Phytochemistry 156 (2018) 33–42

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Three novel mutations in α-galactosidase gene involving in galactomannan degradation in endosperm of curd coconut

Chanrith Phoeurk[a], Jamorn Somana[b,*], Thakorn Sornwatana[b], Somsiri Udompaisarn[b], Somchoke Traewachiwiphak[b], Phanthila Sirichaiyakul[c], Thanawat Phongsak[d], Dumrongkiet Arthan[c]

[a] Department of Bio-Engineering, Faculty of Engineering, Royal University of Phnom Penh, Phnom Penh, 12101, Cambodia
[b] Department of Biochemistry, Faculty of Science, Mahidol University, Bangkok, 10400, Thailand
[c] Department of Tropical Nutrition and Food Science, Faculty of Tropical Medicine, Mahidol University, Bangkok, 10400, Thailand
[d] Program of Chemistry, Faculty of Science and Technology, Thepsatri Rajabhat University, Lopburi, 15000, Thailand

ARTICLE INFO

Keywords:
Curd coconut
Endosperm mutation
Galactomannan
α-galactosidase

ABSTRACT

The deficiency of α-galactosidase activity in coconut endosperm has been reported to cause a disability to hydrolyze oligogalactomannan in endosperm resulting in curd coconut phenotype. However, neither the α-galactosidase encoding gene in coconut nor the mutation type has been identified and characterized in normal and curd coconuts. In this study, cDNA and genomic DNA encoding α-galactosidase gene alleles from a normal and two curd coconuts were successfully cloned and characterized. The deduced amino acid of wild type α-galactosidase contains 398 amino acid residues with a 17 N-terminal amino acids signal peptide sequence. Three mutant alleles, the first 19-amino acids from 67 to 85 (ADALVSTGLARLGYQYVNL) deletion with S137R and the second R216T, were identified from curd coconut plant no.1 while the third P250R was identified from curd coconut plant no. 10. All mutations of α-galactosidase gene were confirmed by the analysis of parental genomic DNA from normal and curd coconuts. Heterologous expression in *Komagataella phaffii* (*Pichia pastoris*) indicated that recombinant P250R, R216T and 19-amino acids deletion-S137R mutant proteins showed no α-galactosidase activity. Only the recombinant wild-type protein was able to detect for α-galactosidase activity. These results are in accordance with the no detection of α-galactosidase activity in developing curd coconut endosperms by tissue staining. While, the accumulation of enzyme activity was present in the solid endosperm of normal coconut. The full-length cDNA and parental genomic DNA sequences encoding α-galactosidase in normal coconut as well as identified curd coconut mutant alleles are reported in Genbank accession no. KJ957156 and KM001681-3. Transcription level of the α-galactosidase gene in mature curd coconut endosperm was at least 20 times higher than normal. In conclusion, absence of α-galactosidase activity caused by gene mutations associates with an accumulation of oligogalactomannan in endosperms, resulting in curd coconut phenotype.

1. Introduction

Coconut (*Cocos nucifera* L.) is a monophyletic genus within the family Arecaceae (Palmae), a member of the subfamily Arecoideae that includes 27 genera and about 600 species of monocotyledonous flowering plants (Perera et al., 1998 & 2003). Coconuts have a diploid genome with 2n = 32 chromosomes and a haplophasic genome size (1C) of around 2.15 billion base pairs (Rohde et al., 2002). The coconut palm is called “tree of life” because it provides almost all the necessities of life including food, drink, oil, medicine, fiber, timber, thatch, and other domestic utensils (Foale, 2003). Curd coconut is an abnormal endosperm phenotype, often found in tall type coconuts (Santoso et al., 1996 and Islam et al., 2009). In normal plantation, only few curd coconut fruits occur only from some coconut plants. Unlike normal coconut, the solid endosperm of curd coconut becomes soft, while its liquid endosperm becomes thick or filling the entire cavity as a soft gel (Islam et al., 2009), as shown in Fig. 1. The curd coconut is in high demand due to its soft fatty and delicious texture and the price of a fruit is about 10 times of a normal coconut (Wattanayothin, 2004, 2005). In addition, the phenotype of curd coconut was found to be a pure genetic factor of a single gene locus, which is inherited in autosomal recessive fashion (Zuniga, 1953).

* Corresponding author.
E-mail address: jamorn.som@mahidol.ac.th (J. Somana).

https://doi.org/10.1016/j.phytochem.2018.08.015
Received 2 September 2017; Received in revised form 25 July 2018; Accepted 17 August 2018
Available online 30 August 2018
0031-9422/

Abbreviations

AGal	α-galactosidase
RACE	Rapid Amplification of cDNA Ends
AOX	Alcohol Oxidase
CnAGal	Coconut α-galactosidase

Galactomannans make up to 61% of the total carbohydrate content of a mature normal coconut kernel (galactose to mannose ratio 1:12); and they play a structural role in the formation of the secondary cell wall in coconut solid endosperm (Balasubramaniam, 1976). Whereas in curd coconut kernel, cell wall was principally 60% water soluble galactomannan with galactose to mannose ratio of 1:2 (Flavier, 1999). In normal coconut or other seed plants, cell wall storage polysaccharides are degraded and mobilized as sucrose, glucose or fructose to supply nutrient during seed germination (Buckeridge, 2010). While the curd coconut contains an apparently normal embryo, it fails to germinate *in vivo* because its endosperm is incapable of supporting embryo germination. However, germination of curd coconuts is successful *in vitro* using tissue culture techniques (De Guzman and Del Rosario, 1964; Ashburner et al., 1995).

Galactomannans, a component of the cell wall, consist of a linear backbone of β-(1 → 4)-linked D-mannose residues to which D-galactose residues are attached by α-(1 → 6) linkages. The degradation of galactomannan requires the presence of three enzymes working together: namely endo-β-mannanase (EC 3.2.1.78), β-mannosidase (EC 3.2.1.25) and α-galactosidase (EC 3.2.1.22) (Reid and Meier, 1973). Each of these enzymes have specific mechanisms for cleaving sugar residues at different sites: endo-β-mannanase randomly cleaves endo (1 → 4)-β-D-mannosidic bonds while β-mannosidase removes terminal (1 → 4)-β-D mannosides from the ends of the oligogalactomannan, and α-galactosidase cleaves the (1 → 6)-α-D galactoside residues from the non-reducing end of oligogalactomannan. Investigations into the comparative activities of these three enzymes within the endosperms of both normal and curd coconuts reveal that in normal coconut, only α-galactosidase activity increased continually with endosperm maturation, and was inversely correlated with the amount of galactomannan. In contrast, only α-galactosidase activity could not be detected in any stage of curd coconut endosperm development except the very last stage (11–12 months of age) where its activity was 8300 folds lower than that of the normal coconut (Mujer et al., 1984; Samonte et al., 1989). Thus, these studies confirmed lacking of endosperm specific coconut α-galactosidase (CnAGal) plays a major role in the *in vivo* degradation of oligogalactomannan in the endosperm. In addition, the deficiency of CnAGal activity affects the accumulation of oligogalactomannan, which causes curd coconut endosperm (Mujer et al., 1984). However, neither the cDNA encoding *CnAGal* gene nor mutations in the gene has been identified and characterized. This study therefore aims to isolate and characterize the wild-type α-galactosidase gene from normal coconut, and identify the types of mutation presenting in the α-galactosidase gene, resulting in the curd coconut phenotype.

