Genes & Genomics (2020) 42:1087–1096
https://doi.org/10.1007/s13258-020-00976-y

Online ISSN 2092-9293
Print ISSN 1976-9571

RESEARCH ARTICLE

Complete genome sequence of *Sphingobium* sp. strain PAMC 28499 reveals a potential for degrading pectin with comparative genomics approach

So-Ra Han[1] · Sung-Min Jang[2] · Young Min Chi[2] · Byeollee Kim[1] · Sang-Hee Jung[3] · Yung Mi Lee[4] · Jun Uetake[5] · Jun Hyuck Lee[6,7] · Hyun Park[2] · Tae-Jin Oh[1,8,9]

Received: 20 January 2020 / Accepted: 17 July 2020 / Published online: 31 July 2020

Abstract

Background *Spingobium sp.* PAMC 28499 is isolated from the glaciers of Uganda. Uganda is a unique region where hot areas and glaciers coexist, with a variety of living creatures surviving, but the survey on them is very poor. The genetic character and complete genome information of *Sphingobium* strains help with environmental studies and the development of better to enzyme industry.

Objective In this study, complete genome sequence of *Spingobium* sp. PAMC 28499 and comparative analysis of *Spingobium* species strains isolated from variety of the region.

Methods Genome sequencingwas performed using PacBio sequel single-molecule real-time (SMRT) sequencing technology. The predicted gene sequences were functionally annotated and gene prediction was carried out using the program NCBI non-redundant database. And using dbCAN2 and KEGG data base were degradation pathway predicted and protein prediction about carbohydrate active enzymes (CAZymes).

Results The genome sequence has 64.5% GC content, 4432 coding protein coding genes, 61 tRNAs, and 12 rRNA operons. Its genome encodes a simple set of metabolic pathways relevant to pectin and its predicted degradation protein an unusual distribution of CAZymes with extracellular esterases and pectate lyases. CAZyme annotation analyses revealed 165 genes related to carbohydrate active, and especially we have found GH1, GH2, GH3, GH38, GH35, GH51, GH51, GH53, GH106, GH146, CE12, PL1 and PL11 such as known pectin degradation genes from *Sphingobium yanoikuiae*. These results confirmed that this *Sphingobium* sp. strain PAMC 28499 have similar patterns to RG I pectin-degrading pathway.

Conclusion In this study, isolated and sequenced the complete genome of *Spingobium* sp. PAMC 28499. Also, this strain has comparative genome analysis. Through the complete genome we can predict how this strain can store and produce energy in extreme environment. It can also provide bioengineered data by finding new genes that degradation the pectin.

Keywords Carbohydrate active enzyme · Comparative genomics · Genome sequencing · Pectin degradation · *Sphingobium* sp.

Introduction

Bacteria of genus *Sphingobium*, family *Sphingomonadaceae*, phylum *Proteobacteria*, are usually Gram-negative, rod-shaped, non-sporulating, motile or non-motile, and strictly aerobic, and they have a color as yellow or whitish brown (Chaudhary et al. 2017; Yabuuchi et al. 1990). In the data obtained from phylogenetic, chemotaxonomic, and physiological analysis, the previously described *Sphingomonas* sensu stricto is now classified into four genuses: *Sphingomonas*, *Novospingobium*, *Sphingobium*, and *Sphopyxis* (Takeuchi et al. 2001). It has been reported that

So-Ra Han, Sung-Min Jang these two authors contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13258-020-00976-y) contains supplementary material, which is available to authorized users.

✉ Hyun Park
hpark@korea.ac.kr

✉ Tae-Jin Oh
tjoh3782@sunmoon.ac.kr

Extended author information available on the last page of the article

Sphingobium members have been isolated from a variety of habitats including oil-contaminated soil, hexachlorocyclo hexane dump site, rhizosphere soil, river sediment, cattle pasture soil, freshwater, copper mine soil, roots, fly ash dumping site, and waste water (Kumari et al. 2009; Maeda et al. 2015; Ushiba et al. 2003; Young et al. 2007).

Biological studies have generally shown frequent surveys of cold areas in the mid-latitudes and polar region (arctic and antarctic), but little has been studied for tropical glaciers except New Guinea (Hope et al. 1976). We can see glaciers that can be seen from the equatorial Africa (three major mountain regions: Mt. Kilimanjaro in Tanzania, Mt. Kenya in Kenya, and the Rwenzori Mountains in Uganda), but it is also true that many people are unaware of and not targeted in previous surveys of glacial biology. Moreover, it is currently on the verge of disappearing due to climate warming, reduced humidity, and reduced cloud. Among them, Rwenzori Mountains in Uganda is expected to disappear by 2020, but also surveys of the region are very insufficient (Mölg et al. 2006; Taylor et al. 2006; Uetake et al. 2015). Microbes separated from cold areas on Earth can produce enzymes with cold activity. It has been reported that some yeast and bacteria separated from Argentinian Patagonia, Himalayan regions, Iceland, and Japan can produce pectinase (Adapa et al. 2014). However, there is a considerable lack of information on enzymes that degradation of pectin that can have cold activity in it. Although the activity of medophilic and cold-active pectinolies has already been reported by filament bacteria, almost all are pectinolytic enzymes that perform optimal activities at 40–45 °C (Dinu et al. 2007; Gadre et al. 2003; Poveda et al. 2018; Saito et al. 2004; Takasawa et al. 1997).

