J Plant Biotechnol (2022) 49:39-45
DOI:https://doi.org/10.5010/JPB.2022.49.1.039

ISSN 1229-2818 (Print)
ISSN 2384-1397 (Online)

Research Article

Comparative analysis of AGPase proteins and conserved domains in sweetpotato (*Ipomoea batatas* (L.) Lam.) and its two wild relatives

Hualin Nie · Sujung Kim · Jongbo Kim · Suk-Yoon Kwon · Sun-Hyung Kim

Received: 9 February 2022 / Revised: 22 February 2022 / Accepted: 22 February 2022

Abstract Conserved domains are defined as recurring units in molecular evolution and are commonly used to interpret the molecular function and biochemical structure of proteins. Herein, the ADP-glucose pyrophosphorylase (AGPase) amino acid sequences of three species of the *Ipomoea* genus *[Ipomoea trifida*, *I. triloba*, and *I. batatas* (L.) Lam. (sweetpotato)] were identified to investigate their physicochemical and biochemical characteristics. The molecular weight, isoelectric point, instability index, and grand average of hyropathy markedly differed among the three species. The aliphatic index values of sweetpotato AGPase proteins were higher in the small subunit than in the large subunit. The AGPase proteins from sweetpotato were found to contain an LbH_G1P_AT_C domain in the C-terminal region and various domains (NTP_transferase, ADP_Glucose_PP, or Glyco_tranf_GTA) in the N-terminal region. Conversely, most of its two relatives (*I. trifida* and *I. triloba*) were found to only contain the NTP_transferase domain in the N-terminal region. These findings suggested that these conserved domains were species-specific and related to the subunit types of AGPase proteins. The study may enable research on the AGPase-related specific characteristics of sweetpotatoes that do not exist in the other two species, such as starch metabolism and tuberization mechanism.

Keywords ADP-glucose pyrophosphorylase, conserved domain, AGPase small subunit, AGPase large subunit, tuberization, sweetpotato

H. Nie · S.-H. Kim (✉)
Department of Environmental Horticulture, University of Seoul, Seoul 02504, Korea
e-mail: pgel2006@gmail.com

H. Nie · S.-Y. Kwon
Plant Systems Engineering Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Republic of Korea

S. Kim
Bioenergy Crop Research Institute, National Institute of Crop Science, Rural Development Administration, Muan 58545, Republic of Korea

J. Kim
Department of Biotechnology, College of Biomedical & Health Sciences, Global Campus. Konkuk University, ChoongJu, 27478, Korea

Introduction

ADP-glucose pyrophosphorylase (AGPase; EC: 2.7.7.27) is a regulatory enzyme that catalyzes the biosynthesis of alpha 1,4-glucans (glycogen or starch) in photosynthetic bacteria and plants (Smith-White and Preiss 1992). In higher plants, it is a heterotetramer composed of two different but closely related subunits (α2β2): “small” (α subunit, 50-54 kDa) and “large” subunits (β subunit, 51-60 kDa) based on the size difference (Ballicora et al. 2004; Smith-White and Preiss 1992). The small subunit is responsible for the catalytic activity, whereas the large subunit plays regulatory roles (Ballicora et al. 2004; Crevillén et al. 2003). These subunits are necessary for the optimal activity of the native enzyme in plants; a lack of one of the subunits will reduce the activity of the AGPase and influence the synthesis of starch (Li and Preiss 1992). In sweetpotato, AGPase is a key enzyme controlling starch synthesis and is considered an important determinant of the sink activity of the roots (Tsubone et al. 2000; Yatomi et al. 1996). Many AGPase genes have been cloned and studied in sweetpotatoes (Lee et al. 2000; Seo et al. 2015; Zhou et al. 2016).

The protein domains can be considered distinct functions and structural units of proteins that are usually identified as repeating (sequence or structural) units (Ingolfsson and Yona 2008; Li et al. 2012). In molecular evolution, these

domains may have been reorganized in different arrangements in protein function annotation (Ingolfsson and Yona 2008), protein structure determination (Marchler-Bauer et al. 2012), and protein engineering (Guerois and Serrano 2001). Conserved domains are defined by a conserved domain database (CDD) as repeating units in molecular evolution, the extent of which can be determined by sequence and structural analysis (Marchler-Bauer et al. 2012).