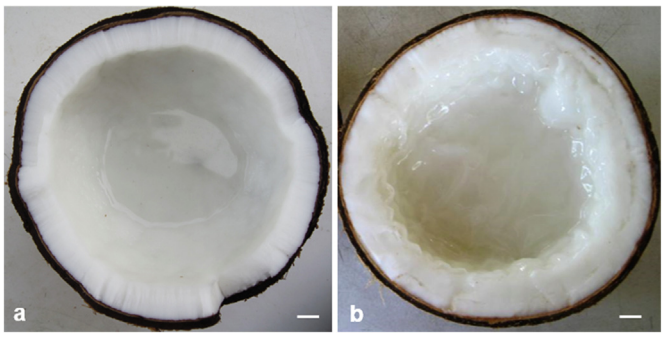


Fig. 1. Comparison of mature coconut endosperm; thin and hard solid endosperm with clear runny water of liquid endosperm in normal coconut (a), thick and fluffy solid endosperm with viscous water of liquid endosperm in curd coconut (b). Scale bars are 1 cm.

2. Results and discussion

2.1. Detection of α-galactosidase in coconut endosperm by tissue staining

Using 5-bromo-4-chloro-3-indolyl-α-D-galactopyranoside (X-α-gal) for tissue staining, AGal activity was determined in the endosperms of both normal and curd coconuts through detection of indigo complex formation. Solid endosperm was not present until fruit age around 6–7 months old and formation of solid endosperm between normal and curd coconut during 6–11 months old were similar. The solid endosperm formed at antipodal end first then gradually expanded to micropylar end (data not shown). Enzyme activity in the solid endosperm of normal coconut was displayed as blue coloration, which follows fruit maturation in starting from the inner layer of the solid endosperm around 8–10 months old then distributing to the outer layer later. In contrast, no blue coloration from enzyme activity was detected in the solid endosperm of curd coconuts (Fig. 2) at all stages. These results suggested the roles of AGal in solid endosperm formation and degradation. It also confirmed curd coconuts lacking of AGal activity. Consequently, galactomannan accumulation occurred in endosperm exhibits curd coconut phenotype. Thus, these results are accordance with lacking of AGal activity in curd coconuts, as previously published (Mujer et al., 1984; Samonte et al., 1989). In addition, X-α-gal tissue staining of the solid endosperm at both micropylar and antipodal ends at different ages of normal coconuts provided the information of activity patterns of AGal during the solid endosperm development. The activity patterns of AGal, which may link to some physiological roles, are required further study to clarify.

2.2. Analysis of cDNA encoding AGals from normal and curd coconuts

An α-galactosidase (EC 3.2.1.22) is one of glycoside hydrolase enzymes, which catalyze the hydrolysis of terminal, non-reducing alpha-D-galactose residues in α-D-galactosides, including galactose oligosaccharides, galactomannans and galactolipids. The full-length cDNAs encoding *CnAGal* genes were identified from both normal and curd coconuts and characterized (Supplementary Fig. S1). The full-length cDNA encoding *CnAGal* gene of normal coconut (KJ957156) consisted of a 1194-nucleotide sequence, encoding 398 amino acids and a calculated molecular weight of 43.575 kD. The encoded amino acid sequence of CnAGal was analyzed by the SignalP 3.0 program (Bendtsen et al., 2004). The predicted signal peptide cleavage site was between positions A17 and G18. The mature protein with 381 amino acids had a calculated molecular weight of 41.703 kD. After analyzed by TargetP 1.1 software, the signal peptide of CnAGal involved in secretory pathway, which may target CnAGal into the extracellular space or to the cell wall of coconut endosperm for the hydrolysis cell wall-associated galactomannan (Bewley, 1997; Minic, 2008). After analysis for the N-glycosylation by NetGLY software, no possible N-linked glycosylation was found. The deduced amino acid sequence of CnAGal showed identity to those of AGals from *Vitis vinifera* (XP_002268711.1), *Nicotiana tabacum* (AEB98600.1), *Phaseolus vulgaris* (AAA73964.1), *Helianthus annuus* (BAC66445.1), *Coffea arabica* (CAI47559.1; Zhu et al., 1995 & 1995), *Arabidopsis lyrata* (XP_002873354.1), *Oryza sativa* (1UAS), and *Zea mays* (NP_001147362.1), 76%,78%, 75%, 75%, 73%, 70%, 73% and 63%, respectively (Fig. 3). Moreover, theoretically estimated molecular weight of the full length CnAGal is similar to those at 43 kDa of AGals previously reported in other plants including *Caffea arabica* L. (coffee bean; Zhu and Goldstein, 1994) *Lycopersicon esculentum* Mill. (tomato; Feurtado et al., 2001) seeds *Oryza sativa* L. var. Nipponbare (rice; Kim et al., 2002), *Cicer arietinum* (chickpea; Singh

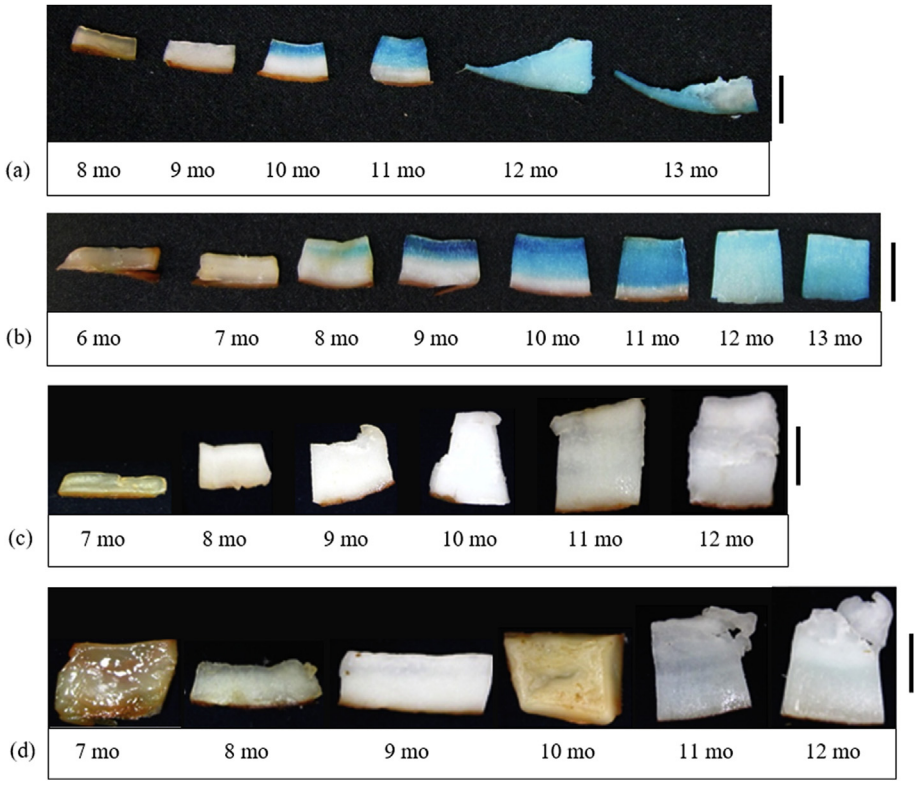


Fig. 2. The α-galactosidase activity of in normal and curd coconut solid endosperms at different stages of endosperm development. (a) Normal endosperm from 8 to 13 months at micropylar end. (b) Normal endosperm from 6 to 13 months at antipodal end. (c) Endosperm of curd coconut from plant no. 1 at micropylar end from 7 to 12 months. (d) Endosperm of curd coconut from plant no. 10 at micropylar end from 7 to 12 months (mo). Scale bars are 1 cm.