Pectin is structurally the most complex component of the cell walls in the dicotyledonous plant. Pectin polysaccharides are present in the primary side walls and middle lamella, and are most abundant in the cell walls of the soft part of the plant and surrounding growing cells (Mohnen 2008). The structural units that produce pectin are usually divided into three main domains: HG (Homogalacturonan), RG I (Rhamnogalacturonan I), and RG II (Rhamnogalacturonan II), and HG and RG I are generally the most dominant quantitative (Voragen et al. 2001). Among them, RG I is the second most abundant polysaccharide of the pectin family. And this is produced of a backbone of alternating a 1,2-linked rhamnose and a 1,4-linked galacturonic acid residues, that is, of repeating disaccharide structural units. RG I degradation enzyme catalyzes the cleavage of galacturonic acid and rhamnose (and the bond between rhamnose and galacturonic acid) in the RG I backbone chain of pectin. They are active on the RG I backbone of pectin and thus strictly specific for cleaving bonds in the repetitive structure (Silva et al. 2016). RG I side chains can be organized in a variety of ways, allowing for highly variable placement, so this is responsible for the chemical and structural complexity of pectin and can constitute 20–35% of the total pectin mass (Mohnen 2008). Enzyme of pectin degradation has an advantage over chemical hydrolysis as the enzyme targets specific enemy bonds of pectin molecules while less specialized in chemical methods (Benen et al. 2002; Schols and Voragen 1996). This is particularly important when partial hydrolysis is required to produce a particular fragments. In order to meet the demand for a particular pectin-induced fructose, a professional mixture of pectinolytic enzymes which isolate and no longer hydrolyze the junctions required to produce such fructose is required (Benoit et al. 2012).

The structural complexity of pectin presents considerable difficulties to the human digestive system, and humans must rely on the coherent action of CAZymes produced by symbiotic gut bacteria to reduce pectin. Studies involving environmental and animal gut bacteria suggest that interception of pectin-degrading interlayers is a prerequisite for initiating plant cell wall degradation, exposing other cell wall polymers, and permitting more heterogeneous bacterial colonization along the plant cell wall (Cheng et al. 1979; Chung et al. 2014). In this study, the *Sphingobium* sp. strain named PAMC 28499. With functional structural analysis of strains, genome sequencing, assembly, and annotation were performed perfectly as a new report separated from the glaciers of Uganda. Uganda is a unique region where hot areas and glaciers coexist, with a variety of living creatures surviving, but the survey on them is very poor. This is the first report on the entire genome of bacteria separated from the glaciers of Uganda, and it is believed to affect many industries as the first report on pectin-degrading enzymes from the complete genome of this strain.

Materials and methods

Isolation of Sphingobium sp. PAMC 28499 and preparation of genomic DNA

Sphingobium sp. PAMC 28499 strain was isolated from the glaciers of Mt. Rwenzori, Uganda using 0.1XR2A Broth Gellan gum. This strain was isolated at 10 °C, but after checking the time it was grown using a pure R2A (MB cell Ltd., Seoul, Korea), it was found to grow up to 25 °C, and we chose grew enough cells at 15 °C and them to get DNA. Genomic DNA was extracted from *Sphingobium* sp. PAMC 28499 using a QIAamp DNA Mini Kit (Qiagen Inc., Valencia, CA, USA), and the quantity and purity were determined using a spectrophotometer (Biochrome, Libra S35PC, UK). The purity was A260/A280, and the genome was processed to determine whether it was good or not. The extracted DNA was checked by agarose gel electrophoresis to evaluate its quality and stored at – 20 °C.

Complete genome sequencing and annotation

Genome sequencing was performed using PacBio sequel single-molecule real-time (SMRT) sequencing technology (Pacific Biosciences, Menlo Park, CA, USA). SMRTbell library inserts (20 kb) were sequenced using SMRT cells. Raw sequence data were generated from 100,966 reads and 580,399,630 bp that were assembled de novo using the hierarchical genome-assembly process (HGAP) protocol (Chin et al. 2013) and RS HGAP4 Assembly in SMRT analysis software [ver. 2.3; Pacific Biosciences (https://github.com/PacificBiosciences/SMRT-Analysis)]. The coding DNA sequences (CDSs) were predicted and annotation was performed with the Rapid Annotation using Subsystem Technology (RAST) server (Aziz et al. 2008). The predicted gene sequences were translated and searched against the National Center for Biotechnology Information (NCBI) non-redundant database, the eggNOG 4.5 database (Huerta-Cepas et al. 2016), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Also identified the features through a circular map of the PAMC 28,760 genome created using the CGView comparison tool (Grant et al. 2012).

