Taiwania, **42**(1): 34-42, 1997

Isolation and Characterization of New Sporamin Gene Members from Sweet Potato (*Ipomoea batatas* Lam.)[1]

Jen-Chih Chen [2], Yih-Ming Chen [2] and Kai-Wun Yeh [2, 3]

(Manuscript received 26 October 1996; accepted 28 November 1996)

ABSTRACT: Two full-length and two partial cDNAs encoding sporamin A have been isolated from sweet potato tuberous roots. Sequence comparisons show that they are very similar with 94-98% homology at nucleotide level, and 80-88% at protein level. All four cDNAs possess multiple alternate polyadenylation signals in the 3' untranslated region (3'-UTR). Genomic Southern blot analysis indicates the presence of a sporamin multigene family in sweet potato. High levels of sporamin mRNAs were detected in developing tuberous roots, but they disappeared at the sprout-germinating stage. Differential expression of these genes was obvious as their mRNAs were present specifically in developing roots, rarely in stems and not in leaves.

KEY WORDS: **multigene family, sporamin, storage protein, sweet potato, 3'-untranslation region (3'-UTR).**

INTRODUCTION

Plants accumulate large quantities of storage proteins in seeds and tuberous organs during development. These proteins may function as nitrogen sinks to affect nutrient movement into kernels (Tsai, 1989) or serve as nutrition source of nitrogen, sulfur, and carbon for germinating seedlings and sprouts (Conlan *et al*., 1995; Staswick, 1990) . Storage proteins usually have several common features: (1) they are the most abundant proteins in the storage organ; (2) most of them have no enzymatic activity; (3) their synthese are under developmental control; (4) they are stored in subcellular structure (e.g. protein bodies); (5) they are rapidly degraded during seed germination or organ propagation; and (6) they are always encoded by a multigene family; therefore, they are typically composed of a group of structurally related polypeptides (Bevan, 1986; Conlan *et al*., 1995; Hattori *et al*., 1989). For these reasons, storage proteins offer an interesting model to study the mechanism of gene regulation in higher plants (Prat *et al*., 1990).

Sporamin, a tissue-specific storage protein (Hattori *et al*., 1985; Maeshima *et al*., 1986) , may account for 60-80% of the total soluble protein in the tuberous roots of sweet potato (*Ipomoea batatas* Lam.). It has a MW of approximately 22,000 daltons (Hattori *et al*., 1990), and is a mixture of closely related polypeptides encoded by a multigene family (Hattori *et al*., 1988). Although the synthesis of sporamin is tissue-specific, these proteins can be induced

1. The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers u17333 (spTi-1), u17334 (spTi-2), u17335 (spTi-3) and u17336 (spTi-4).

2. Department of Botany, National Taiwan University, Taipei 106, Taiwan, Republic of China.
3. Corresponding author. (Fax: 886-2-3622703. e-mail: ykwbppp@cc.ntu.edu.tw).

in the stem and leaf by exogenous supply of sucrose or other metabolic sugars to an excised plant (Hattori *et al*., 1990). These results suggested that the accumulation of sporamin is also regulated by the immediate need for storage, other than a strict developmental control (Prat *et al*., 1990; Staswick, 1990). Cytological studies demonstrated that sporamin was synthesized as a precursor by membrane-bound polysomes. Post-translational processing was essential to remove an extra sequence of 35 or 37 amino acid residues from N-terminal, and the mature sporamin was transported from endoplasmic reticulum to store in vacuole via Golgi apparatus (Nakamura *et al*., 1993; Schroeder *et al*., 1993). The content of sporamin decreased by 30-40% during the early stages of germination. As the sprouts continued to grow, the protein reduced gradually (Lin, 1987), which is characteristic of storage proteins. Although several sporamin cDNAs and genes have been isolated by Hattori *et al* (Hattori and Nakamura, 1988; Hattori *et al*., 1985; Hattori *et al*., 1989) and Wang *et al*. (1995), physiological functions of these proteins and the size of gene members are still unknown. Since the expression of sporamin gene appeared to be closely related to the organ development, as a first step to unveil it's physiological roles, efforts were made to further characterize sporamin gene family.

MATERIALS AND METHODS

Plant Materials

Sweet potato (*Ipomoea batatas* CV. Tainong 57) was grown at the Experimental Field of National Taiwan University for Sample collections.

