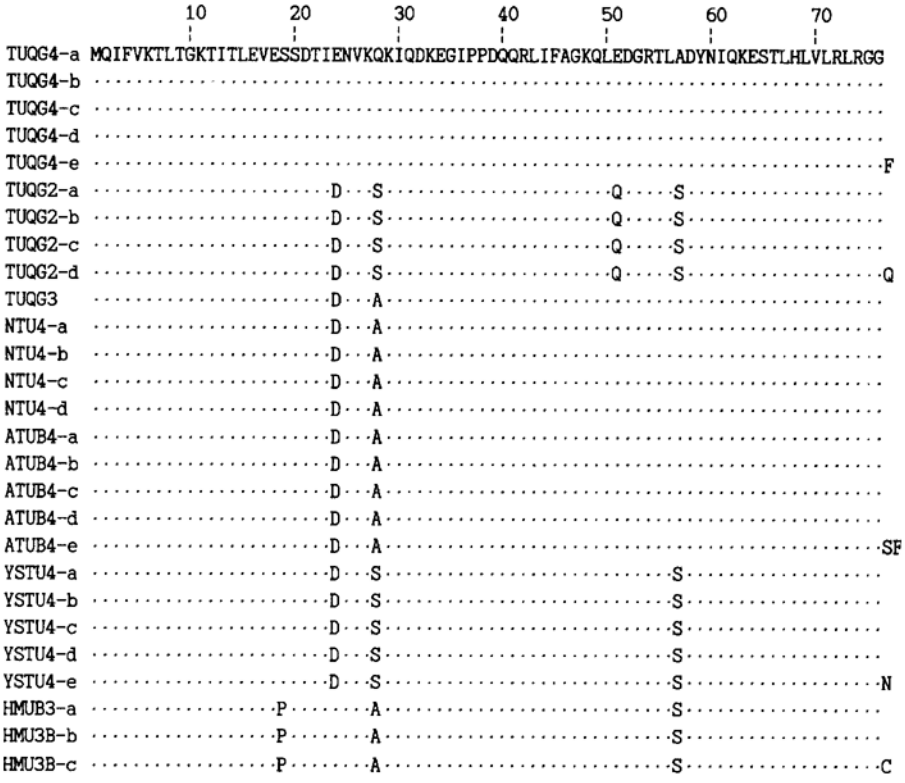

Fig. 3. Homology comparison of the tobacco ubiquitin monomers with the reported ubiquitins at the amino acid level. Amino acid sequence is in one letter symbol. TUQG4-a, -b, -c, -d and -e represent the five ubiquitin monomers shown in Fig. 2. TUQG2-a, -b, -c and -d represent the four ubiquitin monomers in a polyubiqutin gene of *N. tabacum* (Park *et al.*, 1998). TUQG3 represents the ubiquitin monomer in a ubiquitin-extension fusion protein (Park *et al.*, 1996). NTU4-a, -b, -c and -d represent the ubiquitin monomers in the polyubiquitin gene of *N. tabacum* (Genschik *et al.*, 1994). ATUB4-a, -b, -c, -d and -e represent the ubiquitin monomers in the polyubiquitin gene of *A. thaliana* (Burke *et al.*, 1988). YSTU4-a, -b, -c, -d and -e represent the ubiquitin monomers in the polyubiquitin gene of yeast (Ozkaynak *et al.*, 1987). HMUB3-a, -b and -c represent the ubiquitin monomers in the polyubiquitin gene of humans (Baker and Board, 1987).

proteins, i.e., a difference in four amino acids at the most (Fig. 3).

Ubiquitin monomers in TUQG-4 were compared with the ubiquitins from the various sources on the nucleotide sequence. The data presented in percent divergence showed that the nucleotide sequence is most alike among the ubiquitin monomers in one polyubiquitin gene. When the nucleotide sequences of ubiquitin monomers from different polyubiquitin genes were compared, the degree of divergence did not reflect the phylogenetic relationship of the source organism. As an example, the degree of divergence between the first ubiquitin monomers of TUQG-2 and TUQG-4 was 28.2%, and the degree of divergence between the first ubiquitin monomers of TUQG-4 and HMUB3 was 26.0% (Fig. 4).

