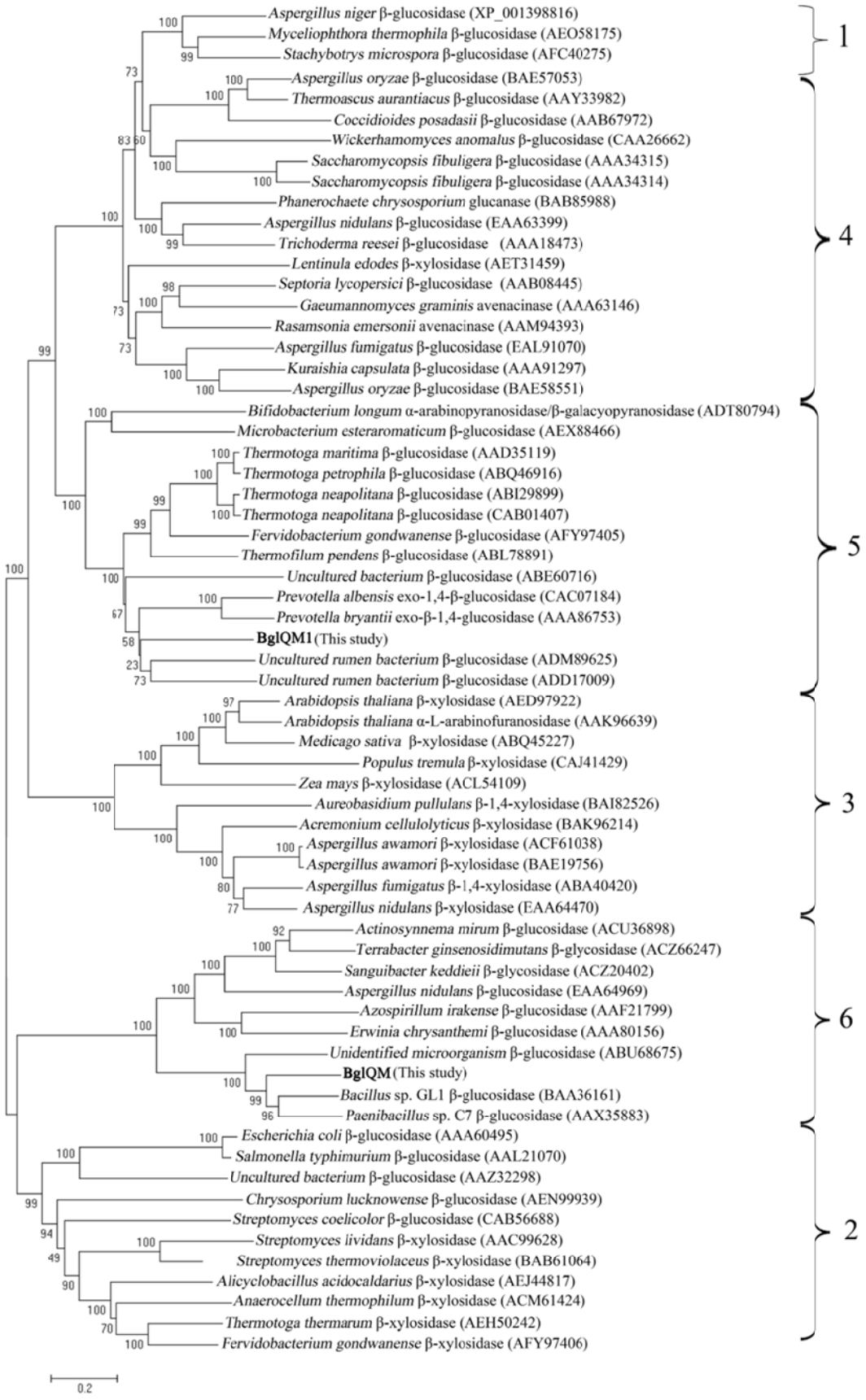
Supplementary Data

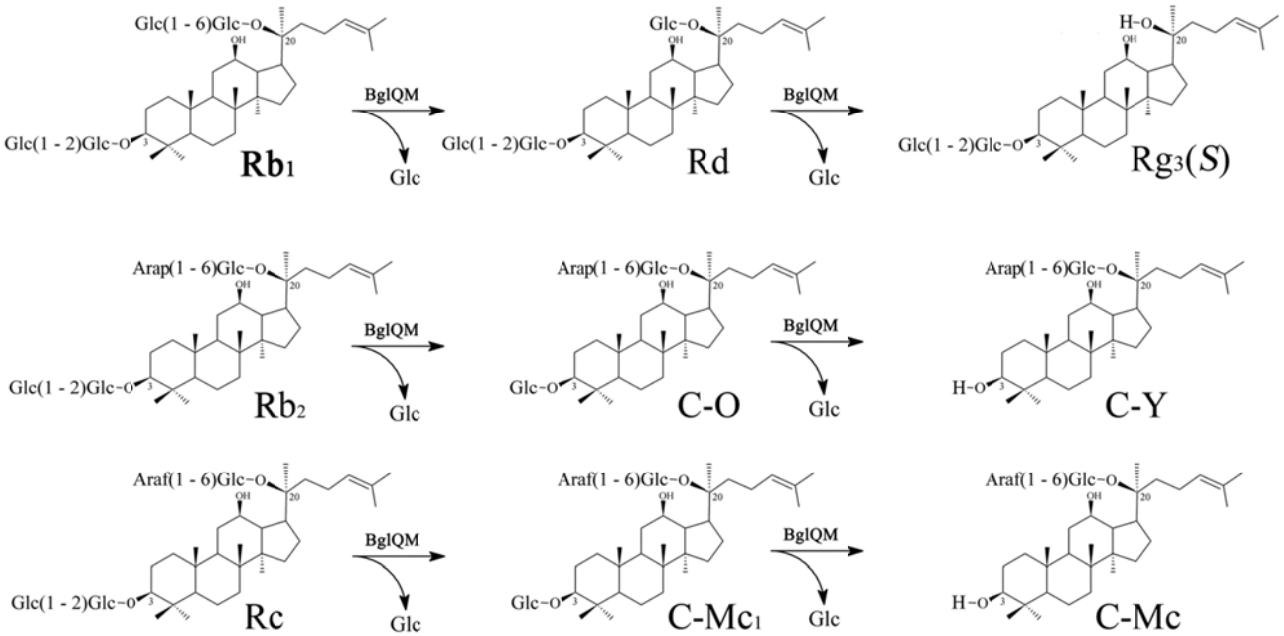
Phenotypic characteristics of strain QM49

Cells are Gram-negative, aerobic, non-spore-forming, non-motile, and rod shaped, 0.3-0.5 μm in diameter and 1.5-2.4 μm in length after culture on R2A agar for 3 days. Colonies grown on R2A agar for 3 days are smooth, circular, translucent, creamy-white and convex. Grows at 4-37°C and in pH 6.0-8.5, but not at 42°C. Optimum growth occurs at 25-30°C and pH 7.0. Growth occurs in the absence of NaCl and in the presence of 1.0% (w/v) NaCl, but not 2.0% (w/v) NaCl. Growth occurs on nutrient agar but not on MacConkey agar. Catalase-positive and oxidase-positive. Nitrate is reduced to nitrite in aerobic conditions. Urease, arginine dihydrolase, β-Galactosidase and β-glucosidase are positive. Protease activity and indole production are negative. Does not produce any acid or gas from glucose. The following compounds are utilized as sole carbon sources: D-glucose, D-mannose, D-maltose, salicin, D-melibiose, L-fucose, L-arabinose, L-rhamnose, *N*-acetyl-D-glucosamine, D-sucrose, D-maltose, and glycogen. The following compounds are not utilized as sole carbon sources: D-mannitol, gluconate, caprate, adipate, malate, citrate, phenyl-acetate, D-sorbitol propionate, caprate, valerate, L-histidine, 2-ketogluconate, 3-hydroxy-butyrate, 4-hydroxy-benzoate, L-proline, D-ribose, itaconate, suberate, malonate, acetate, lactate, L-alanine, 5-ketogluconate 3-hydroxy-benzoate, and L-serine.