Sweetpotato (*Ipomoea batatas* (L.) Lam.) is a hexaploid (2n = 6x = 90) perennial tuberization crop belonging to the family Convolvulaceae (Welbaum 2015). Two non-tuberization diploid *Ipomoea* species, *I. trifida* (H.B.K.) G. Don (2n = 2x = 30) and *I. triloba* L. (2n = 2x = 30), have been reported to be the putative progenitors of sweetpotato, which are commonly considered to be model species for sweetpotato research (Roullier et al. 2013; Wu et al. 2018). In this study, we aimed to screen the AGPase genes from sweetpotato and its two related species to investigate the conserved domains of the coding proteins. The differences in these domains can be used to confirm the molecular functions of the AGPase proteins in sweetpotato and its two relatives.

Methods

Identification of AGPase amino acid sequences

Sweetpotato Genomics Resource (http://sweetpotato.plantbiology.msu.edu/index.shtml) and NCBI databases (https://www.ncbi.nlm.nih.gov/) were used to identify the AGPase domain-containing proteins in the three species. The amino acid sequence of the AGPase protein *IbAGPa1* (BAF47744.2) was used as the driver sequence for BLAST-search.

The ProtParam (http://www.expasy.org/tools/protparam.html) of ExPASy (Expert protein analysis system, https://www.expasy.org/) tool was used to compute the physicochemical characteristics of AGPase proteins in the three species, including the number of amino acids, molecular weight, theoretical isoelectric point (pI), instability (II) and aliphatic index (AI), and grand average of hydropathy (GRAVY) (Gasteiger et al. 2005).

Multiple-sequence alignment and phylogenetic tree structure

The amino acid sequences of the AGPase proteins in FASTA formats were used for multiple-sequence alignment using the CLC Sequence Viewer 7.6 software (CLC bio, Aarhus, Denmark). A neighbor-joining phylogenetic tree was constructed using MEGA X 10.1 software (Pennsylvania State University, US) with the following parameters: bootstrap analysis of 1,000 replicates, Poisson correction method, and pairwise deletion (Kumar et al. 2018).

Conserved domain analysis

Pfam (http://pfam.janelia.org/), SMART (http://smart.embl-heidelberg.de/), and CDD (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) were used to explore the conserved domains of the AGPase proteins. The selected conserved domains were drawn using DOG 2.0.1 software (Ren et al. 2009).

Results

Identification of AGPase proteins

Forty-five AGPase domain-containing proteins from *I. batatas* (26 accessions), *I. trifida* (10 accessions), and *I. triloba* (9 accessions) were identified and used for various analyses (Table 1). The sizes of these proteins were distinctly different; the amino acids ranged from 165 to 525 and the molecular weights (MW) ranged from 18.35 to 58.19 kDa.

The isoelectric point (pI), which represents the average pH of the molecule without a net electrical charge or electrically neutrality, was 4.71-9.53 in all categories. The average pI of *I. batatas*, *I. trifida*, and *I. triloba* AGPase were 6.83, 7.11, and 6.47, respectively. The instability index (II), which represents the stability and instability of a polypeptide at $\leq$ 40 and > 40, respectively, indicated 40 or less in AGPase of *I. batatas*. In contrast, some AGPases of the *I. trifida* and *I. triloba* were 40 or more. The aliphatic index (AI), which represents the relative volume of the aliphatic side chains of a polypeptide, was similar in the three species, but there were differences between subunits of *I. batatas* AGPase. Higher AI values were observed for the small subunits than the large subunits of the *I. batatas* AGPase. The grand average of hydropathy (GRAVY), which was analyzed to determine the hydropathy of AGPase, showed that *I. batatas* had different characteristics from the other two species. All *I. batatas* AGPases showed negative values, whereas some of the *I. trifida* and *I. triloba* AGPases had positive values.

