Isolation and characterization of cDNA clones encoding a novel subfamily sporamin B in sweet potato

JOLLANDA EFFENDY1,*, DON R. LABONTE2, DARDA EFFENDY3
1Department of Plant Breeding, Faculty of Agriculture, Universitas Pattimura. Jl. Ir. M. Putuhena, Poka Campus, Ambon 97233, Maluku, Indonesia. Tel.: +62-911-322499, *email: jeffendy@alumni.sfu.ca
2School of Plant, Environmental, and Soil Sciences, Louisiana State University Agricultural Center. 131 J.C. Miller Hall, Baton Rouge, LA 70803, USA
3Department of Agronomy and Horticulture, Faculty of Agriculture, Institut Pertanian Bogor. Jl. Meranti, IPB University Campus Dramaga, Bogor 16680, West Java, Indonesia


Abstract. Effendy J, Labonte DR, Efendi D. 2019. Isolation and characterization of cDNA clones encoding a novel subfamily sporamin B in sweet potato. Biodiversitas 20: 3033-3041. Three cDNAs (IbH23, IbH41, and IbH74) encoding sporamin of sweet potato, were isolated from a skinning injury cDNA library. The open reading frame consisted of 288 nucleotides and the deduced polypeptide sequences consist of 96 amino acids with missing 127 amino acid residues at 5’. The sequences shared identity [42-62% (IbH23 and IbH41) and 40-61% (IbH74)] to sporamins A and B previously identified in other sweet potato species. The IbSpors contained three of four signature cysteine residues of sporamin family (Cys135, Cys198, Cys209). In addition to similarity to sporamins, these cDNA clones also showed similarities to miraculin (29-35%); Kunitz trypsin inhibitor (29-36%); factor Xa inhibitor (32-38%); subtilisin inhibitor (26-37%); seed protein (32%); tumor-related protein (36% in IbH23 and IbH41); and [wound-responsive protein (30%), sexual organ responsive protein (29%), and proteinase inhibitor (26%)] only can be found in IbH74. The three partial mRNA transcripts showed induced and transient expression in response to skinning injury. The phylogenetic analysis revealed that they belonged to a new member of sporamin gene subfamily B and may play key roles in plant defense and wound repair in response to skinning injury in storage roots of sweet potato.

Keywords: Cysteine residues, Ipomoea batatas, miraculin, storage root, wounding

INTRODUCTION

Sweetpotato is the second most important storage crop after cassava (FAO 2016). World sweet potato production is estimated at 130 Mt (FAO 2016). It is the main staple food in 82 countries, in which almost all of them are developing countries (FAO 2016). Indonesia is the world’s fifth-largest sweet potato producer in the world with total production of 2.270 Mt and total harvested area of 0.137 MHa with an average yield of 16.6 ton/ha (FAO 2016). The high starch content in sweet potato fulfills calorie needs in many countries. As valuable starch reserves, however, post-harvest loss was huge due to wound injury especially in underprivileged countries (Booth and Burton 1983; Rees et al. 2003).

Storage substances such as starch, lipids, and proteins are accumulated at different stages of plant development (Fujitava 2002). They can be found in vegetative and reproductive tissues and function as reserves for further use (Fujitava 2002). Storage proteins function as biological reservoirs for nitrogen, sulfur, and carbon (Shewry 2003). They are usually found in seed and tuberous roots of plants species to provide nutrients to support the growth of new plants (Shewry 2003). Vegetative and reproductive tissues highly accumulate proteins in protein bodies, so that they do not interfere with other cellular functions (Shewry 2003). Examples include patatin in potato tuber (Dutt et al. 2014), dioscorins in yam (Xue et al. 2015), glutelin in wheat (Scherf et al. 2016), ocatin in Andean tuber crop oca (Flores et al. 2002) and sporamin in sweet potato (Effendy et al. 2013; Effendy et al. 2017; Ponniah et al. 2017).