DNA and RNA Extraction

Tubers of sweet potato weighted with 20 ~ 50g were freshly harvested from field and instantly ground into fine powder. The standard method was performed to isolate genomic DNA following Sambrook *et al*. (1989). For RNA isolation, a rapid and efficient method was used (Yeh *et al*., 1991)

Southern, Nothern blot and cDNA screening

Southern and Nothern gel blot were performed following the standard protocol (Sambrook *et al*., 1989). cDNA library was constructed in λ gt11 generated from sweet potato poly $(A)^{+}$RNA. A sp-B cDNA, a putative sporamin-antisense gene (unpublished data), was employed as prob to screen sporamin cDNA clones. Random primer labeling kit (Promega) was used for probe labeling.

RESULTS AND DISCUSSION

After screening of a λ gt11 cDNA prepared from sweet potato tuberous roots, using a ^{32}p-labeled cDNA fragment corresponding to a sporamin antisense gene from sweet potatoes (unpublished data), four cDNA clones (spTi-1 to 4) were isolated. The cDNA inserts were ca. 880, 860, 720 and 550 bp long, respectively. The sequence of spTi-1and 2 included a complete open reading frame, while those of spTi-3 and 4 contained partial ORF. DNA gel blots from sweet potato genomic DNA digested with *EcoR* I, *Sst* I, *Hind* III, *Kpn* I or *Pst* I

and hybridized under a high stringency with spTi-1 cDNA probe, revealed the presence of 10 more hybridizing fragments in some digestions. This result indicates that a large number of sporamin genes constitutes a multiple gene family in sweet potato (Fig. 1).

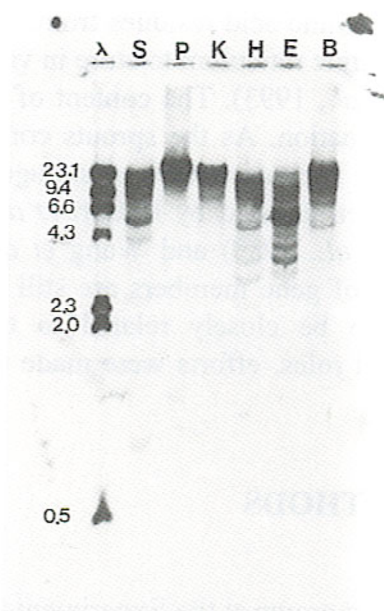

Fig.1. Genomic Southern blot analysis of spoamin genes in sweet potato. Genomic DNA, 10ug each, was digested with a restriction enzyme and subsequently resolved on 0.8% agarose gel, blotting and hybridizing with ^{32}p-labelled sporamin cDNA. S: *Sac* 1, P: *Pst* 1, K: *Kpn* I, H: *Hind* III, E: *EcoR* 1, B: *Bam* HI.

The complete nucleotide sequence of spTi-1, 2, 3 and 4 clones were shown in Fig. 2. Sequence comparisons show that spTi-1 is nearly identical with pIM0335 published by Hattori *et al.* (1988), and the other three clones, spTi-2, 3 and 4 are obviously new members of sporamin genes. These four clones share 94 - 98% of homology among themselves and with those of sporamin subfamily A gene members, including pIM023, and gSPOA1 (Hattori and Nakamura, 1988; Hattori *et al*., 1989) as well. However, the homology of these four clones with subfamily B members, including pIM0336, pIM0553 and gSPOB1 (Hattori and Nakamura, 1988), are 79 - 82%. These results suggest that spTi-1, 2, 3 and 4 belong to sporamin subfamily A. Analysis of the 3'-UTR sequence of the four cDNA clones and pIM0335 clone notably showed a prominent similarity among these clones with the exception of few base deletions and substitutions. Though full identity was found between ORF of spTi-1 and pIM0335, slight difference was present in their 3'-UTR region, with the addition of extra sequence TTTAATTCTCC to spTi-1 in comparison to pIMO335 (Fig. 3).