The two putative intron sequences of TUQG-4 in the first ubiquitin monomer is characteristic. Intron sequences very rich in A and T are common in plant polyubiquitin genes, but they are mostly located in the 5' upstream untranslated region (Christensen *et al.*, 1992; Genschik *et al.*, 1994; William and Garbarino, 1996). So far, we have fully sequenced two polyubiquitin genes from *N. tabacum*, one from this report and the other from the previous report (Park *et al.*, 1998), and both of them had intron sequences in the first ubiquitin monomer. The polyubiquitin genes with the intron sequence in the open reading frame reported are from *A. thaliana* and *Caenorhabditis elegans*. The *A. thaliana* polyubiquitin gene is considered a pseudogene with amino acid substitution between ubiquitin monomers (Callis *et al.*, 1995), and the *C. elegans* polyubiquitin gene has the intron sequence at the same position in the ubiquitin repeating units (Graham *et al.*, 1989).

The ubiquitin gene is unique among the repeated polyprotein genes. There are genes coding for perfectly or imperfectly repeated polyproteins (Slabaugh *et al.*, 1989; Hourcade *et al.*, 1990; Atkinson *et al.*, 1993; Colby and Williams, 1993), but polyubiquitin is the only example whose repeating unit always starts with ATG, the codon for methionine, or the first codon of the eukaryotic mRNA. So far, no evidence has been raised showing the importance of methionine for ubiquitin function. Thus it is not likely that evolutionary pressure has worked to make every ubiquitin start with methionine. It has been suggested in *A. thaliana* that gene duplication coupled with unequal crossing-over brought about the polyubiquitination and a sudden

TUQG4-a	1																									
TUQG4-b	2	17.1																								
TUQG4-c	3	18.9	15.0																							
TUQG4-d	4	14.0	20.6	14.5																						
TUQG4-e	5	16.7	14.5	14.9	14.9																					
TUQG2-a	6	28.2	25.6	22.8	29.1	29.8																				
TUQG2-b	7	29.1	25.6	21.5	26.0	28.5	10.5																			
TUQG2-c	8	31.7	24.2	27.6	30.4	28.1	14.9	16.7																		
TUQG2-d	9	28.1	21.9	24.1	25.9	23.9	16.7	19.7	17.1																	
TUQG3	10	26.9	24.2	22.5	25.1	27.3	18.1	18.1	22.5	21.6																
NTU4-a	11	20.7	23.3	24.2	22.0	24.7	23.3	22.5	25.6	22.5	17.1															
NTU4-b	12	22.0	25.6	24.2	22.9	26.0	22.0	21.6	24.7	24.7	19.3	13.2														
NTU4-c	13	25.6	24.7	22.0	22.0	24.7	19.8	22.0	25.1	25.1	18.4	14.0	12.7													
NTU4-d	14	25.1	25.6	23.3	22.9	26.0	18.9	20.7	24.2	25.1	18.0	17.1	13.6	6.6												
ATUB4-a	15	20.7	22.5	22.0	23.8	24.7	18.9	19.4	23.8	23.8	21.1	19.3	21.5	18.9	21.1											
ATUB4-b	16	22.5	21.1	22.5	20.3	23.8	22.5	21.6	20.7	22.9	18.4	18.0	20.6	19.3	18.4	21.5										
ATUB4-c	17	21.6	25.6	25.6	24.7	25.1	22.9	22.5	22.5	23.8	17.5	18.9	18.0	18.9	19.7	19.3	18.0									
ATUB4-d	18	23.3	20.3	24.2	24.7	25.1	22.9	23.3	20.7	26.0	18.9	23.2	23.2	19.7	18.0	20.6	16.7	19.3								
ATUB4-e	19	25.4	25.0	24.6	23.2	26.1	21.1	20.2	23.2	24.4	17.1	18.4	19.7	17.5	18.0	21.1	15.8	16.2	16.2							
YSTU4-a	20	23.3	23.3	23.7	25.1	24.6	22.4	23.7	25.4	23.2	24.2	23.8	24.2	23.8	22.5	22.5	23.3	25.1	23.8	27.2						
YSTU4-b	21	23.3	23.3	20.6	22.0	21.5	23.7	24.1	24.6	23.2	21.1	19.8	22.0	19.8	23.3	20.7	20.3	21.1	24.2	22.8	14.9					
YSTU4-c	22	22.5	23.8	25.0	23.3	24.1	25.9	27.6	26.3	27.2	25.6	25.6	25.6	23.3	23.3	21.1	25.1	25.6	25.6	28.5	13.6	17.1				
YSTU4-d	23	20.3	21.1	21.5	22.9	19.7	25.0	28.9	29.4	23.2	25.1	19.8	22.5	21.6	22.5	22.9	22.9	22.0	26.9	26.3	19.7	13.6	14.0			
YSTU4-e	24	23.2	23.7	23.2	22.8	24.4	23.7	25.0	27.2	26.9	25.6	23.3	25.1	24.7	23.8	22.0	19.4	25.1	26.0	27.8	15.4	17.1	9.6	14.0		
HMUB3-a	25	26.0	26.0	29.4	26.9	26.8	18.0	20.2	17.1	23.2	18.1	20.3	20.7	20.3	19.8	19.4	21.6	19.4	21.6	19.7	26.8	22.8	26.8	26.3	25.9	
HMUB3-b	26	25.6	25.1	27.6	26.4	25.9	17.1	19.7	18.0	21.1	18.9	20.7	20.3	20.7	20.3	20.3	20.3	18.1	21.6	20.6	24.6	21.9	26.8	25.9	25.9	3.5
HMUB3-c	27	25.9	25.0	27.2	25.9	26.1	18.0	20.6	18.9	22.2	20.3	22.5	21.6	22.0	21.6	20.7	21.6	18.9	21.1	21.4	25.0	22.8	27.6	26.8	26.9	5.7	2.2
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27