Sporamin, a vacuolar storage protein, is exclusively found in storage roots of sweet potato (Hattori et al. 1989). It is a major storage protein and represents 60-80% of the total soluble protein of sweet potato storage roots (Maeshima et al. 1985). It shares significant amino acid sequence identity with some Kunitz-type trypsin inhibitors (Hattori et al. 1991). There are two classes of sporamins, sporamin A (spor A) and sporamin B (spor B) (Hattori et al. 1985). Based on electrophoresis analysis and the intensity of bands, Maeshima et al. (1985) found that the ratio of spor A to spor B content was 2:1. They showed that the molecular weight of Spor A was 31000 Da and spor B was 22000 Da (Maeshima et al. 1985).

Previous studies showed that sporamin belongs to a multigene family (Hattori et al. 1989). Various sporamin genes have been isolated from many Ipomoea species (Altschul et al. 2005). BLASTX search showed that Spor A genes have been isolated from Ipomoea batatas, I. trifida, I. leucantha, and I. nil, while spor B genes have been isolated from I. batatus and I. nil only (Altschul et al. 2005).

Sporamin has attracted the attention of an increasing number of researchers due to its role against pests and herbivores in the vegetative tissues of plants (Yeh et al. 1997; Ding et al. 1998; Rajendran et al. 2014). The fact that...
sporamin levels are highly regulated in response to a wide range of biotic and abiotic stress responses showed the importance of this gene in stress-related responses (Senthilkumar and Yeh 2012; Effendy et al. 2013; Qiu et al. 2013; Rajendran et al. 2014; Effendy et al. 2017).

Here, we report the isolation and characterization of sweet potato partial cDNA clones encoding sporamin which is a new subfamily of spor B. This sporamin may play key roles in plant defense and wound repair in response to skinning injury in storage roots of sweet potato. The cDNAs sequences, similarity, and phylogenetic analysis among related genes are discussed in this paper.

### MATERIALS AND METHODS

#### Plant materials

Freshly harvested storage roots of sweet potato cultivar LA 07-146 were washed, blot-dried, and carefully scraped with a razor scraper (Titan 11030; Star Asia-USA, Renton, WA) to remove the thin outer pigmented skin as detailed in Effendy et al. (2013). The roots were skinned and the skinned roots were peeled to the same thickness (1.2 mm) at 0 (control), 2, 4, 8, and 12 h after skinning and the peels were immediately frozen in liquid nitrogen and stored at -80 °C for RNA extraction. Three independent roots were used for each time points as replicates.

#### RNA isolation, cDNA preparation, and ACP-based gene-fishing PCR

Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Valencia, CA) as outlined in manufacturer’s instructions (Effendy et al. 2013; 2017). The quantity and quality of the total RNA were determined using a spectrophotometer (ND-1000; Nanodrop Technologies, Wilmington, DE). An aliquot of 2 µg RNA of three biological replicates extracted from storage root tissues for each time point was pooled to capture rare transcripts that may have escaped in one of the replicates as a result of technical errors during gene fishing experiment. Also, 3 µg RNA extracted from storage roots at 8 and 12 h were pooled together to make the sample size of four for convenience with PCR handling during gene fishing.

First-strand cDNA synthesis was performed using a Gene- Fishing™ DEG premix kit (Seegene, Rockville, MD) as previously described (Effendy et al. 2013). Second strand cDNA synthesis and subsequent PCR amplification were performed in a single tube as previously described (Effendy et al. 2013). Twenty ACP primers were used during the second-stage PCR and subsequent PCR amplification were performed in a single tube using the protocol as detailed in the GeneFishing™ DEG premix kit manual (Seegene, Rockville, MD) in present study to capture the DEGs.