Table 1 Biochemical and physicochemical characteristics of AGPase proteins in the three species

Species	Accession No.	Subunit	Amino acids	Molecular weight (MW)	Isoelectric point (pI)	Instability index (II)	Aliphatic index (AI)	Grand average of hydropathy (GRAVY)
I. batatas	BAF47744.2	Small	522	57155.24	6.74	39.79	91.24	-0.178
I. batatas	AFL55400.1	Small	522	57143.19	6.74	39.50	90.48	-0.188
I. batatas	AAS66988.1	Small	522	57188.32	6.74	39.42	91.23	-0.166
I. batatas	AAA19648.1	Small	303	33530.51	5.52	35.06	96.30	-0.129
I. batatas	CAA86726.1	Small	302	33374.32	5.39	35.14	96.62	-0.115
I. batatas	CAA58473.1	Small	427	47300.22	6.13	36.29	97.12	-0.119
I. batatas	AFL55401.1	Small	523	57164.19	8.02	37.38	90.15	-0.194
I. batatas	BAF47745.1	Small	523	57178.21	8.02	37.38	90.34	-0.190
I. batatas	AAS66987.1	Small	523	57179.24	8.02	36.64	90.52	-0.183
I. batatas	AFL55399.1	Large	525	58055.43	8.92	34.29	88.44	-0.164
I. batatas	AGB85112.1	Large	525	57990.31	8.82	33.14	87.80	-0.158
I. batatas	BAF47749.1	Large	525	58117.46	8.93	35.26	87.50	-0.164
I. batatas	AFL55398.1	Large	518	57269.40	6.37	29.97	85.08	-0.178
I. batatas	BAF47748.1	Large	518	57269.36	6.25	29.73	85.08	-0.177
I. batatas	AGB85111.1	Large	517	57376.52	6.41	28.99	84.29	-0.190
I. batatas	AFL55396.1	Unknown	517	57577.74	7.01	35.32	86.36	-0.245
I. batatas	BAF47746.1	Large	517	57616.78	6.69	36.61	87.31	-0.234
I. batatas	CAB52196.1	Unknown	450	50090.21	5.38	35.94	89.04	-0.168
I. batatas	BAF47747.1	Large	515	57562.13	7.08	31.74	88.99	-0.204
I. batatas	AFL55397.1	Large	515	57485.94	6.44	32.78	88.80	-0.194
I. batatas	AGB85109.1	Large	517	57527.64	6.44	37.97	87.50	-0.237
I. batatas	CAB55495.1	Unknown	490	54707.53	7.14	36.97	89.33	-0.227
I. batatas	AGB85110.1	Large	515	57559.03	6.31	31.13	89.55	-0.212
I. batatas	AAC21562.1	Large	517	57686.94	7.55	38.55	86.92	-0.234
I. batatas	CAB55496.1	Large	385	43443.49	5.35	32.30	85.82	-0.224
I. batatas	CAB51610.1	Large	306	34636.48	5.13	37.96	86.63	-0.300
I. trifida	itf11g03360.t1	Unknown	522	57155.24	6.74	39.79	91.23	-0.178
I. trifida	itf13g19620.t1	Large	525	58186.57	9.01	34.65	87.89	-0.170
I. trifida	itf02g13930.t1	Unknown	523	57178.21	8.02	37.40	90.15	-0.194
I. trifida	itf01g13780.t1	Unknown	351	39640.79	9.53	65.48	93.02	-0.191
I. trifida	itf00g32520.t1	Unknown	351	39204.50	5.40	46.38	99.46	0.111
I. trifida	itf09g27040.t1	Small	474	52547.38	6.15	47.76	85.99	-0.240
I. trifida	itf06g21950.t1	Large	517	57244.40	6.37	28.90	84.87	-0.174
I. trifida	itf08g03850.t1	Large	517	57594.29	8.50	28.36	85.98	-0.201
I. trifida	itf05g24300.t1	Unknown	416	46019.99	5.76	33.92	99.81	0.057
I. trifida	itf10g06320.t1	Unknown	427	48406.64	5.64	37.09	99.53	0.111
I. triloba	itb02g09380.t1	Unknown	523	57164.19	8.02	37.38	90.15	-0.194
I. triloba	itb11g03360.t1	Unknown	522	57155.24	6.74	39.79	91.23	-0.178
I. triloba	itb13g23180.t1	Large	266	29618.76	5.68	32.92	92.74	-0.106
I. triloba	itb09g31010.t1	Small	475	52687.57	6.16	48.56	86.63	-0.236
I. triloba	itb06g20570.t1	Large	517	57203.30	6.51	29.78	83.73	-0.185
I. triloba	itb08g03970.t1	Large	517	57626.35	8.50	28.36	85.42	-0.206
I. triloba	itb09g17690.t1	Unknown	165	18349.10	4.71	32.45	92.24	0.049
I. triloba	itb05g25020.t1	Unknown	416	46032.99	5.76	33.46	99.57	0.050
I. triloba	itb11g22920.t4	Unknown	415	45485.48	6.23	41.54	100.48	0.045