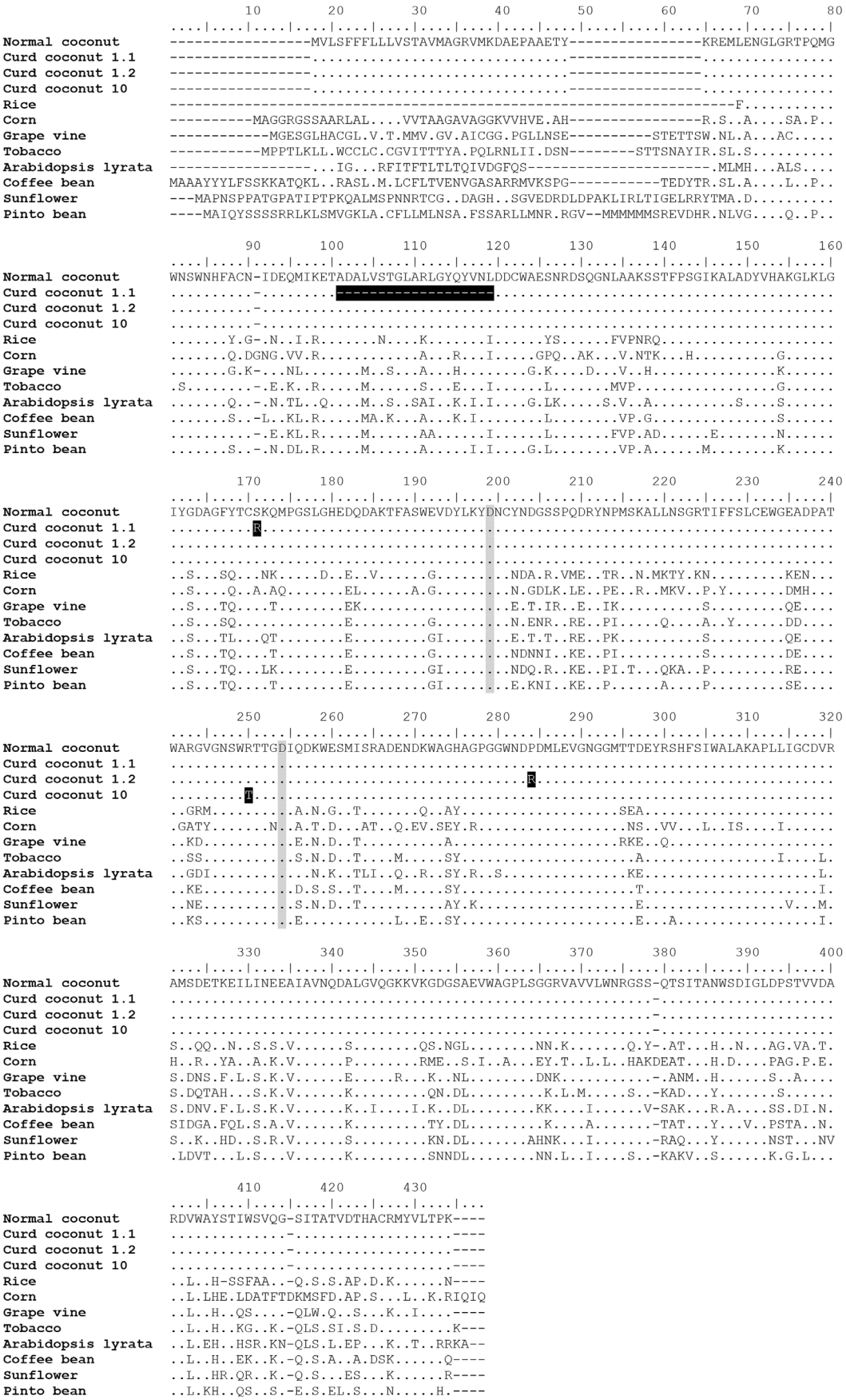
and Kayastha, 2012), *Phaseolus coccineus* (runner bean; Du et al., 2013). CnAGal belongs to glycoside hydrolase family 27 (GH27) based on the classification of glycosidase enzymes using the similarity of amino acid sequences (Henrissat, 1991). In addition, the active site pocket of CnAGal contains the catalytic nucleophile Asp-165 and the catalytic acid/base Asp-220 (Figs. 3 and 7).

2.3. Identification of mutation types in curd coconut genes

All full-length of cDNAs encoding *CnAGal* genes isolated from normal and curd coconuts were sequenced and characterized to determine all mutation types occurred in *CnAGal* genes. The full-length of parental genomic DNA encoding *CnAGal* genes from normal coconut, curd coconut no. 1 two alleles (1.1 and 1.2) and no. 10 comprises of 7629, 7623, 7629 and, 7626 bp, respectively (Supplementary Fig. S1). Additionally, the full-length parental DNA encoding the *CnAGal* gene contains 15 exons separated by 14 introns (Fig. 6 and Supplementary Fig. S1). In curd coconut no.1, the first allele (1.1, Genbank accession no. KM0011681) containing two mutation regions were 57-nucleotides deletion (19-amino acids deletion from 67 to 85 (ADALVSTGLARLGYQYVNL) together with *C4*049A compared with the wild type genomic sequence or S137R compared with the wild type amino acid sequence which designated Δexon3-S137R respectively. In case of Δexon3 mutation, comparing with wild type and its genomic information, the 57-nucleotides deletion at mRNA level caused by three nucleotides (CAG) deletion at base 1956–1958. The deletion located at the end of intron 2 and exon 3 junction, resulting in exon 3 skipping by missplicing of mRNA and exclude 57 nucleotides from exon 3. The other allele of curd coconut plant no. 1 (1.2, KM0011682) was found to have C5817G (P250R). While G5546C (R216T) mutation was found in curd coconut plant no. 10, KM0011683 (Fig. 3 and Supplementary Fig. S1). Comparison of genomic DNA sequences between curd and the normal coconuts, three mutations in the parental gDNA encoding *CnAGal* gene were corresponding to three mutations found in cDNA sequences isolated and characterized from curd coconut from plant no. 1 and 10 (Supplementary Fig. S1). Two different mutant alleles were found in the same AGal gene of curd coconut plant no.1. It resulted from heterozygous allele. However, only a mutant allele of AGal gene found in curd coconut plant no. 10. This either resulted from homozygous allele or unfound for the other mutant allele.

2.4. Expression of wild type and mutant CnAGal genes

To understand the consequences of the curd phenotype caused by mutations at the protein and enzyme levels, the recombinant wild-type and mutant CnAGal proteins (rwtCnAGal and rmtCnAGal) were expressed in *K. phaffii* to detect for their enzyme activity and analyze for protein profiles. A PCR-mediated gene construct engineered to represent wild type cDNA encoding the protein sequence of the wtCnAGal from positions 18 to 398 corresponding to the mature enzyme was successfully cloned into pPICZαB vector, designated as *pPICZαB/wtCnGal*. Similarly, *Δexon3-S*137R, *R*216T and *P*250R mutant cDNAs were also successfully cloned into pPICZαB, designated *pPICZαB/Δexon3-S*137R, *pPICZαB/R*216T and, *pPICZαB/P*250R respectively. All recombinant plasmids were also confirmed for the successful integration of wild-type and mutant *CnAGal* genes into *K. phaffii* by colony PCR techniques (data not shown). Cell-free supernatants (CFS) AGal activity and specific activity from a selected clone from each construct was shown in Table 1. Cell-free supernatants of 10 clones of *K. phaffii* harboring *pPICZαB/wtCnAGal* produced rwtCnAGal containing enzyme specific activity at 0.578 ± 0.105U/mg, using 5 mM pNPAGal as a substrate. Under the same conditions, CFS of 10 clones of *each K. phaffii* harboring *pPICZαB/Δexon3-S*137R, *pPICZαB/R*216T and, *pPICZαB/P*250R produced rΔexon3-S137R, rP250R and rR216T proteins, respectively. All CFS from expressed mutant constructs showed no AGal activity (Table 1, lane 4–9, Fig. 5 and Supplementary Fig. S2). SDS-PAGE analysis of CFS of transformants harboring *pPICZαB* as a control showed no band at both 41.7 and 39.7 kDa (lane 1, Fig. 5). The rwtCnAGal (lane 2–3; Fig. 4) showed protein band around 43 kDa, which is higher than the predicted CnAGal without signal peptide at 41.7 kDa similar to rR216T (lane 6–7; Fig. 4) and rP250R (lane 8–9; Fig. 4). However, r*Δexon3-S*137R (lane 4–5, Fig. 5) showed no protein



(caption on next page)

Fig. 3. Alignment of deduced amino acids among CnAGals of normal and curd coconut and AGals from other plants by ClustalW and Gene doc program. Coconut (*Cocos nucifera*) KJ957156, rice (*Oryza sativa*) AB039671, corn (*Zea mays*) EU954969, grape vine (*Vitis vinifera*) XM_002268675, tobacco (*Nicotiana tabacum*) HQ877670, *Arabidopsis lyrata* XP_020879139, coffee bean (*Coffea arabica*) AJ877911, sunflower (*Helianthus annuus*) AB092594, and pinto bean (*Phaseolus vulgaris*) U12927. Black highlights represent mutant residues. Gray highlights are major catalytic residues. The . are residues that identical to normal coconut and – are missing residues in the alignment.