The sequence alignment shows that three or four polyadenylation recognition consensus sequence AATAAA or AATAAG (Dean *et al.*, 1986) repeatedly occurred at the 30 to 100 nucleotides upstream of the poly(A) addition site in all four sporamin genes. A G/T cluster, TGTGTTTGT (similar to a signal of YGTGTTYY in mammalian cells), was found immediately distal to the poly(A) site of spTi-1, 2, 3 cDNA genes. This consensus sequence was identified to have a function in the RNA processing event in viral and mammalian mRNA formation (McDevit *et al*., 1984; McLauchlan *et al*., 1985). It is interesting to see that another conserved plant polyadenylation signal AATAAT, which was found in all of the rbcs genes family (Dean *et al*., 1986), was also present in spTi-1 and pIMO335 (Fig. 3). Although the significance of sequence diversity is still unclear, it may offer a flexibility for processing and polyadenylation of the sporamin multiple gene family. The variability in the processing and polyadenylation of mRNA may affect the mRNA stability and expression level (Dean *et al*., 1986).

```
spTi-1   aattaaacatcattacctcttcgcttt.ctcccaattaaggttgtcatct
spTi-2      -....-------------------t-------....-.--------.

               1                                        44
spTi-1   gccaccATGAAAGCCCTCACACTGGCACTCTTCTTAGCCCTTTCCCTCTA
spTi-2   ......---------T----------T-----------T-----------

         45                                               94
spTi-1   TCTCCTCCCCAATCCCGCCCATTCCAGGTTCAATCCCATCCGCCTCCCCA
spTi-2   ---------------A----------------------------------

         95                                              144
spTi-1   CCACACACGAACCCGCCTCCTCTGAAACTCCAGTACTGGACATCAACGGC
spTi-2   -------------------------------------C------------
spTi-3                                                   --

         145                                             194
spTi-1   GACGAGGTCCGCGCCGGCGGGAACTACTACATGGTCTCCGCCATATGGGG
spTi-2   --------------------------------------------------
spTi-3   --------------------------------------------------

         195                                             244
spTi-1   AGCCGGCGGGGAGGGCTAAGACTCGCCCACTTGGACATGATGTCCAAAT
SPTi-2   --------------------------------------------------
SPTi-3   --------------------------------------------------

         245                                             294
spTi-1   GCGCCACGGACGTCATCGTATCCCCCAACGACTTAGACAACGGCGACCCC
spTi-2   --------------------------------------------------
spTi-3   ------------------------------------------------T-
spTi-4                                                    -

         295                                             344
spTi-1   ATCACCATCACGCCGGCGACGGCCGACCCGGAATCCACCGTGGTCATGGC
spTi-2   --------------------------------------------------
spTi-3   --------------------------------------------------
spTi-4   --------------------------------------------------

         345                                             394
spTi-1   GTCGACGTACCAGACTTTCCGGTTCAACATCGCCACCAACAAGCTCTGCG
spTi-2   ---------------------------------------------G----
spTi-3   ------------------------------------------A--G----
spTi-4   --------------------------------------------------
         395                                             444
SPTi-1   TGAACAACGTGAACTGGGGAATCCAGCACGACAGCGCGTCCGGGCAGTAT
SPTi-2   --------------------------------------------------
SPTi-3   -A------------------------T---------------T-----C-
SPTi-4   --------------------------------------------------
```

```
            445                                               494
spTi-1      TTCCTGAAAGCCGGCGAGTTTGTGTCCGACAATAGCAACCAGTTCAAGAT
spTi-2      -----------------------C--------------------------
spTi-3      --G--------------------C--------------------------
spTi-4      --------------------------------------------------

            495                                               544
spTi-1      TGAGCTGGTGGATGCCAACCTTAACTCCTACAAACTCACTTACTGTCAGT
spTi-2      ----G---------------------T-----------------------
spTi-3      ----G-T-------------------T--------------------T--
spTi-4      --------------------------T-----------------------

            545                                               594
spTi-1      TCGGCTCCGATAAATGCTACAACGTCGGCAGATTCCACGACCACATGTTG
pGRM-TIB    ----------C-------------------------------C-------
spTi-3      ----------C------------------------A-A----C-------
spTi-4      --------------------------------------------------

            595                                               644
spTi-1      AGGACCACGCGTTTGGCTCTCTCCAATTCTCCCTTCGTTTTTGTCATCAA
spTi-2      -----------C--------------------------------------
spTi-3      ------------------------TC--A-------T-------G-----
spTi-4      --------------------------------------------------

            645            660
spTi-1      ACCTACCGATGTGTAAtgtaacactgaaaagcgccggttatgaggttgca
spTi-2      ---------------------------------.----------------
spTi-3      --------T-------.....--g-------t------------------
spTi-4      --------------------------------------------------

spTi-1      tggtagctatgcaacgttgccactttgacaacgttgtacgtgtaagaata
spTi-2      ---------------------.----------------------------
spTi-3      --t-----------------------------------------------
spTi-4      --------------------------------------------------

spTi-1      aacatgcaacaaatccgagctggtatggttgtgtaaatcctaaataaatc
spTi-2      ----------------a---------------------------------
spTi-3      --------------------g-----------------------------
spTi-4      ----------------a---------------------------------

spTi-1      cgaagaaataataa...ggataaaatattatcctgtgtttgttttaattcttc
spTi-2      t-----...-----...----------------------------
spTi-3      t-----...-----taa--------------------------------
spTi-4      t-----...-----...------------
```