#### Cloning of Differentially Expressed Genes (DEGs)

Based on presence/absence or relative intensity of ACP-PCR products between control and skinning samples, the fragments corresponding to DEGs were excised and extracted following the manufacturer’s instruction (QIAquick® Gel Extraction Kit Qiagen, Valencia, CA). The DEGs were cut from the gel and cloned into pGEM®-T Easy vector (Promega, Madison, WI). The positive colonies from DEGs were confirmed by colony PCR using M13F and M13R primers. Plasmids isolated from these clones were single-pass sequenced with T7 primer in an ABI 3730x1 genetic analyzer according to the manufacturer’s instruction.

#### Nucleotide and deduced amino acid sequencing analyses of DEGs

DNA sequences were processed manually to remove the vector backbone and the poly (A) tail and searched against the nonredundant nucleotide and protein database of NCBI using BLASTN and BLASTX interface (http://www.ncbi.nlm.nih.gov/BLAST). Three positive cDNAs clones showed similarity to a storage protein sporamin were named IbH23, IbH41, and IbH74 were used for further analysis. A total of 100 hits from BLASTX search that showed similarities to IbH23, -41, and -74 were pooled based on the type of genes and the number of occurrence of each gene, and for Ipomoea sp. also based on the type of sporamins. Alignment and comparison of sequences were made using the ClustalW program (http://www.ebi.ac.uk/clustalw) to find the homologs and to predict the functions of the differentially expressed genes/gene products to generate multiple sequence alignment in which three samples were chosen to represent each gene.

Open reading frame and protein prediction were made using the NCBI ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Conserved amino acid sequences were indicated by BOXSHADE sequences [http://www.ch.embnet.org/software/BOX_form.html]. The unrooted phylogenetic tree was generated from the deduced amino acid sequences for IbSpor cDNA clones and other Spor homologs from other Ipomoea species and other genes from other plants that showed similarities by using the bootstrap method with 1,000 bootstrap iterations. This tree represents the branching order but does not indicate the root or location of the last common ancestor. To construct the phylogenetic tree, the .dnd file generated in ClustalX2 was displayed in FigTree v1.4.3.

### RESULTS AND DISCUSSION

#### Isolation and analysis of candidate partial cDNAs using ACP GeneFishing

To examine the effect of skinning injury on storage root RNA population and to determine how rapidly the changes in mRNA population occurred following the onset of a skinning injury, time-course experiments were carried out. RNA was isolated from freshly harvest storage roots exposed to skinning injury at 0, 2, 4, and combined (8 + 12) h and used for ACP-PCR. IbH23 and IbH74 showed higher expression at 8 + 12 h, while IbH41 showed a high expression level at 2 h following the onset of skinning injury treatment (Figure 1 and Table 1). Under non-stress conditions, sporamin is constitutively expressed in storage conditions.
root of sweet potato (Wang et al. 2002). In this study, ACP-PCR products of IbH23, -41, and -74 have different peaks in response to skinning injury. These results showed that upon the onset of skinning injury, sporamin mRNA transcripts were transiently and up-regulated. Rapid induction of spor at 2 h in IbH41 might indicate its defense role in wound healing. This study and others showed that spor was differentially expressed in leaves and roots in response to wounding (Yeh et al. 1997; Effendy et al. 2013; Effendy et al. 2017), abiotic (Senthilkumar and Yeh 2012) and biotic (Qiu et al. 2013) stresses, suggesting that spor may be a stress-responsive gene. These expression patterns at the transcriptional levels revealed that the complex regulatory mechanism of sporamin was modulated by environmental cues. Therefore, these cDNAs could be considered as key target genes to be used in genetic engineering to study molecular evolution, regulatory mechanisms and physiological functions in plants.