Conserved domain analysis

Six types of conserved domains that showed different distributions were included in the AGPase proteins of these three species (Fig. 1b, Table 2). Most of the *I. trifida* and *I. triloba* AGPases had only the NTP_transferase domain and some had two conserved domains: NTP_transferase at the N-terminal and Hexapep or Cpn60_TCP1 at the C-terminal. On the other hand, the *I. batatas* AGPase proteins had four types of conserved domains (NTP_transferase, LbH_G1P_AT_C, ADP_Glucose_PP, and Glyco_tranf_GTA_type); each of them had two conserved domains. All of the *I. batatas* AGPase proteins had the LbH_G1P_AT_C domain at the C-terminals, but the N-terminals differed according to the subunit. The N-terminal of all large subunits of *I. batatas* AGPase proteins has the NTP_transferase domain only except for CAB51610.1, whereas all small subunits have ADP_Glucose_PP domain except for CAB55496.1, AAA19648.1, and CAA86726.1. The proteins with this exception all had partial sequences and had the Glyco_tranf_GTA_type domain at the C-terminals.

Phylogenetic analysis

The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). Fig. 1a presents the optimal tree with the sum of the branch length = 29.09. This analysis involved 45 amino acid sequences and 512 positions. The conserved domains were labeled on the amino acid sequences (Fig. 1a). The length and type of the domain were different for each species. Based on the phylogenetic tree, AGPase proteins from these species were grouped together according to large and small subunit type.

Discussion

AGPase is an important factor involved in the tuberous root of sweetpotatoes because it is a vital enzyme in starch synthesis (Tsubone et al. 2000; Yatomi et al. 1996). Although it is also present in *I. trifida* and *I. triloba*, as well as in plants of the genus *Ipomoea*, they all have different physiological properties from sweetpotatoes, such as non-tuberization. Therefore, AGPase is believed to have different structures or different functions in plants of the genus *Ipomoea*. The AGPase identification of sweetpotatoes and two non-tuberous *Ipomoea* species performed in this study is very important for understanding the relationship between plants of the genus *Ipomoea* and the functions of each species. Sweetpotato is a polyploid crop of *I. trifida*, but it is unclear if it is autopolyploidy or allopolyploidy (Roullier et al. 2013; Wu et al. 2018). The amount of AGPases increased by whole-genome duplication in sweetpotatoes from its relatives. This result is consistent with a study showing that the number of *rboh* genes in the polyploid plant, *Gossypium hirsutum*, was higher than its progenitor plants *G. raimonddi* and *G. arboreum* (Wang et al. 2020). Moreover, some AGPases in *I. trifida* and *I. triloba* exhibited an II value $\geq$ 40, which means an unstable state, but there was no AGPase representing an II value $\geq$ 40 in *I. batatas* (Table 1). This suggests that some of the genes that were unstable during the evolution of *I. batatas* may have been deleted.