Fig. 2. Sequence aligment of four cDNA clones, spTi-1, spTi-2, spTi-3 and spTi-4. Among them, Ti-1 and Ti-2 are full length. Dash "-" indicates similar nucleotide base, and dot "." indicate deletion.

```
              v10       v20       v30       v40       v50
spTi-1    taaTGTAACACTGAAAAGCGCCGGTTATGAGGTTGCATGGTAGCTATGCA
IM0335    taaTGTAACACTGAAAAGCGCCGGTTATGAGGTTGCATGGTAGCTATGCA
spTi-2    taaTGTAACACTGAAAAGCGCCGGTTATGAGGTTGCATGGTAGCTATGCA
spTi-3    taa.....cAgtGAAAAGtgCCGGTTATGAGGTTGCATGttAGCTATGCA
spTi-4    taaTGTAACACTGAAAAGCGCCGGTTATGAGGTTGCATGGTAGCTATGCA

              v60       v70       v80       v90       v100
spTi-1    aCGTTGCCACTTTGACAACGTTGTACGTGTAAGAATAAACATGCAACAAA
IM0335    aCGTTGCCACTTTGACAACGTTGTACGTGTAAGAATAAACATGCAACAAA
spTi-2    aCGTTGCC-cTTTGACAACGTTGTACGTGTAAGAATAAACATGCAACAAA
spTi-3    aCGTTGCCACTTTGACAACGTTGTACGTGTAAGAATAAACATGCAACAAA
spTi-4    aCGTTGCCACTTTGACAACGTTGTACGTGTAAGAATAAACATGCAACAAA

              v110      v120      v130      v140      v150
spTi-1    tCCGAGCTGGTATGGTTGTGTAAATCCTAAATAAATCCGAAGAAATAATA
IM0335    tCCGAGCTGGTATGGTTGTGTAAATCCTAAATAAATCCGAAGAAATAATA
spTi-2    tCCaaGCTGGTATGGTTGTGTAAATCCTAAATAAATCtgAAGA---AATA
spTi-3    tCCGAGCggGTATGGTTGTGTAAATCCTAAATAAATCtgAAGAAATAATA
spTi-4    tCCaaGCTGGTATGGTTGTGTAAATCCTAAATAAATCtgAAGA---AATA

              v160      v170      v180
spTi-1    aGGATAAAATATTATCCTGTGTTTGTTTTAATTCTCC(A)n
IM0335    aGGATAAAATATTATCCTGTGTTTGT(A)n
spTi-2    aGGATAAAATATTATCCTGTGTTTGTTTT(A)n
spTi-3    aGGATAAAATATTATCCTGTGTTTGTTTT(A)n
spTi-4    aGGATAAAATATT(A)n
```

Fig. 3. Sequence analysis of 3'-untranslated region among four pSPTi clones and pIM0335 (Nakamura,1989). All colnes show alternate polyadenylation signal in 3'-UTR sequence. The consensus polyadenylation recognition sequences AATAAA, AATAAT, and AATAAG were singly underlined. G/T cluster was doubly underlined.

In order to study the expression of sporamin in various tissue, sweet potato plants (*Ipomoea batatas* Lam. cv. Tainong 57) were grown at Agronomy farm of National Taiwan University for 4 months. Total cell RNA was isolated by the method of Yeh *et al.* (1991) from leaves, stems, and developing tuberous roots for northern blot analysis. Developing tuberous roots were separated into rooting tissue, young, mature and germinating tuberous organ. The sporamin genes were highly expressed during root developement. As sweet potato grew into tuberous form, the sporamin genes continued to increase their expression, and reached a maximum level at the mature stage. However, during germination these genes ceased to express (Fig 4a). Similarly, no signal was detected in leaves and only little in stems (Fig. 4b). These results indicate that sporamin genes is tissue-specific and is only expressed in the tuberous organ.