Isolation of three cDNAs encoding sporamin from storage roots of sweet potato

Three ACP primers used in the experiment demonstrated changes in the amplification of DEGs in storage roots after skinning injury at 0, 2, 4, (8 + 12) h. Three cDNA clones were isolated and their sequences were determined. The cDNA products of interest were excised, re-amplified and then cloned. The insert of each cDNA clone was sequenced. Each of the cDNA sequences was submitted to the NCBI server for comparison to sequences available in the GenBank nucleotide and expressed sequence tag (EST) databases using BLAST programs. The IbH23, -41, and -74 cDNAs contained 388 bp, 414 bp, and 414 bp of nucleotides, respectively, each containing an open reading frame of 288 bp (Figure 2). Thus, the IbH23, -41, and -74 cDNAs covered partial length of the sporamin mRNA sequences. Based on their percent identity, these three clones shared high identity among themselves. IbH23 and IbH41 shared 99% identity; IbH23 and IbH74 shared 92% identity; and IbH41 and IbH74 shared 94% identity (Figure 2).

Figure 1. ACP-PCR products of mRNA isolated from non-treated (0 h), and skinning injury treatment for 2, 4, and (8 + 12) h of sweet potato storage roots. ACP-PCR was performed with primers ACP17, ACP18, and ACP19. Arrows marked indicate ACP-PCR products that were excised from the gel and cloned. Differential display of sweet potato cDNA GeneFishing showed amplification products of IbH23, IbH41, and IbH74 using ACP arbitrary primers ACP17, ACP18, and ACP19, respectively at 0, 2, 4, and (8 + 12) h after skinning treatments. ACP: annealing control primer; M: DNA ladder marker. Arrows indicate differential amplified cDNA bands.

Table 1. Accumulation of ACP-PCR products corresponding to RNA in response to skinning injury treatment in sweet potato

<table>
<thead>
<tr>
<th>cDNA</th>
<th>Type</th>
<th>0 h</th>
<th>2 h</th>
<th>4 h</th>
<th>(8 + 12) h</th>
</tr>
</thead>
<tbody>
<tr>
<td>IbH23</td>
<td>Skinned up-regulated</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>IbH41</td>
<td>Skinned transient regulated</td>
<td>+</td>
<td>+++</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>IbH74</td>
<td>Skinned up-regulated</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
</tbody>
</table>

Type up-regulated and transient regulated correspond to groups of ACP-PCR products with different accumulation patterns. ’+’, ’+++’ and ’++++’ denotes the presence and relative intensity of ACP-PCR products on agarose gels.
The sequence alignment showed that polyadenylation recognition consensus sequences (AATAAG) occurred at 3′ for IbH23, -41 for IbH41 and IbH74; while sequences (AATAAA) occurred at 63 for IbH41 and IbH74.

Analysis of the 3′ untranslated region (UTR) region of the three cDNA clones and a sporamin B subfamily pIMO336 (Hattori et al. 1990) definitely showed a low (27/96) and medium (54/100) identity among these clones without and with the exception of the deletions and substitutions of several bases, respectively (Figure 3). The 3′ UTR of IbH23, -41, and -74 with pIMO336 encoding spor B from sweet potato showed a shifting stop codon (taa) in IbH23, -41, and -74 when compared with pIMO336 (Murakami et al. 1986).

The sequence alignment showed that polyadenylation recognition consensus sequences (AATAAA) occurred at 37 for IbH23; 20 for IbH41 and IbH74; while sequences (AATAAG) occurred at 63 for IbH41 and IbH74.

Figure 2. Sequence alignment of three partial cDNA clones IbH23, IbH41, and IbH74 in response to skinning injury in sweet potato. Star (*) represented same nucleotides, empty space () represented 1/3 nucleotide differences and dash (-) represented deletions.