A difference in the domain composition of AGPase was observed between sweetpotatoes and the other *Ipomoea* plants; *I. batatas* has a more complex composition (Fig. 1b). The N-terminal of the small subunit and the C-terminal in sweetpotatoes were composed differently from the domains of the two species. These results suggest that LbH_G1P_AT_C at the C-terminal and ADP_Glucose_PP and Glyco_tranf_GTA_type at the N-terminal of the small subunit contribute to the different functions and regulations than non-tuberous relative plants. Many studies have shown that genes can be orthologs or paralogs by domain architectures, such as the insertion and deletion of new domains during evolution (Björklund et al. 2005; Forslund et al. 2011). Although this study cannot confirm the homolog genes of each AGPase in the genus *Ipomoea* plants, the evolutionary process of the genome among these plants, including AGPase, is expected to be revealed through further studies.

Conclusion

Sweetpotato AGPases have relatively conserved domains compared to *I. trifida* and *I. triloba*. The small subunit of AGPase showed complex structures in sweetpotatoes compared to the other two species. Sweetpotato AGPase had the LbH_G1P_AT_C domain in the C-terminal region, which was not present in *I. trifida* and *I. triloba*. This suggests that the structure of AGPase in sweetpotato, which is different from the other two species, plays important roles in certain functions of sweetpotatoes, such as starch biosynthesis and tuber formation. More isolation studies and further examination of gene expression will be needed to clarify the functional role of sweetpotato-specific domains in tuberization.

Table 2 Conserved domain prediction of the AGPase in the three species

Species	Accession No.	Amino acid	Conserved domain 1				Conserved domain 2			
			ID	Name	Start	End	ID	Name	Start	End
I.batatas	BAF47744.2	522	cd02508	ADP_Glucose_PP	103	352	cd04651	LbH_G1P_AT_C	390	516
I.batatas	AFL55400.1	522	cd02508	ADP_Glucose_PP	103	352	cd04651	LbH_G1P_AT_C	390	516
I.batatas	AAS66988.1	522	cd02508	ADP_Glucose_PP	103	352	cd04651	LbH_G1P_AT_C	390	516
I.batatas	AAA19648.1	303	cd00761	Glyco_tranf_GTA_type	1	147	cd04651	LbH_G1P_AT_C	171	297
I.batatas	CAA86726.1	302	cd00761	Glyco_tranf_GTA_type	1	146	cd04651	LbH_G1P_AT_C	170	296
I.batatas	CAA58473.1	427	cd02508	ADP_Glucose_PP	1	257	cd04651	LbH_G1P_AT_C	295	421
I.batatas	AFL55401.1	523	cd02508	ADP_Glucose_PP	104	353	cd04651	LbH_G1P_AT_C	391	517
I.batatas	BAF47745.1	523	cd02508	ADP_Glucose_PP	104	353	cd04651	LbH_G1P_AT_C	391	517
I.batatas	AAS66987.1	523	cd02508	ADP_Glucose_PP	104	353	cd04651	LbH_G1P_AT_C	391	517
I.batatas	AFL55399.1	525	cd04181	NTP_transferase	93	307	cd04651	LbH_G1P_AT_C	393	519