Fig. 4. Homology comparison of the ubiquitin monomers at the nucleotide level. Percent divergences between two ubiquitin monomers are shown. TUQG4-a, -b, -c, -d and -e; TUQG2-a, -b, -c and -d; TUQG3; NTU4-a, -b, -c and -d; ATUB4-a, -b, -c, -d and -e; YSTU4-a, -b, -c, -d and -e; and HMUB3-a, -b and -c are as in Fig. 3.

change in the level of the nucleotide sequence divergence along a ubiquitin gene (Callis *et al.*, 1995; Sun *et al.*, 1997). From TUQG-4 and TUQG-2, the polyubiquitin genes of *N. tabacum*, we were not able to find an indication of unequal crossing-over. The fact of ubiquitin always starting with methionine, the very wide range of polymerization of ubiquitin, i.e., from 3 repeats in *Phytophtora infestans* (Pieterse *et al.*, 1991) to 52 repeats in *Trypanosoma cruzi* (Swindle *et al.*, 1988), and the rather complex organization of the genes while maintaining the unit of ubiquitin gene, i.e., polyubiquitin genes in *A. thalina* code for ubiquitins with identical amino acid sequence (Callis *et al.*, 1995; Sun *et al.*, 1997), and that polyubiquitin genes in *N. tabacum* code for ubiquitins with amino acid substitutions, differentiate the gene duplication and gene conversion process in the development of polyubiquitin genes from those involved in other repeated genes.

Whatever the mechanism of polyubiquitination is, the process must have progressed quickly. Differences in the structure and expression patterns of ubiquitin genes of *N. sylvestris* and *N. tabacum* might indicate fast recombination events occuring in polyubiquitin genes. In *N. tabacum*, 2.6-kb and 1.9-kb transcripts are heat-stress inducible (Seo *et al.*, 1996), and in *N. sylvestris* a 1.6-kb transcript is stress inducible (Genschik *et al.*, 1992). This difference simply cannot be explained by the allopolyploidization of *N. sylvestris* and *N. tomentosiformis* which established *N. tabacum* (Leitch and Bennett, 1997). Considering that the number of copies of the polyubiquitin gene in the chromosome between these two species is similar and that the size of the polyubiquitin transcripts detected in these two species probably reflects multiples of a ubiquitin monomer coding sequence, the difference detected from these two species may imply the occurrence of a fast recombination event of polyubiquitin genes while maintaining the gene structure of the ubiquitin monomer after the allopolyploidization.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Genetic Engineering Research Fund of the Ministry of Education, Republic of Korea, to C.B. Hong.

LITERATURE CITED

Atkinson, A.H., R.L. Heath, R.J. Simpson and A.E. Clarke. 1993. Proteinase inhibitors in *Nicotiana alata* stigmas are derived from a precursor protein which is processed into five homologous inhibitors. *Plant Cell* **5**: 203-213.

Baker, R.T. and P.G. Board. 1987. The human ubiquitin gene family: structure of a gene and pseudogenes from the Ub B sub family. *Nucleic Acids Res.* **26**: 443-463.

Burke, T.J., J. Callis and R.D. Vierstra. 1988. Characterization of a polyubiquitin gene from *Arabidopsis thaliana*. *Mol. Gen. Genet.* **213**: 435-443.

Callis, J., T. Carpenter, C.W. Sun and R.D. Vierstra. 1995. Structure and evolution of genes encoding polyubiquitin-like proteins in *Arabidopsis thaliana* ecotype Columbia. *Genetics* **139**: 921-939.

Christensen, A.H., R.A. Sharrock, and P.H. Quail. 1992. Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplast by electroporation. *Plant Mol. Biol.* **18**: 675-689.