In summary, we present four new members of sporamin genes from sweet potato tuberous roots. Sequences analysis suggest that these cDNA colnes may be classified into sporamin gene subfamily A. DNA sequence comparisons show that they share 94-98% sequence homology among these four cDNA clones and with previously isolated cDNA/genomic clones, e.g. pIM023, pIM0335, gSPOA1 (Hattori and Nakamura, 1988) and

gSPOR 5-31 (Wang *et al.*, 1995). The 3'-UTR sequence of these genes displays the occurrence of alternate polyadenylation event and processing in the mRNAs formation. Therefore, it may be postulated that the level of gene expression and mRNA stability among the multigene members may vary in the tuberous roots of sweet potato. Furthermore, the expression is root-specific, and closely related with the development of tuberous roots.

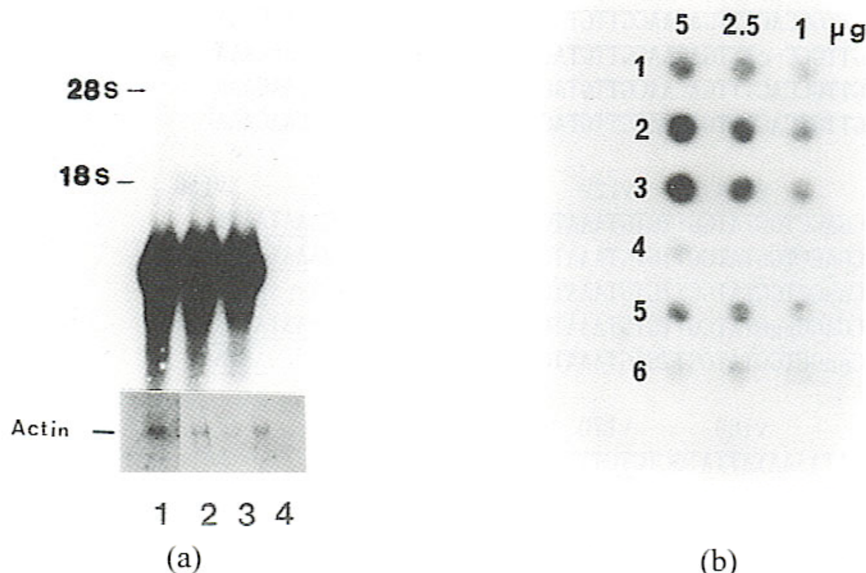

Fig. 4. Tissue-specific and developmental expression of sporamin genes. (a) Twenty *u*g of total RNA extracted from tuberous roots at different developmental stages were electrophoresed, blotted and hybridized with sporamin cDNA. A hybridizing signal by action probe served as the control for loading was shown at the bottom. (b) Different quantities of RNA (5, 2.5 and 1 μg) extracted from different sweet potato tissues were dot blotted and hybridized with the ^{32}P-labelled sporamin cDNA. Roots (1), young tuberous roots (2), mature tuberous roots (3), sprout-germinating tuberous roots(4), stems (5) and leaves (6).

ACKNOWLEDGEMENT

This work was supported, in part by National Science Council, ROC, to K.-W. Yeh under the grant of NSC 84-2311-B-002-050 B14.

LITERATURE CITED

Bevan, M., R. Barker, A. Goldsbrough, M. Jarvis, T. Kavanagh and G. Iturriga. 1986. The structure and transcription start site of a major potato tuber protein gene. Nucl. Acids Res. **114**: 4625-4638.

Conlan, R. S., L. A. Griffiths, J. A. Napier, P. R. Shewry, S. Mantell and C. Ainsworth. 1995. Isolation and characterization of cDNA clones representing the genes encoding the major tuber storage protein (dioscorin) of yam (*Dioscorea cayenensis* Lam.). Plant Mol. Biol. **28**: 369-380.

Dean, C., S. Tamaki, P. Dunsmuir, M. Favreau, C. Katayama, H. Dooner and J. Bedbrook. 1986. mRNA transcripts of several plant genes are polyadenylated at multiple sites *in vitro*. Nucl. Acids Res. **14**: 2229-2241.

Hattori, T. and K. Nakamura. 1988. Gene coding for the major tuberous root protein of sweet potato: Identification of putative regulatory sequence in the 5' upstream region. Plant Mol. Biol. **11**: 417-426.

Hattori, T., T. Nakagawa, M. Maeshima, K. Nakamura and T. Asahi. 1985. Molecular cloning and nucleotide sequence of cDNA for sporamin, the major soluble proteinof sweet potato tuberous roots. Plant Mol. Biol. **5**: 313-320.