I.batatas	AGB85112.1	525	cd04181	NTP_transferase	93	307	cd04651	LbH_G1P_AT_C	393	519
I.batatas	BAF47749.1	525	cd04181	NTP_transferase	93	307	cd04651	LbH_G1P_AT_C	393	519
I.batatas	AFL55398.1	518	cd04181	NTP_transferase	88	363	cd04651	LbH_G1P_AT_C	386	512
I.batatas	BAF47748.1	518	cd04181	NTP_transferase	88	363	cd04651	LbH_G1P_AT_C	386	512
I.batatas	AGB85111.1	517	cd04181	NTP_transferase	87	362	cd04651	LbH_G1P_AT_C	385	511
I.batatas	AFL55396.1	517	cd04181	NTP_transferase	87	362	cd04651	LbH_G1P_AT_C	385	511
I.batatas	BAF47746.1	517	cd04181	NTP_transferase	87	362	cd04651	LbH_G1P_AT_C	385	511
I.batatas	CAB52196.1	450	cd04181	NTP_transferase	20	295	cd04651	LbH_G1P_AT_C	318	444
I.batatas	BAF47747.1	515	cd04181	NTP_transferase	85	360	cd04651	LbH_G1P_AT_C	383	509
I.batatas	AFL55397.1	515	cd04181	NTP_transferase	85	360	cd04651	LbH_G1P_AT_C	383	509
I.batatas	AGB85109.1	517	cd04181	NTP_transferase	87	362	cd04651	LbH_G1P_AT_C	385	511
I.batatas	CAB55495.1	490	cd04181	NTP_transferase	60	335	cd04651	LbH_G1P_AT_C	358	484
I.batatas	AGB85110.1	515	cd04181	NTP_transferase	85	360	cd04651	LbH_G1P_AT_C	383	509
I.batatas	AAC21562.1	517	cd04181	NTP_transferase	87	362	cd04651	LbH_G1P_AT_C	385	511
I.batatas	CAB55496.1	385	cd00761	Glyco_tranf_GTA_type	2	230	cd04651	LbH_G1P_AT_C	253	379
I.batatas	CAB51610.1	306	cd00761	Glyco_tranf_GTA_type	1	151	cd04651	LbH_G1P_AT_C	174	300
I.trifida	itf11g03360.t1	522	cd04181	NTP_transferase	94	367				
I.trifida	itf13g19620.t1	525	cd04181	NTP_transferase	94	371				
I.trifida	itf02g13930.t1	523	cd04181	NTP_transferase	95	368				
I.trifida	itf01g13780.t1	351	cd04181	NTP_transferase	243	299				
I.trifida	itf00g32520.t1	351	cd04181	NTP_transferase	127	182				
I.trifida	itf09g27040.t1	474	cd04181	NTP_transferase	56	322				
I.trifida	itf06g21950.t1	517	cd04181	NTP_transferase	86	363				
I.trifida	itf08g03850.t1	517	cd04181	NTP_transferase	86	363				
I.trifida	itf05g24300.t1	416	cd04181	NTP_transferase	11	205	pfam00132	Hexapep	297	329
I.trifida	itf10g06320.t1	427	cd04181	NTP_transferase	109	161	pfam00118	Cpn60_TCP1	175	212
I.triloba	itb02g09380.t1	523	cd04181	NTP_transferase	95	368				
I.triloba	itb11g03360.t1	522	cd04181	NTP_transferase	94	367				
I.triloba	itb13g23180.t1	266	cd04181	NTP_transferase	1	112				
I.triloba	itb09g31010.t1	475	cd04181	NTP_transferase	57	323				
I.triloba	itb06g20570.t1	517	cd04181	NTP_transferase	86	363				
I.triloba	itb08g03970.t1	517	cd04181	NTP_transferase	86	363				
I.triloba	itb09g17690.t1	165	cd04181	NTP_transferase	2	30	cd04181	NTP_transferase	38	85
I.triloba	itb05g25020.t1	416	cd04181	NTP_transferase	11	205	pfam00132	Hexapep	297	329
I.triloba	itb11g22920.t4	415	cd04181	NTP_transferase	10	211	pfam00132	Hexapep	300	328