Ciechanover, A. 1994. The ubiquitin-proteasome proteolytic pathway. *Cell* **79**: 13-21.

Colby, C. and S.M. Williams. 1993. The distribution and spreading of rare variants in the histone multigene family of *Drosophila melanolochgaster*. *Genetics* **135**: 127-133.

Drexler, H.C.A. 1997. Activation of the cell death program by inhibition of proteasome function. *Proc. Natl. Acad. Sci. USA* **94**: 855-860.

Futterer, J. and T. Hohn. 1996. Translation in plants-rules and exceptions. *Plant Mol. Biol.* **32**: 159-189.

Genschik, P., Y. Parmentier, A. Durr, J. Marbach, M. C. Criqui, E. Jamet and J. Fleck. 1992. Ubiquitin genes are differentially regulated in protoplast-derived cultures of *Nicotiana sylvestris* and in response to various stresses. *Plant Mol. Biol.* **20**: 897-910.

Genschik, P., J. Marbach, M. Uze, M. Feuerman, B. Plesse and J. Fleck. 1994. Structure and promoter activity of a stress and developmentally regulated polyubiquitin-encoding gene of *Nicotiana tabacum*. *Gene* **148**: 195-202.

Graham, R.W., D. Jones and E.P. Candido. 1989. UbiA, the major polyubiquitin locus in *Caenorhabditis elegans*, has unusual structural features and is constitutively expressed. *Mol. Cell. Biol.* **9**: 268-277.

Hochstrasser, M. 1995. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Curr. Opin. Cell Biol.* **7**: 215-223.

Hourcade, D., D.R. Miesner, C. Bee, W. Zeldes and J. P. Atkinson. 1990. Duplication and divergence of the amino-terminal coding region of the complement receptor 1 (CR1) gene. An example of concerted (horizontal) evolution within a gene. *J. Biol. Chem.* **265**: 974-980.

Hu, Q. and H.R. Henney Jr. 1997. An Acanthamoeba polyubiquitin gene and application of its promoter to the establishment of a transient transfection system. *Biochim. Biophys. Acta* **1351**: 126-136.

Joshi, C.P. 1987. An inspection of the domain between putative TATA box and translation start site in 79

plant genes. *Nucleic Acids Res.* **15**: 6643-6653.

Leitch, L.J. and M.D. Bennett. 1997. Polyploidy in angiosperms. *Trends Plant Sci.* **2**: 470-476.

Ozkaynak, E., D. Finley, M.J. Solomon and A. Varshavsky. 1987. The yeast ubiquitin genes: a family of natural gene fusions. *EMBO J.* **6**: 1429-1439.

Park, S.M., J.H. Kim, Y.K. Lee, J.H. Pyee, C.I. Cheon and C.B. Hong. 1996. Nucleotide sequence of a ubiquitin-extension protein gene in *Nicotiana tabacum* and its expression pattern upon heat shock. *Mol. Cells* **7**: 125-130.

Park, S.M., J.H. Kim and C.B. Hong. 1998. Structure of a polyubiquitin gene in *Nicotiana tabacum*. *Mol. Cells* **8**: 107-112.

Pieterse, C.M., E.P. Risseeuw and L.C. Davidse. 1991. An in planta induced gene of *Phytophthora infestans* codes for ubiquitin. *Plant Mol. Biol.* **17**: 799-811.

Rothnie, H.M. 1996. Plant mRNA 3'-end formation. *Plant Mol. Biol.* **32**: 43-61.

Seo, J.W., J.H. Pyee, C.B. Hong and Y.N. Hong. 1996. Three classes of the heat-shock response of the ubiquitin gene family in *Nicotiana tabacum*. *Mol. Cells* **6**: 27-32.

Slabaugh, M.B., N.A. Roseman and C.K. Mathews. 1989. Amplification of the ribonucleotide reductase small subunit gene: analysis of novel joints and the mechanism of gene duplication in vaccinia virus. *Nucleic Acids Res.* **17**: 7073-7088.

Sun, C.W., S. Griffen and J. Callis. 1997. A model for the evolution of polyubiquitin genes from the study of *Arabidopsis thaliana* ecotypes. *Plant Mol. Biol.* **34**: 745-758.

Swindle, J., J. Ajioka, H. Eisen, B. Sanwal, C. Jacquemot, Z. Browder and G. Buck. 1988. The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. *EMBO J.* **7**: 1121-1127.

William, R.B. and J.E. Garbarino. 1996. The role of ubiquitin in plant senescence and stress responses. *Trends Plants Sci.* **1**: 331-335

Received April 13, 1998
Accepted May 29, 1998