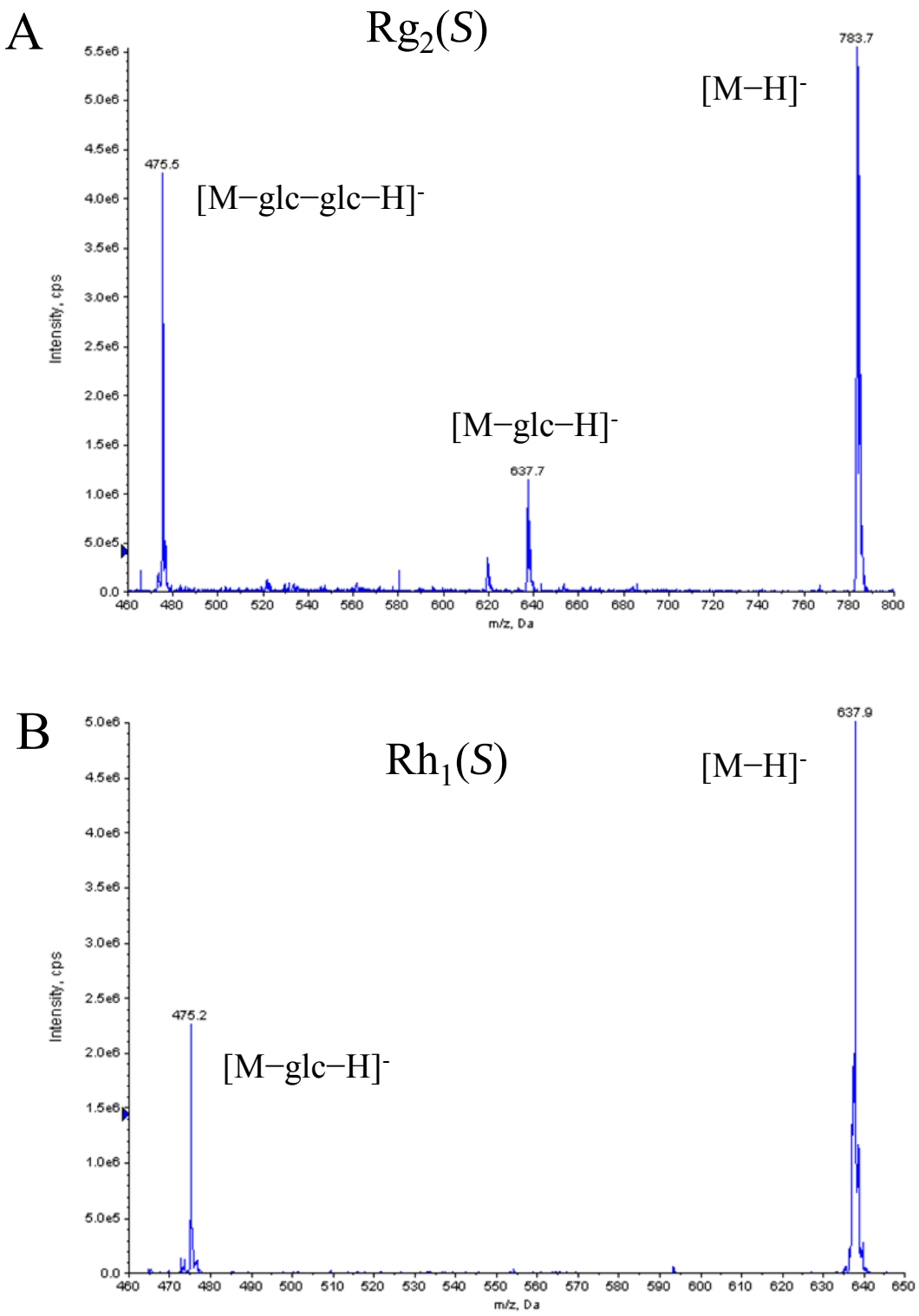


Supplementary Fig. S1.

Supplementary Fig. S1. Phylogenetic analysis of characterized glycoside hydrolases family 3. Amino acid sequences were obtained from the NCBI/EMBL database and CAZy database (accession numbers are indicated on the tree). This tree was made using the neighbor-joining method (4) with a Kimura two-parameter distance matrix (2) and pairwise deletion. Bootstrap values expressed as percentages of 1000 replications greater than 50% are shown at the branch points. The bar represents 20 amino acid residues substitutions per 1000 amino acid residues.



Supplementary Fig. S2. Transformation pathways of ginsenosides Rb_1, Rb_2, Rc, and Rd by recombinant BglQM, respectively.



Supplementary Fig. S3. Negative ion ESI–MS–MS spectra of biotransformed ginsenosides by BglQM: A, MS–MS spectrum of ginsenoside $Rg_2(S)$; B, MS–MS spectrum of ginsenoside $Rh_1(S)$.

Table S1. Isolated strains with ginsenoside-converting abilities

Strain name	Most closest type strain	Similarity[a] (%)	b.p. [b]	Re	Rg_1	Rb_1
				Transformed ginsenosides		
QM01	*Paenibacillus glycanilyticus* DS-1^T	99.3	728	Rg_2(*S*)	Rh_1(*S*)	Rd
QM04	*Burkholderia soli* GP25-8^T	99.0	728	Rg_2(*S*)	Rh_1(*S*)	Rd, F_2, C-K
QM05	*Dyella koreensis* BB4^T	99.6	734	-	-	Rd, F_2
QM06	*Dyella koreensis* BB4^T	99.6	734	-	Rh_1(*S*)	Rd, F_2
QM08	*Dyella koreensis* BB4^T	99.6	734	-	-	Rd, F_2
QM12	*Dyella koreensis* BB4^T	99.6	734	-	Rh_1(*S*)	Rd, F_2
QM15	*Bacillus acidiceler* CBD 119^T	99.7	738	PPT	PPT	Rd
QM18	*Asticcacaulis biprosthecium* ACM 2498^T	98.8	673	-	-	Rd, F_2, C-K
QM20	*Dyella koreensis* BB4^T	99.6	734	-	-	Rd, F_2, C-K
QM21	*Dyella koreensis* BB4^T	99.6	734	-	-	F_2
QM26	*Dyella japonica* XD53^T	98.5	734	-	-	Rd, F_2
QM28	*Dyella japonica* XD53^T	98.5	734	-	-	Rd, F_2
QM45	*Dyella koreensis* BB4^T	99.6	734	-	Rh_1(*S*)	F_2
QM48	*Dyella koreensis* BB4^T	99.6	734	-	Rh_1(*S*)	F_2
QM49	*Mucilaginibacter gossypiicola* Gh-48^T	99.9	1450	Rg_2(*S*)	Rh_1(*S*)	C-K