Table 1
Determination of α-galactosidase activity in cell free supernatant expressed by *K. phaffii* transformed with the wild type and mutant mature *CnAgal* constructs.

CnAgal construct	Protein conc. (mg/ml)	Enzyme activity (U/ml)	Specific activity (U/mg)
wt	25.9 ± 1.7	14.92 ± 2.61	0.578 ± 0.105
Δexon3-S137R	25.0 ± 1.4	0.00 ± 0.00	0.00 ± 0.00
R216T	25.9 ± 2.1	0.00 ± 0.00	0.00 ± 0.00
P250R	27.0 ± 1.6	0.00 ± 0.00	0.00 ± 0.00
pPICZαB	27.4 ± 1.6	0.00 ± 0.00	0.00 ± 0.00

Note: Mean ± SD of protein concentration, enzyme activity, and specific activity of were calculated from 10 independent positive integrant clones from each plasmid construct.

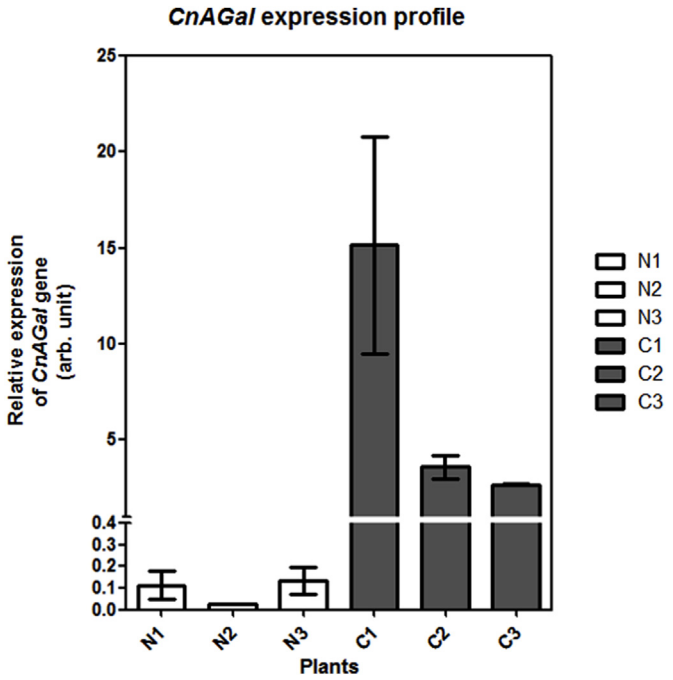


Fig. 4. Relative *CnAGal* expression between wild type and curd coconuts determined by qPCR. Total RNAs isolated from solid endosperm of three independent fruits of normal (N1, N2, and N3) and curd (C1, C2, and C3) coconuts were converted to first strand cDNA, and were analyzed by $2^{-\Delta\Delta Ct}$ method. Error bars represent *SD*, n = 2.

band as predicted at 39.7 kDa or anywhere nearby. In fact, the calculated theoretical molecular weight of full length CnAGal is 43 kDa. Although, no possible N-glycosylation was found, the different the calculated theoretical molecular weight is likely smaller than the estimation by SDS-PAGE analysis. This might result from interaction of protein to gel matrix in each buffer system, O-glycosylation or, any post-translation modifications of CnAgal. The results also showed the effect of mutations Δexon3-S137R, R216T, and P250R on protein stability. The absent band of r*Δexon3-S1*37R might result from protein misfolding and decreased protein stability which prone to be degraded by protease intracellularly (lane 4–5, Fig. 5). Similarly, compared with wild-type protein, rP250R showed faint band at 43 kDa (lane 8–9, Fig. 5). The results also indicated that P250R mutation might affect to

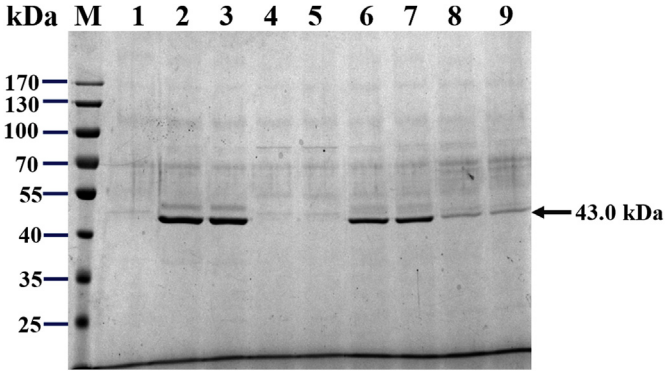


Fig. 5. SDS-PAGE analysis of recombinant wild-type and recombinant CnAGal proteins in cell-free supernatant with normalized total protein to 500 μg in each lane. *K. phaffii* transformed with the pPICZαB empty vector (lane 1) and 2 independent expressed clones of each construct for pPICZαB harboring CnAGal (wild type, lane 2–3; Δexon3-S137R, lane 4–5; R216T, lane 6–7; P250R, lane 8–9). Lane M represents protein marker. The black arrow indicates approximated position of 43 kDa protein size. Number labels indicate protein sizes marker in kDa.

the conformational changes of protein, leading to decreased protein stability. The presence and absence of all recombinant mutant proteins were confirmed by their expression by 10 random transformants harboring wild-type and mutant genes (Supplementary Fig. S2). In curd coconuts, heterologous expression results of mutant *AGal* genes are accordance with curd coconut lacking of AGal activity *in vivo* (Balasubramaniam and Mathew, 1986).

2.5. Differential expression of cnAGal in normal and curd coconut

Differential transcription level study of wild-type and mutant *AGal* genes in mature solid coconut endosperm at 12 months old age revealed that transcription level of mutant *AGal* genes in curd coconuts were unanimously higher than in normal ones. The fold differences were between 20 and 583 times of normal coconuts (Fig. 4 and Supplementary Table S1). The increment of transcription level of *AGal* in curd coconuts might be from lacking of feedback inhibition by sugar-mediated signal from metabolic substrates or products of galactomannan or galactose in cell wall degradation pathway at the transcription level or by some other signals from altered conditions in curd coconut endosperm. A similar event was found in *shrunken2* and *brittle2* genes encoding for ADP-glucose pyrophosphorylase subunits for starch accumulation in maize kernel. The more severe mutation the more increased transcription level of the genes. These also effected to transcription level of other related genes in the same pathway (Giroux et al., 1994). However, the precise mechanism to control expression level of *AGal* in coconut endosperm is required further studies.