Figure 3. Sequence analysis of 3′-untranslated region among three partial cDNA clones of sweet potato in response to skinning injury and pIMO336 from sweet potato. All sweet potato clones showed putative eukaryotic polyadenylation signal in 3′-UTR sequence. The consensus polyadenylation recognition sequences AATAAA and AATAAG were singly and double underlined, respectively.
Similarity of sporamins with miraculin, KTIs, and other genes

Sporamin is exclusively found in storage roots of sweet potato (Hattori et al. 1989). Spors A and B have been isolated from Ipomoea species. A total of 100 hits were retrieved by BLASTX search. Our partial cDNAs showed that spors A and B were greatest (57/100). The highest percentage occurrence of spor A has been isolated from I. batatas (SporA_Ibat: 22/100) followed by I. trifida (SporA_Itri: 7/100), I. leucantha (SporA_Ileu: 3/100) and the lowest occurrence in I. nil (SporA_Inil: 1/100) were distributed equally in IbH23, -41, and -74. In contrast, spor B was only found in I. batatas (SporB_Ibat: 11/100) and I. nil (SporB_Inil: 13/100).

Besides spors A and B, the second most abundant gene that showed similarity to IbH23, -41, and -74 was miraculin (mirac). The highest occurrence was found in IbH23 and IbH41 with the equal amount of Mirac (15/100) and less than that of IbH23 and IbH41 in IbH74 (6/100 occurrence) (Figure 4). Mirac has been studied in eggplants infested with Colorado potato beetle (López-Galiano et al. 2017) and citrus plants under insect infestation and salinity stress (Podda et al. 2014). The third most abundance gene that showed similarity to IbH23, -41, and -74 was Kunitz trypsin inhibitors (KTInhib) with the occurrence being similar in IbH23 and IbH41 (7/100) and a slightly higher in IbH74 (8/100). Another was Factor Xa Inhibitor (FXaInhib), with an occurrence a slightly higher in IBH74 (3/100) while in IbH23 and IbH41 the occurrence was less (2/100). Another type of inhibitor was a proteinase inhibitor (ProtInhib) exclusively found only in IbH74 (Figure 4). In this study IbH23, -41, and -74 showed similarities with three inhibitors such as KTInhib (Guerra et al. 2015; Chan et al. 2017), FXaInhib (Salu et al. 2014), subtilisin inhibitor (CLSInhib) (Bunyatang et al. 2016; Yu et al. 2017), and proteinase inhibitor (Chen et al. 2014; Kidrič et al. 2014; Fischer et al. 2015; Shamsi et al. 2016). IbH23, -41, and -74 showed 29-36% identity with KTInhib which are commonly found in seeds of Leguminoseae family. These inhibitors are toxic to insects, pests, and pathogens (Cruz et al. 2013), and function in regulating endogenous proteolytic activities in storage organs as well as a defense against mechanical wounding and insects, pathogens and herbivory attacks (Macedo et al. 2016). Only IbH74 showed similarities with these three inhibitors, while IbH23 and IbH41 only showed similarities with two inhibitors (Figure 4).

The three partial cDNA has equal amount of CLSI gene (CLSInhib:1/100). This study found that tumor-related protein (TRProt) gene was specific to IbH23 and IbH41, while genes such as albumin (Album) (Bhunia et al. 2014), wound-responsive protein (WRProt) (Wong et al. 2013), and sexual organ-responsive protein (SRProt) were exclusively found only in IbH74 (Figure 4). Therefore, these cDNAs especially IbH74 should be studied further to reveal its potential roles in protecting crop plants against insects, pathogens, herbivores, wounding as well as its function as a storage protein.