[a]Similarity is based on 16S rRNA gene sequences.

[b]b.p.: base pair

Table S2. Comparison of conserved sequence motifs of BglQM with family 3 Glycoside hydrolases*.

Glycoside hydrolases of Family 3 [a]	157-169[b]	231-244[c]	449-462[d]
BglQM (This study)	VGACIKHFVANNQ	EWGFEGFVMSDWYA	IVISRISGEGYDRK
U. mic	VGACIKHFVANNQ	EWGFEGFVMSDWYA	IVISRISGEGYDRK
B. GL1	VGACIKHFVVNEQ	EWGFEGFVMTDWFA	YILTRISGEGVDRK
P. C7	VGATLKHFAANDQ	EWGFDGVVMTDWGA	VVLYRVSGEGWDRR
P. TS12	VGTSLKHFAVNNQ	EWGHEGIVVSDWGA	GLPDRYESEGYDRT
T. nea	VGACIKHFVANNQ	EWGFEGFVMSDWYA	IVISRISGEGYDRK
A. acu	VVATAKHYILNEQ	ELGFQGFVMSDWGA	VFVNSDAGEGYISV
F.men	VGACIKHFVANNQ	EWGFDGFVMSDWYA	VVISRISGEGYDRK
K. mar	IAATVKHFVCNDL	EWKWDGMLMSDWFG	GLNGEWETEGYDRE
H. Vul	VAACAKHFVGDGG	TLKFKGFVISDWEG	AIVAVGEHPYTETK
		▲	

*. Multiple sequence alignment of BglQM and selected glycoside hydrolase family 3 enzymes by using Clustal Omega (1.1.0).

a. Full species names and Genbank IDs of the glycoside hydrolases in family 3 are as follows, *Mucilaginibacter* sp. QM49 β-glucosidase [BglQM (This studty)], JX403802; Unidentified microorganism β-glucosidase (U. mic), ABU68675; *Bacillus* sp. GL1 β-glucosidase (*B.* GL1), BAA36161; *Paenibacillus* sp. C7 β-glucosidase(*P.* C7), AAX35883; *Paenibacillus* sp. TS12 glucosylceramidase (*P.* TS12), BAC16750; *Thermotoga neapolitana* β-glucosidase (*T. nea*), ABI29899; Aspergillus aculeatus β-glucosidase (*A. acu*), BAA10968; Flavobacterium meningosepticum β-glucosidase (*F.men*), AAB66561; Kluyveromyces marxianus β-glucosidase (*K. mar*), ACY95404; *Hordeum vulgare* subsp. Vulgare β-D-glucan exohydrolase isoenzyme (*H. Vul*), AF102868. BglQM, U. mic, *B.* GL1 and *P.* C7 are in subfamily 6. The conserved residues of *P.* TS12, *T. nea, A. acu, F.men have been analysised* (1, 3, 5) .

b. Conserved sequence containing a putative carbohydrate-binding site (3). The conserved residues are shown in box.

c. Conserved sequence containing the catalytic nucleophile (▲) (1). The conserved residues are shown in box.

d. Conserved sequence containing the catalytic acid (1, 3, 5). The conserved glutamic residues are shown in box, including those that have been experimentally determined previously, which are highlighted in gray.

References:

1. **Joseph V, Maria H and Geoffrey F.** 1999. Three-dimensional structure of a barley b-D-glucan exohydrolase, a family 3 glycosyl hydrolase.Structure. **7:**179-190.
2. **Kimura, M.** 1983. The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, United Kingdom.
3. **Preeyanuch T, Lauren SM, Ana CA, Prachumporn TK and Harry B.** 2013. Identification of the acid/base catalyst of a glycoside hydrolase family 3 (GH3) β-glucosidase from Aspergillus niger ASKU28. BBA-GEN. SUBJECTS. **1830:** 2739–2749.
4. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:**406-425.
5. **Suzuki K, Sumitani J, Nam YW, Nishimaki T, Tani S and Wakagi T.** 2013. Crystal structures of glycoside hydrolase family 3 β-glucosidase 1 from Aspergillus aculeatus. Biochem. J. **452:** 211-221.