2.6. Structure modelling of wild-type and mutant CnAGal proteins

In vitro expression studies, wild type and mutant recombinant *CnAGal* genes expressed in *K. phaffii* obviously showed the consequences of the curd phenotype caused by mutant *CnAGal* gene, resulting in decreased protein stability and enzyme activity levels, as described above. Due to its deduced amino acids of wtCnAGal showing 73% identity to rice AGal, modeled structure of wtCnAGal was established based 3-dimensional structure of rice AGal (blue color)

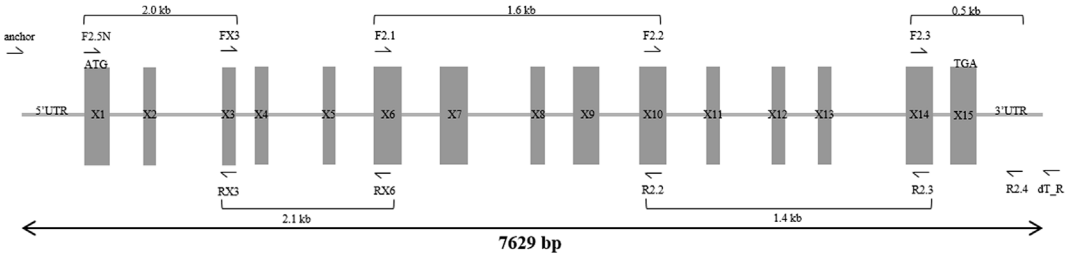


Fig. 6. Specific-gene primer binding sites, their location, and the size of each fragment amplified by each pair of primers used to obtain for the full length of α-galactosidase gene on genomic DNA sequence or cDNA. X represents "Exon". Size of PCR products can be estimated from given length and pairing of each forward primer (F) and each reverse primer (R). See Table 2 for sequence details.

containing α-galactose (yellow), which has been already elucidated by the x-ray crystallography technique (Fujimoto et al., 2003). The SWISS-MODEL web site (Guex and Peitsch, 1997; Schwede et al., 2003; Arnold et al., 2006) was used to predict the modeled structure of both wild-type and mutant CnAGal from their protein sequences. Modeled structure of wtCnAGal (green color) superimposed with rice AGal (blue color) containing α-galactose (yellow) was shown the small difference motif, as represent in red color. When compared with the modeled structure of wtCnAGal, that of Δexon3-S137R mutant showed the missing of conserved α1 helix domain, resulting in disruption of a tertiary protein structure (Fig. 7e). While S137R mutation appears to be quite close to the active site of wtCnAGal containing galactose (Fig. 7a), but unclear for its function. Thus, the synergistic effect of Δexon3 and S137R may cause totally losing enzyme stability and activity. Further studies may be required to investigate the roles of S137R at the active site by site-directed mutagenesis together with *in vitro* expression. In homology to wtCnGal, R216 in coconut, which is equivalent to R181 in rice, which serves as a ligand-binding residue by forming hydrogen

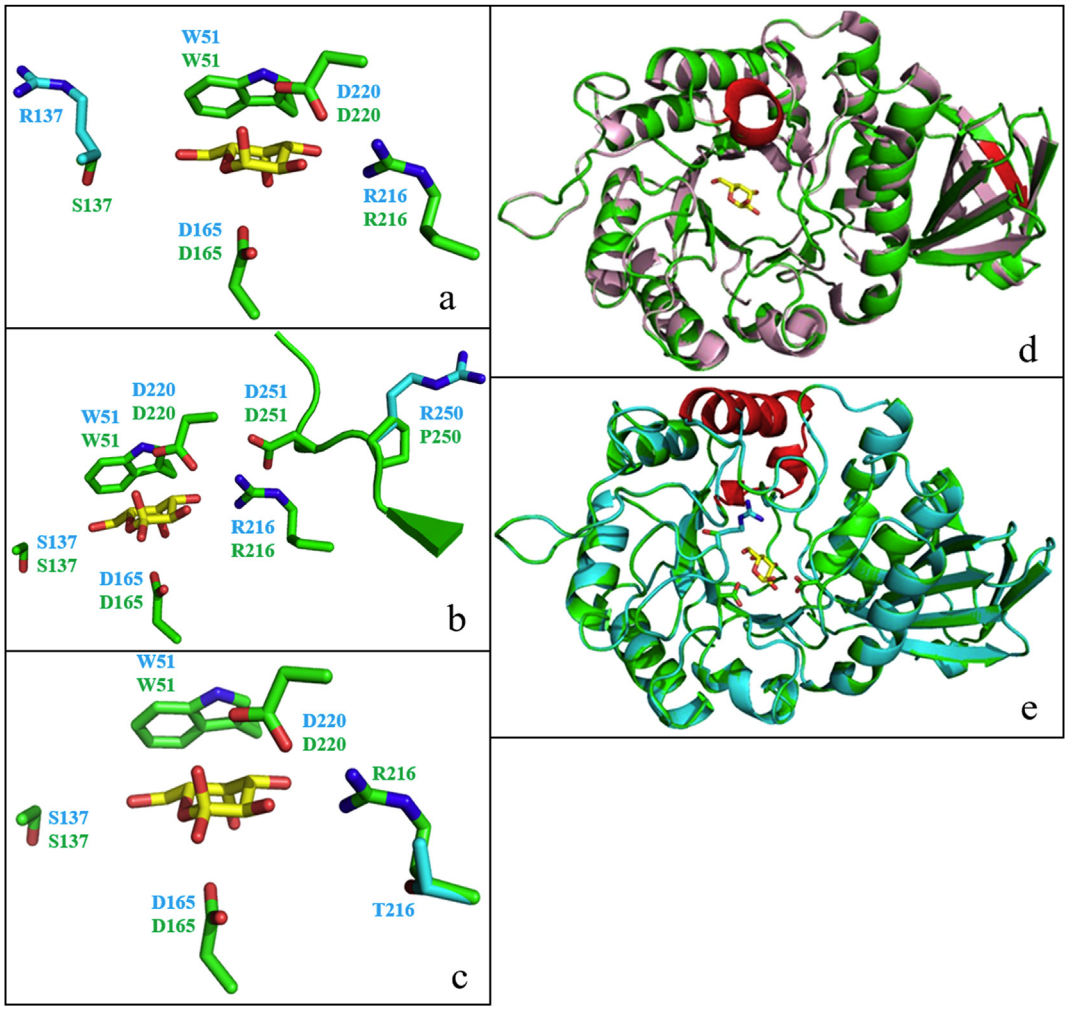


Fig. 7. Predicted modeled structure of wtCnAGal and mutant amino acid residue substitutions in comparison. The enzyme structure data was based on rice Agal. The green amino acids residues are wtCnAGal and blue residues are mutant conformation with proximity to α-galactose (yellow). The mutant proteins are Δexon3-S137R (a), P250R (b) and, R216T (c). The pink ribbon model is rice AGal conformation superimposed with green ribbon model from wtCnAGal. Red ribbon regions are part of different structure between rice and coconut (d). The pink ribbon model of Δexon3-S137R is superimposed with green ribbon model wtCnAGal. The red ribbon region is missing region in the conformation of Δexon3-S137R (e). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

bond with the oxygen atom of galactose (Fujimoto et al., 2003). Thus, losing of AGal activity of recombinant mutant R216T is involved in ligand-binding and stabilizing of enzyme-substrate complex during enzymatic catalysis (Fig. 7c). Modeled structure of P250R mutant showed no conformation change (data not shown) and P250 is not located in the catalytic pocket (Fig. 7b). However, P250 was in a well-conserved domain in AGal family 27 from many plants. Based on function of proline for the real protein folding, P250 may stabilize the loop structure of CnAGal. Therefore, P250R mutant may cause the regional or global enzyme conformation changes, resulting in loss of enzyme activity and decreased protein stability.

3. Conclusion

In normal and curd coconuts, full-length cDNA and parental genomic DNA sequences encoding endosperm specific AGal were sequenced and characterized. At least, three mutation alleles were identified in *CnAGal* gene. *In vitro* expression in *K. phaffii* showed all recombinant mutant protein containing no AGal activity. Loss of AGal activity in curd coconut causes no hydrolysis and accumulation of oligogalactomannan in endosperm, leading to curd coconut phenotype. Curd coconut phenotype is obvious in the late stage of coconut endosperm development that corresponds to degradation of galactomannan during germination stage of embryo in normal coconut endosperm.