Sporamin has four signature cysteine residues, for example, Cys85, Cys135, Cys198, and Cys209. Multiple sequence alignment shows that the IbH23, -41, and -74 cDNAs are missing the first 126 amino acid residues. The reason is that it requires an effective reverse transcriptase to the extreme 5’ end of the mRNA. Therefore, IbH23, -41, and -74 only have the last 3 Cys regions (Cys135, Cys198, and Cys209) of sporamin. Although spors A and B have four Cys residues, the significant difference between spors A and B from Ipomoea sp. with other genes from other plants was the Cys85 and Cys84 positions, respectively.
Figure 5. Alignment of the deduced amino acid sequences of partial cDNAs of *IbH23*, *IbH41*, and *IbH74* encoded sporamin from *Ipomoea batatas* (Ip); spors A and B from *I. batatas* (Ib), *I. leucantha* (Il), *T. trifida* (It); miraculin (Mirac) from *Camelina sativa* (Cs), *Nicotiana tabacum* (Nt), *Theobroma cacao* (Tc); KTs from *Glycine max* (Gm), *Arachis duranensis* (Ad), *Arabidopsis thaliana* (At); subtilisin inhibitor from *Canavalia ensiformis* (Cn); tumor-related protein (TRP) from *Nicotiana glauca* x *Nicotiana langsdorffii* (NggXN); and albumin (Album) from *Theobroma cacao*. The corresponding GenBank accession numbers are noted at the end of each line. Cysteine residues are shown as black bold letters on a gray background.
Phylogenetic relationship among IbSpors cDNAs and Spors from Ipomoea species, miraculin, KTIs and other genes from various plant species

In the present study, IbH23 and IbH41, as well as IbH74, were used to study their phylogenetic relationship with sporamins from ESTs available in databases. An unrooted phylogenetic tree was created in order to determine the relationship between the sporamins from Ipomoea sp., and other related genes from other plant species. As shown in Figure 6, the spors proteins from various Ipomoea sp. could be divided into three distinct classes (I to III). Class I contained spor A, Class II contained spor B and Class III contained IbH23, -41, and -74. Other genes except spors were categorized in different classes. The tree shown in Figure 6 indicated that a distinction can be made between spors A and B from Ipomoea sp. compared with IbH23, -41, and -74 which appears to be less related. These results were supported by the BLASTX search in which the sequences shared identity of 42-62% (IbH23 and IbH41) and 40-61% (IbH74) to spors A and B previously identified in other sweet potato species. Based on these low to medium shared identity with spors A and B in other Ipomoea sp., these IbH23, -41, and -74 could be placed in a novel spor B subfamily.

**Figure 6.** Phylogenetic relationship among IbSpors (IbH23, IbH41, and IbH74) (Class III with an \[-\] and Spors from Ipomoea sp. and other genes from other plant species. The amino acid sequences were aligned by the ClustalX2 program, and the neighbor-joining tree was drawn using FigTree v1.4.3. The corresponding GenBank accession numbers are noted in the phylogenetic tree.
From phylogenetic analysis, multiple alignment and domain analysis of IbH23, -41, and -74 in sweet potato in response to skinning injury, we concluded that spor A family members may be conserved among *I. batatas*, *I. leucantha*, *I. trifida* and *I. nil*, while spor B were equally conserved in *I. batatas* and *I. nil*. The difference in Cys residues at Cys85 position may affect protein disulfide bonding that could lead to different classes in phylogenetic relationship between spors A and B with other genes. Our results suggested that IbH23, -41, and -74 are presumably new members of spor genes subfamily B. The phylogenetic analysis together with the Cys domain motif analysis presented here will facilitate the functional annotation and study of sweet potato sporamins.

In conclusion, DNA sequence comparisons show that these three cDNAs share 92-99% identity among themselves and 40-62% identity with spors A and B. ACP-PCR products showed skinned up-regulated (IbH23 and IbH74) and transient regulated (IbH41) expressions. Amino acid and phylogenetic analysis suggest that these partial cDNA clones can be classified into a new sporamin gene subfamily B. The ORF contained three conserved cysteine residues that usually found in spors A and B, except for the first cysteine due to missing ORF, the first amino acid at 5′. The 3′-UTR sequence of these genes may have an impact on understanding the role of sporamin, not only function as a major storage protein in sweet potato but may be applied in other tuberous roots that spoiled easily due to postharvest physiological deterioration such as in cassava.

**ACKNOWLEDGEMENTS**

Funding for this work was provided by Borlaug Fellowship 2011 from the United State Department of Agriculture – Foreign Agricultural Service to JE.

**REFERENCES**