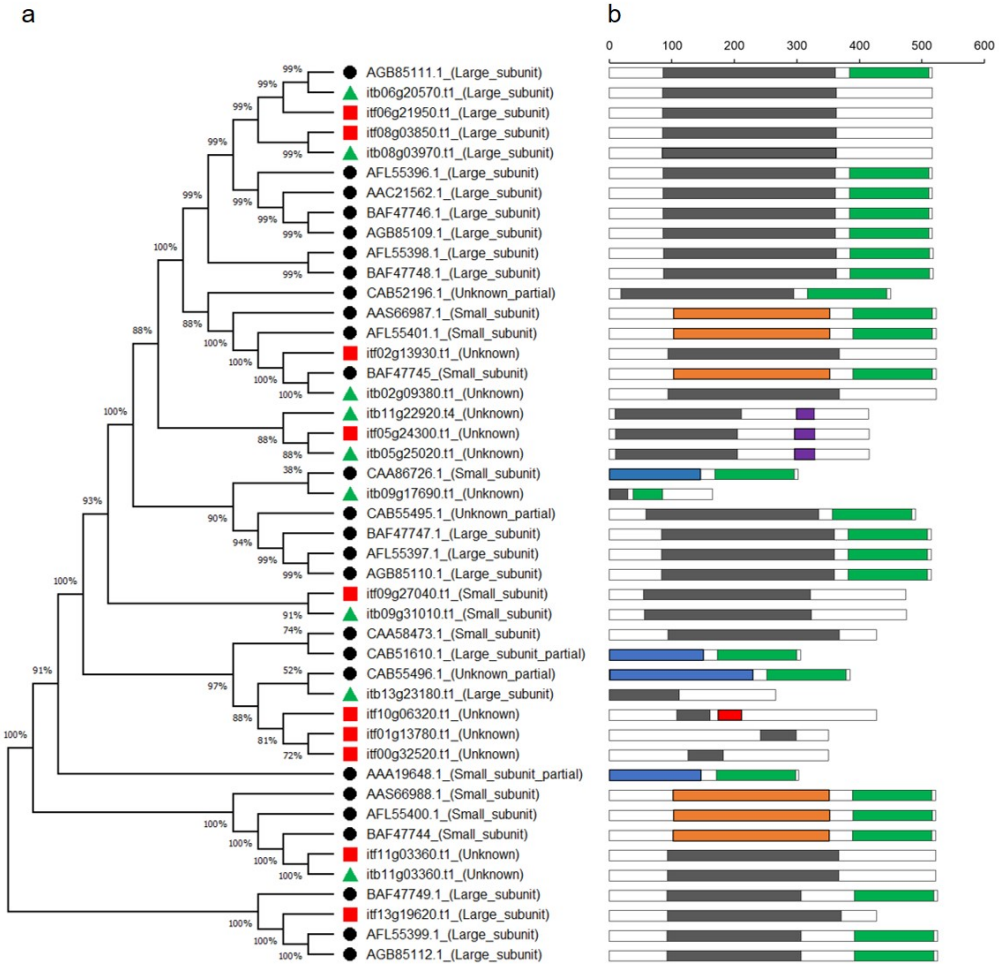

Fig. 1 Phylogenetic tree (a) and domain structure (b) of the AGPase proteins in *Ipomoea batatas* (black circles), *I. trifida* (red quadrangles), and *I. triloba* (green triangles). The numbers at the nodes indicate the bootstrap values. The conserved domains are indicated by colored blocks on the right. Gray, NTP_transferase; green, LbH_G1P_AT_C; blue, Glyco_tranf_GTA_type; purple, Hexapep; red, Cpn60_TCP1; orange, ADP_Glucose_PP

References

Ballicora MA, Iglesias AA, Preiss J (2004) ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. Microbiol Mol Biol Rev 67(2):213-225. https://doi.org/10.1128/mmbr.67.2.213-225.2003

Björklund ÅK, Ekman D, Light S, Frey-Skött J, Elofsson A (2005) Domain Rearrangements in Protein Evolution. J Mol Biol 353(4):911-923. https://doi.org/10.1016/j.jmb.2005.08.067

Crevillén P, Ballicora MA, Mérida A, Preiss J, Romero JM (2003) The different large subunit isoforms of *Arabidopsis thaliana* ADP-glucose pyrophosphorylase confer distinct kinetic and regulatory properties to the heterotetrameric enzyme. J Biol Chem 278(31):28508-28515. https://doi.org/10.1074/jbc.M304280200

Forslund K, Pekkari I, Sonnhammer EL (2011) Domain architecture conservation in orthologs. BMC Bioinformatics 12:326. https://doi.org/10.1186/1471-2105-12-326

Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy Server. In: Walker JM (eds) The proteomics protocols handbook. Humana Press, New Jersey, pp 571-607. https://doi.org/10.1385/1-59259-890-0:571

Guerois R, Serrano L (2001) Protein design based on folding models. Curr Opin Struct Biol 11:101-106. https://doi.org/10.1016/S0959-440X(00)00170-6

Ingolfsson H, Yona G (2008) Protein domain prediction. Methods Mol Biol 426:117-143. https://doi.org/10.1007/978-1-60327-058-8_7

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. Mol Biol Evol 35:1547-1549. https://doi.org/10.1093/molbev/msy096