4. Experimental

4.1. Plant materials and chemicals

For the purposes of total RNA and DNA extraction and enzyme activity staining, normal coconut fruits (cultivar Maphrao Yai; Thai Tall) at the 11–12 month developmental stage, were obtained from Rubber Research Station in Khan Thuli Commune, Tha Chana District, Surat Thani province and a private plantation in Nakhon Chaisri District, Nakhon Pathom province, Thailand. Curd coconuts were collected from two different plants: plant no. 1 represented typical curd coconut that its fruits incapable of germination and the other plant no. 10, the only plant that some fruits capable of germination; both from a curd coconut plantation on an isolated island in Vajiralongkorn dam, Thong Pha Phum District, Kanchanaburi province, Thailand. X-α-Gal (5-Bromo-4-Chloro-3-indolyl-α-D-galactopyranoside) was purchased from Clontech; *p*-nitrophenol (pNP) *p*-nitrophenyl-α-D-galactopynoside (*p*NPGal), and yeast nitrogen base from Sigma-Aldrich. Tryptone and yeast extract from Scharlab. Restriction enzymes were from New England Biolabs. *E. coli* strain DH5α and pGEM®_T Easy vector were from Promega. PCR primers were synthesized from IDT, as shown in Table 2. *Komagataella phaffii* (*Pichia pastoris*) strain GS115 and *Pichia* expression vector (pPICZαB) were from Invitrogen.

4.2. Enzyme activity staining

Solid endosperms or embryo parts of normal and curd coconuts at

Table 2
Degenerated and gene specific primers and their applications for coconut α-galactosidase cDNA, gDNA synthesis and qPCR.

Primer name	Sequence 5′-3′	Application
ALGALCOCO_F	AAGCTYGGVATYTACTCHGATGC	Degenerated primers for amplification of α-galactosidase gene.
ALGALCOCO_R	CATBCCHCCATTTCCCACYTCMA	
CnAlGal_F1	ACVCCKCMSATGGGGTGGAAYAGC	
CnAlGal_F2	GGYGGMTGGAAYGAYCCKGACATG	
CnAlGal_R2	CATGTCMGGRTCRTTCCAKCCRCC	
SKdT_R	CGCTCTAGAACTAGTGGATCCTTTTTTTTTTTTTTTTTTTTTTTTVN	3′ end cDNA synthesis.
OligodT-anchor primer	GACCACGCGTATCGATGTCGACTTTTTTTTTTTTTTTV	5′ end cDNA amplification.
Anchor primer	GACCACGCGTATCGATGTCGAC	
F2.5N	GGGGCTATGGTGCTATCCTTCTTC	Specific-gene primers for 5′/3′ RACE, full length cDNA synthesis, amplifying each fragment of α-galactosidase isoform 2 gene on gDNA sequences, and sequencing.
ALPHACOCO_F2.2_5′OUT	TTCCTGCTCCTTGTTTCGACGGCCGTG	
ALPHACOCO_F2.1START	ATGGCCGGTCGAGTCATGAAGG	
ALPHACOCO_FX2	CAGCTGGAATCACTTCGCTTGCAACAT	
ALPHACOCO_FX3	ATGCACTCGTATCCACTGGTCTCGC	
ALPHACOCO_F2.1	CACTTGGACATGAAGATCAAGATGCAAAGA	
ALPHACOCO_F2.2	TCTCGTGCAGATGAGAATGACAAATGGGC	
ALPHACOCOF_EXON6	GCTCTTCTGAACTCTGGGAGAACC	
ALPHACOCO_F2.3	GGTAGGGTAGCTGTGGTCCTATGGAA	
ALPHACOCO_RX2	TCCTTAATCATTTGCTCGTCGATG	
ALPHACOCO_RX3	ATTAACATACTGATACCCCAGCCTG	
ALPHACOCO_RX5	AAATTCCAAGCTTTAGCCCTTTTGC	
ALPHACOCO_RX6	GGTCTTTGCATCTTGATCTTCATGTCCA	
ALPHACOCO_R2.1N	CTGCATCAGAGTAAATCCCAAGCTT	
ALPHACOCO_R2.1	TCCTGTGGACTGCTGCCATCATTATAAC	
ALPHACOCO_R2.2	GCCCATTTGTCATTCTCATCTGCACGAGA	
ALPHACOCO_R2.3	TTCCATAGGACCACAGCTACCCTACC	
ALPHACOCO_R2.3_3′OUT	CATTATTCCTAATATCCAGAAACGCTGAGCA	
R2.5N	GGTATAGTTGCCACTAGAGTAGCACATC	
ALPHACOCO_R2.4	GATGCCAATTAGAAGCACACAAGCTC	
aGalXhoF	CCCCCCCTCGAGAAAAGAGAGGCTGAAGCTGGTCGAGTCATGAAGGATGCCGAA	PCR amplification and cloning of *cnAGal* gene into *Pichia* expression vector
aGalXbaR	CCCCCCTCTAGATCACTTTGGTGTCAAGACATACA	
aGalXbaHR	CCCCCCTCTAGATCAATGATGATGATGATGCTTTGGTGTCAAGACATACA	
qAGalF	GGGGCTATGGTGCTATCCTTCTTCT	Specific *cnAGal* target gene and internal control *β-actin* gene for qPCR analysis
qAGalR	CAAGGCCATTCTCCAGCATCTCTC	
Co-ActinF	ATGGTGAAGGCTGGATTTGCTGG	
Co-ActinR	GCATCCTTTTGGCCCATCCC	

different developmental stages (7–13 months) were stained for expression patterns of AGal. First, the coconut tissues were cut into thin layers as longitudinal section with 1–2 cm diameter. The staining solution (0.1 M Na-citrate buffer pH 5.5, 2 mM $K_3Fe(CN)_6$, 0.05% v/v Triton X-100, 50 μg/mL X-α-gal) was freshly prepared. Coconut tissues were soaked in the staining solution overnight at 30 °C with gentle shaking. The staining solution was removed the next day and 70% ethanol was added to wash coconut tissues until they became white with blue precipitation. Stained tissues were preserved in 50% ethanol and 50% glycerol (v/v) after photograph taken.

4.3. Total RNA extraction

Total RNA was extracted from coconut tissue (mature endosperm, embryo, and testa layer) using a combination of Fruit-mate™ buffer (TAKARA Bio Inc.) and Trizol reagent (Invitrogen). About 100 mg of freshly chopped tissue was added with 1 mL of Fruit-mate™ buffer and ground with grinding sand. The fine slurry was collected and centrifuged at 13,000*g* for 5 min. An equal amount of supernatant was divided and transferred into two 1.5 mL tubes, and 0.5 mL of Trizol reagent were added to each tube. The extraction step further performed following the instruction manual of Trizol reagent RNA extraction kit. Total RNA was subsequently used as a template for the first strand cDNA synthesis.

4.4. PCR amplification of partial fragment of cDNA encoding CnAGal gene

Total RNA was used for the synthesis of the first strand cDNA using Superscript III reverse transcriptase and the oligo$(dT)_{20}$ primer. The degenerate forward primer, namely CnAlGal_F1 was designed based on the alignment of conserved deduced amino acids from the nucleotide sequences of plant family 27 AGals classified in including *Coffea arabica* (AJ877911.1), *Coffea arabica* (L27992.1), *Coffea canephora* (AJ877912.1), *Nicotiana tabacum* (HQ877670.1), *Arabidopsis thaliana* (AY093200.1), *Oryza sativa* (AB039671.1), *Zea may* (NM_001153890.1). The reverse primer SKdT_R was used in order to amplify the partial fragment of *CnAGal* gene. The PCR-amplified products were then cloned into the pGEM-T vector (Promega, USA), transformed into *E. coli* DH5α, and DNA sequencing carried out by the Macrogen company.