Hattori, T., N. Yoshida and K. Nakamura. 1989. Structural relationship among the members of a multigene family coding for the sweet potato tuberous roots. Plant Mol. Biol. **13**: 563-572.

Hattori, T., S. Nakagawa and K. Nakamura. 1990. High-level expression of tuberous root storage protein genes of sweet potato in stems of plantlets grown *in vitro* on sucrose medium. Plant Mol. Biol. **14**: 595-604.

Lin, Y.-H. 1987. Trypsin inhibitors of sweet potato: Review and prospect. in Recent Advances in Botany. In: Hsing, Y.-I. and C.-H. Chou. (eds.). Institute of Botany. pp. 179-185. 149. Academia Sinica, Taipei, R. O. C.

Maeshima, M., T. Sasaki and T. Asahi. 1986. Characterization of major proteins in sweet potato tuberous roots. Phytochem. **24**: 1899-1902.

McDevit, M. A., M. J. Imperial, A. Haseeb and J. R. Nevins. 1984. Requirement of a downstream sequence for generation of a poly(A) addition site. Cell **37**: 994-999.

McLauchlan, J., D. Gaffney, L. Whitton and B. Clements. 1985. The consensus sequence YGTGTTYY located downstream from the AATAAA signal is required for efficient formation of mRNA 3' termini. Nucl. Acids Res. **13**: 1347-1368.

Nakamura, K., K. Matsuoka, F. Mukumoto and N. Watanabe. 1993. Processing and transport to the vacuole of a precursor to sweet potato sporamin in transformed tobacco cell line BY-2. J. of Exp. Bot. **44**: 331-338.

Prat, S., W. B. Frommer, M. Hofgen, M. Keil, J. Kobmann, M. Koster -Topfer, X. J. Liu, B. Muller, H. Pena-Cortes, M. Rocha-Sosa, J. J. Sanchez-Serrano, U. Sonnewald and L. Willmitzer. 1990. Gene expression during tuber development in potato plant. FEBS Letters **268**: 334-338.

Sambrook, J., E. -F. Fritsch, T. Manniatis. 1989. Molecular cloning: A Laboratory Mannual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schroeder, M. R., O. N. Borkhsenious, K. Matsouka, K. Nakamura and N. V. Raikhel. 1993. Colocalization ofbarley lectin and sporamin in vacuole of transgenic tobacco plants. Plant Physiol. **101**: 451-458.

Staswick, R. E. 1990. Novel regulation of vegetative storage protein genes. Plant Cell **2**: 1-6.

Tsai, C.-Y. 1983. Genetics of storage protein in maize. Plant Breeding Rev. **11**: 103-1138. AVI Publishing Company Westport, CT, USA.

Wang, S.-J., C.-T. Lin, K.-C. Ho, Y.-M. Chen and K.-W. Yeh. 1995. Nucleotide sequence of a sporamin gene in sweet potato. Plant Physiol. **108**: 829-830.

Yeh, K.-W., R, H. Juang and J.-C. Su. 1991. A rapid and efficient method for RNA isolation from plants with high carbohrdrate content. Focus **13**: 102-103.

甘藷中 Sporamin 基因群的選殖及特性研究[1]

陳仁治[2]、陳益明[2]、葉開溫[2,3]

(收稿日期：1996年10月26日；接受日期：1996年11月28日)

摘　　要

本文中報告了四個 Sporamin cDNA 基因，其中有二個是全長的基因，另二個則不完整；此四個 cDNA 基因其序列與以往發表的皆不盡相同，但相似度則達 90%以上；根據 Gene Bank 的資料分析，四者皆屬 Sporamin A 亞群分子；從南方雜合的資料顯示，其含有相當多的子基因，北方轉印資料顯示，此基因在葉子部不表現，在莖部則微量存在，然而在塊根中則大量表現，故此基因被推論為扮演甘藷塊根中貯藏性蛋白質的功能。

關鍵詞：多基因型，sporamin 基因群，貯藏性蛋白質，甘薯，3'-尾終端不轉譯區 (3'-UTR)。

1. The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers u17333 (spTi-1), u17334 (spTi-2), u17335 (spTi-3) and u17336 (spTi-4).

2. 國立臺灣大學植物學系，臺北市106，臺灣，中華民國。

3. 通信聯絡員。