Lee SS, Bae JM, Oh MS, Liu, JR, Harn CH (2000) Isolation and characterization of polymorphic cDNAs partially encoding ADP-glucose pyrophosphorylase (AGPase) large subunit from sweet potato. Mol Cells 10(1):108-112. https://doi.org/10.1007/s10059-000-0108-3

Li BQ, Hu LL, Chen L, Feng KY, Cai YD, Chou KC (2012) Prediction of protein domain with mRMR feature selection and

analysis. PLOS ONE 7(6):e39308. https://doi.org/10.1371/journal.pone.0039308

Li L, Preiss J (1992) Characterization of ADP-glucose pyrophosphorylase from a starch-deficient mutant of *Arabidopsis thaliana* (L). Carbohydr Res 227:227-239. https://doi.org/10.1016/0008-6215(92)85074-A

Marchler-Bauer A, Zheng C, Chitsaz F, Derbyshire MK, Geer LY, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Lu S, Marchler GH, Song JS, Thanki N, Yamashita RA, Zhang D, Bryant SH (2012) CDD: conserved domains and protein three-dimensional structure. Nucleic Acids Res 41(D1):D348-D352. https://doi.org/10.1093/nar/gks1243

Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X (2009) DOG 1.0: illustrator of protein domain structures. Cell Res 19(2): 271-273. https://doi.org/10.1038/cr.2009.6

Roullier C, Duputié A, Wennekes P, Benoit L, Bringas VMF, Rossel G, Tay D, McKey D, Lebot V (2013) Disentangling the origins of cultivated sweetpotato (*Ipomoea batatas* (L.) Lam.). PLOS ONE 8(5):e62707. https://doi.org/10.1371/journal.pone.0062707

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4(4):406-425. https://doi.org/10.1093/oxfordjournals.molbev.a040454

Seo SG, Bea SH, Jun BK, Kim ST, Kwon SY, Kim SH (2015) Overexpression of ADP-glucose pyrophosphorylase (*IbAGPaseS*) affects expression of carbohydrate regulated genes in sweetpotato [*Ipomoea batatas* (L.) Lam. cv. Yulmi]. Genes Genom 37:595-605. https://doi.org/10.1007/s13258-015-0289-y

Smith-White BJ, Preiss J (1992) Comparison of proteins of ADP-glucose pyrophosphorylase from diverse sources. J Mol Evol 34(5):449-464. https://doi.org/10.1007/BF00162999

Tsubone M, Kubota F, Saitou K, Kadowaki M (2000) Enhancement of tuberous root production and Adenosine 5'-Diphosphate Pyrophosphorylase (AGPase) activity in sweetpotato (*Ipomoea batatas* Lam.) by exogenous injection of sucrose solution. J Agro Crop Sci 184(3):181-186. https://doi.org/10.1046/j.1439-037x.2000.00396.x

Wang W, Chen D, Liu D, Cheng Y, Zhang X, Song L, Hu M, Dong J, Shen F (2020) Comprehensive analysis of the *Gossypium hirsutum* L. respiratory burst oxidase homolog (*Ghrboh*) gene family. BMC Genomics 21:91. https://doi.org/10.1186/s12864-020-6503-6

Welbaum GE (2015) Vegetable production and practices. CAB International, Oxfordshire

Wu S, Lau KH, Cao Q et al (2018) Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. Nat Commun 9:4580. https://doi.org/10.1038/s41467-018-06983-8

Yatomi M, Kubota F, Saito K, Agata W (1996) Evaluation of root sink ability of sweetpotato (*Ipomoea batatas* Lam.) cultivars on the basis of enzymatic activity in the starch synthesis pathway. J Agro Crop Sci 177(1):17-23. https://doi.org/10.1111/j.1439-037X.1996.tb00587.x

Zhou Y, Chen Y, Tao X, Cheng X, Wang H (2016) Isolation and characterization of cDNAs and genomic DNAs encoding ADP-glucose pyrophosphorylase large and small subunits from sweetpotato. Mol Genet Genomics 291(2):609-620. https://doi.org/10.1007/s00438-015-1134-3