4.5. Cloning of full-length CnAGal gene by rapid amplification cDNA end (RACE)

The 5′/3′-RACE method was performed according to the Roche's RACE kit 2nd generation instruction protocol. For 3′-RACE, the CnAlGal_F1 and ALPHACOCO_F2.1 primers were designed using the partial DNA sequence of *CnAGal* gene. The first strand cDNA was synthesized by Superscript III reverse transcriptase (Invitrogen) using Oligo $(dT)_{20}$ primer and total RNA as a template. The first strand cDNA was used as a template together with CnAlGal_F1 and Oligo$(dT)_{20}$ primers to firstly amplify the 3′ end of cDNA then this amplified product was used as a template together with ALPHACOCO_F2.1 and SKdT_R primers in a nested PCR to obtain the amplicon of 3′-RACE. The ALPHACOCO_R2.1 and ALPHACOCO_R2.2 specific primers used for 5′-RACE were designed from the partial DNA sequence of *CnAGal* gene together with the first strand cDNA synthesized by the activity of Transcriptor reverse transcriptase (Roche) using an ALPHACOCO_R2.1 primer and total RNA as a template, and purified by the High Pure PCR Product Purification Kit (Roche). The poly (A) tailing was added to the purified first stand cDNA at 3′ end by recombinant Terminal transferase enzyme, according to the Roche's instruction protocol. This purified first stand cDNA template with poly (A) tail together with the ALPHACOCO_R2.2 and OligodT-anchor primers were used for the PCR-amplification of 5′-RACE product. PCR-amplified products of 3′-RACE and 5′-RACE were then gel-purified, ligated into the pGEM-T easy vector and transformed into competent *E. coli* strain DH5α. Proper construction of the pGEM-T easy vector harboring partial *CnAGal* gene was confirmed by DNA sequencing. The full-length cDNA was PCR-amplified by using two pairs of specific primers: ALPHACOCO_F2.25′OUT & ALPHACOCO_R2.4 and F2.5N & ALPHACOCO_R2.3_3′OUT, designed from 5′ and 3′ cDNA fragments. Total RNA was used for the synthesis of the first strand cDNA using Superscript III reverse transcriptase and the oligo$(dT)_{20}$ primer. The first strand cDNA as a template together with ALPHACOCO_F2.2_5′OUT and ALPHACOCO_R2.3_3′OUT primers were used for amplifying the full-length *CnAGal* cDNA by PCR using KOD hot start DNA polymerase (Merck). F2.5N ALPHACOCO_R2.4 primers and the first amplified product used as a template were further used for the Nested PCR amplification using KOD hot start DNA polymerase (Merck) to obtain the full-length cDNA encoding *CnAGal*. The PCR-amplified products were cloned into pGEM-T easy and then sequenced. The theoretical molecular weight of proteins were determined by ExPASy web base to compute pI and MW (http://web.expasy.org/compute_pi/).

4.6. Genomic DNA extraction and PCR amplification

Genomic DNA was extracted from approximately 100 mg fresh maternal coconut leaf, both normal and curd, using the NucleoSpin® Plant II kit (Macherey-Nagel) following the PL1 buffer method. DNA quality and quantity were checked by running on 1% agarose gel or by Nanodrop 2000 spectrophotometer (Thermo Fisher Scientific) at wavelengths 260 nm and 280 nm. The extracted genomic DNA sample was diluted to 1:20 (to a final concentration around 20–30 ng/μL) and used as template for PCR amplification. All fragments of AGal gene were amplified by Phusion DNA polymerase (Thermo Fisher Scientific) or KOD hot start DNA polymerase using the primer pairs shown in Fig. 6. Target PCR product bands were cut to purify, A-overhangs added prior to cloning into pGEM®_T Easy vector and then subjected to DNA sequencing. Alignment of DNA and protein sequences of coconut AGal among wild type and mutant or with other structurally related plant AGals was performed using the Bioedit program.

4.7. Quantitative real time PCR analysis

To observe differential expression of *CnAGal* between wild-type and curd coconuts, 3 independent of normal and curd coconut fruits at age about 12-month-old were extracted RNA from solid endosperm and converted to first strand cDNA as describe in section 4.3–4.4. The primers for *CnAGal* (target) and *Cnβ-Actin* (control) genes were designed specifically for qPCR amplification (Table 2). cDNA and specific primers were added to KAPA SYBR® FAST Universal 2X qPCR Master Mix (KAPA Biosystems), and qPCR was performed using Mx3000P qPCR System (Agilent Technologies) following the standard protocol as described in Traewachiwiphak et al. (2018). Data analysis was carried out by $2^{-\Delta\Delta Ct}$ method. Control without template or reverse-transcription were included in the calculation.

4.8. Construction of recombinant yeast expression plasmid harboring CnAGal gene

The cDNA coding for mature *CnAGal* gene was amplified from pGEM-T Easy-α-Gal II clones with two pairs of primers (aGalXhoF & aGalXbaR). The PCR products and the pPICZαB vector were digested with *Xho*I and *Xba*I, and ligated together between the *Xho*I and *Xba*I sites. These plasmids were transformed to *E. coli* DH5α, and grown on LB with 2 μg/mL bleomycin agar plates. Positive clones were selected for inoculation and plasmid extraction. The pPICZαB vector, the recombinant pPICZαB/*wtCnAGal* and pPICZαB/*mtCnAGal* plasmids (5–10 μg) were linearized by *Sac*I, and then electroporated into electrocompetent *K. phaffii* GS115 cells by electroporation using 0.2 cm-gapped cuvettes at 1500 V, 5 ms single square wave pulse with Gene Pulser (Bio-Rad Laboratories). Bleomycin-resistant transformants were

screened from YPD plates containing 6 μg/mL bleomycin.

4.9. Expression of recombinant wild type and mutant CnAGal in K. phaffii

The selected positive clones of *K. phaffii* integrants harboring pPICZαB/*wtCnAGal,* pPICZαB/*mtCnAGal*, and pPICZαB were cultured in the same condition as follows: *K. phaffii* integrants harboring recombinant plasmid were inoculated in BMGY broth (1% yeast extract, 2% peptone, 1% glycerol, 400 μg/L biotin, and 0.1 M potassium phosphate, pH 6.0) and then cultivated in a shaking incubator at 30 °C and 250 rpm, until OD_{600} reach to 2–6. The cells were harvested and resuspended in BMMY medium (1% yeast extract, 2% peptone, 400 μg/L biotin, 1% methanol, and 0.1 M potassium phosphate, pH 6.0) (EasySelect Pichia Expression kit, 2010). To maintain production of recombinant AGal, methanol at 0.5% final concentration was added in the culture medium every 24 h to induce protein secreted into the medium under plasmid AOX promoter (Daly and Hearn, 2005). At day 5, cell-free supernatant (CFS) collected by centrifugation were analyzed by 10% SDS-PAGE and enzyme activity assays.

4.10. Prediction of modeled structure of CnAGal'

The SWISS-MODEL web site (Guex and Peitsch, 1997; Schwede et al., 2003; Arnold et al., 2006) was used to predict the modeled structure of both wild-type and mutant CnAGal from their protein sequences.

4.11. Enzyme assay

The AGal activity was assayed by incubation of 50 μL of the enzyme solution with 20 μL of 0.1 M sodium citrate buffer pH 5.5 and 30 μL of 2 mM *p*NPGal for 10 min at 30 °C. The reaction was stopped by adding 300 μL of 0.5 M Na_2CO_3. The release of *p*-nitrophenol was measured by the absorbance at 405 nm. A unit of enzyme activity is defined as the amount of enzyme releasing 1 μmole of *p*-nitrophenol per second at pH 5.0 and 30 °C. Specific activities are expressed as units per mg protein. Protein concentrations of the enzyme solution were determined with the absorbance at 280 nm.

4.12. SDS-PAGE analysis

10% SDS–PAGE was performed according to the method (Laemmli, 1970) and each lane was loaded with CFS containing 500 μg of total protein. Gels were stained in Coomassie Brilliant Blue R-250.

Acknowledgements

This work was financially supported by the ASEAN UNINET On-Place Thailand Scholarship Program, the Thailand Research Fund-Commission of Higher Education-Mahidol University MRG5380133, and the New Researcher Supporting Fund, Faculty of Science, Mahidol University, Thailand, . D.A. received the Dean's fund 2011 from the Faculty of Tropical Medicine, Mahidol University, Thailand. We also cordially thank Mr. Somchai Wattanayothin, the former head of coconut breeding program, Department of Agriculture, Ministry of Agriculture and Cooperative, Thailand for normal coconut samples, Dr. Utai Charanasri, the owner of the curd coconut island for curd coconut samples.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.phytochem.2018.08.015.

References

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.
Ashburner, G.R., Faure, M.G., Tomlinson, D.R., Thompson, W.K., 1995. A guide to the zygotic embryo culture of coconut palms (*Cocos nucifera* L.). ACIAR Techn. Rep. 36, 2–16.
Balasubramaniam, K., 1976. Polysaccharides of the kernel of maturing and mature coconuts. J. Food Sci. 41, 1370–1373.
Balasubramaniam, K., Mathew, C.D., 1986. Purification of α-D-galactosidase from coconut. Phytochemistry 25, 1819–1821.
Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795.
Bewley, J.D., 1997. Breaking down the walls- a role for endo-beta-mannanase in release from seed dormancy? Trends Plant Sci. 2, 464–469.
Buckeridge, M.S., 2010. Seed cell wall storage polysaccharides: models to understand cell wall biosynthesis and degradation. Plant Physiol. 154, 1017–1023.
Daly, R., Hearn, M.T.W., 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. J. Mol. Recogn. 18, 119–138.
De Guzman, E.V., Del Rosario, D.A., 1964. The growth and development of *Cocos nucifera* L. Makapuno embryo *in vitro*. Philipp. Agric. 48, 82–94.
Du, F., Zhu, M., Wang, H., Ng, T., 2013. Purification and characterization of an α-galactosidase from *Phaseolus coccineus* seeds showing degrading capability on raffinose family oligosaccharides. Plant Physiol. Biochem. 69, 49–53.
EasySelect *Pichia* expression kit, 2010. For Expression of Recombinant Protein Using PPICZ and PPICZα in Pichia pastoris. Manual part no. 25-0172. Invitrogen Life Technology.
Feurtado, J.A., Banik, M., Bewley, J.D., 2001. The cloning and characterization of α-galactosidase present during and following germination of tomato (*Lycopersicon esculentum* Mill.) seed. J. Exp. Bot. 52, 1239–1249.
Flavier, M.E., 1999. Makapuno (An Abnormal Endosperm Phenotypes of Coconut) Endosperm Wall Polysaccharides: Isolation, Characterization and Industrial Application. Plant Physiology and Biochemistry, University library of the Philippines at Los Baños.
Foale, M., 2003. The coconut odyssey: the bounteous possibilities of the tree of life. ACIAR Monogr. 101, 33–35.
Fujimoto, Z., Kaneko, S., Momma, M., Kobayashi, H., Mizuno, H., 2003. Crystal structure of rice α-Galactosidase complexed with D-galactose. J. Biol. Chem. 278 (22), 20313–20318.
Giroux, M.J., Boyer, C., Feix, G., Guex, Hannah, L.C., 1994. Coordinated transcriptional regulation of storage product genes in the maize endosperm. Plant Physiol. 106, 713–722.
Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. Electrophoresis 18, 2714–2723.
Henrissat, B., 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 280, 309–316.
Islam, M.N., Cedo, M.L., Namuco, L.O., Borromeo, T.H., Aquilar, E.A., 2009. Effect of fruit age on endosperm type and embryo germination of Makapuno coconut. Gene Conserv. 8, 708–722.
Kim, W.D., Kobayashi, O., Kaneko, S., Sakakibar, Y., Park, G.G., Kusakabe, I., 2002. Alpha-galactosidase from cultured rice (*Oryza sativa* L. var Nipponbare) cells. Phytochemistry 61, 621–630.
Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680–685.
Minic, Z., 2008. Physiological roles of plant glycoside hydrolases. Planta 227, 723–740.
Mujer, C.V., Tamirez, D.A., Mendoza, E.M.T., 1984. α-D- galactosidase deficiency in coconut endosperm: its possible pleiotropic effects in Makapuno. Phytochemistry 23, 893–894.
Perera, L., Russell, J.R., Provan, J., McNichol, J.W., Powell, W., 1998. Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L.) accessions from Sri Lanka by means of AFLP profiling. Theor. Appl. Genet. 96, 545–550.
Perera, L., Russell, J.R., Provan, J., Powell, W., 2003. Studying genetic relationships among coconut varieties/populations using microsatellite markers. Euphytica 132, 121–128.
Reid, J.S.G., Meier, H., 1973. Enzymatic activities and galactomannan mobilization in germinating seeds of fenugreek (*Trifonella foenum-graecum* L. Leguminosae). Secretion of α-galactosidase and β-mannosidase by the aleurone layer. Planta 112, 301–308.
Rohde, W., Sniady, V., Herrán, A., Estioko, L., Sinje, S., Marseillac, N., Berger, A., Lebrun, P., Becker, D., Kullaya, A., Rodrigue, J., Billotte, N., Ritter, E., 2002. Construction and exploitation of high-density DNA marker and physical maps in the perennial tropical oil crops coconut and oil palm: from biotechnology towards marker-assisted breeding. Burotrop Bull. 20, 13–14.
Samonte, J.L., Mendoza, E.M.T., Ilag, L.L., De La Cruz, N.B., Ramirez, D.A., 1989. Galactomannan degrading enzymes in maturing normal and makapuno and germinating normal coconut endosperm. Phytochemistry 28 (9), 2269–2273.
Santoso, U., Kubo, K., Ota, T., Tadokoro, T., Maekawa, A., 1996. Nutrient composition of kopyor coconuts (*Cocos nucifera* L.). Food Chem. 57 (2), 299–304.
Schwede, T., Kopp, J., Guex, N., Peitsch, M.C., 2003. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res. 31, 3381–3385.
Singh, N., Kayastha, A.M., 2012. Purification and characterization of α-galactosidase from white chickpea (*Cicer arietinum*). J. Agric. Food Chem. 28, 3253–3259.
Traewachiwiphak, S., Yokthongwattana, C., Ves-Urai, P., Charoensawan, V.,

Yokthongwattana, K., 2018. Gene expression and promoter characterization of heat-shock protein 90B gene (*HSP90B*) in the model unicellular green alga *Chlamydomonas reinhardtii*. Plant Sci. 272, 107–116.
Wattanayothin, S., 2004. Study on hybrids of MaphraoKathi in Thailand. TNCEL Newslett. 2, 14–15 (in Thai).
Wattanayothin, S., 2005. Study on hybrids of MaphraoKathi in Thailand. TNCEL J. l3, 6–7 (in Thai).
Zhu, A., Goldstein, J., 1994. Cloning and functional expression of a cDNA encoding coffee bean α-galactosidase. Gene 140, 227–231.
Zhu, A., Leng, L., Monahan, C., Zhang, Z., Hurst, R., Lenny, L., Goldstein, J., 1995. High level expression and purification of coffee bean α-galactosidase produced in the yeast *Pichia pastoris*. Arch. Biochem. Biophys. 324 (1), 65–70.
Zuniga, L.C., 1953. The probable inheritance of the Makapuno character of the coconut. Philipp. Agric. 36, 402–413.