Carbohydrate active enzyme

CAZyme genes such as glycoside hydrolases (GH), glycosyl transferase (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), auxiliary activities (AA), and carbohydrate binding module (CBM) were predicted by running HMMER3 (Eddy 2011) scan using HMMs profile downloaded from dbCAN2 (version: dbCAN HMMdb release 7.0) with an e-value cutoff of 1e-15 and coverage cut off of > 0.35. Also we used DIAMOND (Buchfink et al. 2015) with E value < 1e−102 and Hotpep (Busk et al. 2017) with frequency > 2.6, Hits > 6, which were used to improve the accuracy of expected genes. In addition, the results of CAZyme gene clusters (CGCs) were checked and able to upgrade the genome information with help of a CAZy database (https://www.cazy.org/).

Comparative genomic analysis of CAZymes

Average nucleotide identity (ANI) analysis was performed to determine the relationship of *Sphingobium* species with other strain as predicted by complete genome sequence and confirmed of possibility in this comparative analysis. ANI values between the genomes were calculated using method based on MUMmer alignment (Richter and Rosselló-Móra 2009). There are already many genomes of *Sphingobium* sp. strains registered available in NCBI GenBank. To compare the CAZyme of the widely registered *Sphingobium* sp. strains, among have a complete genome about strains and this is only complete genome were chosen except for WGS to finally select 19 strains. We downloaded genome sequence for each 19 strains from NCBI database and as mentioned earlier re-annotated of CAZyme using the dbcan2 server. The CAZyme gene parts in the each other complete genome were identified and compared, and their function was also confirmed using BLASTP of NCBI.

Results

Complete genome sequence

The complete genome of *Sphingobium* sp. PAMC 28499 contains a circular chromosome of 4,880,615 bp with 64.5% GC content. The CDSs were predicted, and annotation was performed with the RAST server (Aziz et al. 2008). The predicted gene sequences were translated and searched against the NCBI nonredundant database, the EggNog/COG database, and the KEGG database. A total of 4,432 CDSs were predicted, and the coding region accounted for 90.2% of the *Sphingobium* sp. PAMC 28499 genome; 61 tRNA and 12 rRNA were predicted in the complete genome (Table S1). A total of 4,581 genes were assigned a putative function. The genes were classified into 26 COG functional categories by eggNOG 4.5 database (https://eggnog.embl.de). Figure S1 shows a circular map of the PAMC 28499 of *Sphingobium* sp. genome, which was generated using the CGView Comparison Tool (Grant et al. 2012). Using this tool, tRNA, rRNA, and GC contents were represented and verified on the genome map (Fig. S1A), and the characteristics of the gene were identified by presenting graphs for each category of 26 COGs (Fig. S1B). Checking with COG result of our strain, has larger number of genes (> 100) in category M (235), O (152), T (200), J (157), K (247), L (203), C (240), E (268), G (220), H (113), I (157), and P (265). Each of these categories is meaningful, with the following meanings: M, Cell wall/membrane/envelope biogenesis; O, Posttranslational modification, protein turnover, and chaperones; T, Signal transduction mechanisms; J, Translation, ribosomal structure, and biogenesis; K, Transcription; L, Replication, recombination and repair; C, Energy production and conversion; E, Amino acid transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; P, Inorganic ion transport and metabolism. Within these many categories, the result for many genes in the G category means that there are many genes associated with carbohydrate transport and metabolism and also suggests that our strains are related to CAZyme. Because this is carbohydrates play an important role in many essential metabolic pathways (Maughan 2009). And known components of bacterial cell walls, including peptidoglycan, lipid polysaccharide (LPS), and outer membrane proteins, are classified into M categories. So we can

see that more than 235 proteins can be largely divided into five types: peptidoglycan biosynthesis-related genes, lipid glycogen biosynthesis-related genes, OMP coding genes, unknown genes, among them, those that degradation enzyme pectin or peptidoglycan (Yang et al. 2007). These results help to predict the characteristics of the strain that degradation the cell walls and get energy.

Conjecture of functional impact about pectin degradation enzyme from PAMC 28499

Newly annotated using dbCAN2, we were able to investigate interesting CAZyme genes related to pectin. The genome of *Sphingobium* sp. PAMC28499 encodes 74 GHs, 49 GTs, 5 PLs, 28 CEs, 2 CBMs, and 7 AAs. CAZyme annotations are based upon the protein models derived from the annotation of the genomes. Based on CAZyme results, many related genes that degradation enzymes of pectin were annotated and detailed results are summarized in Table 1. The CAZyme annotation of strain identified 4 putative β-galactosidase (GH2), 2 putative endorhamnogalacturonase, 3 putative β-galactosidase (GH35), 9 putative GH43, 3 putative α-$_L$-arabinofuranosidase (GH51), 1 putative endo-1,4-β-galactanase, 2 putative unsaturated rhamnogalacturonyl hydrolase (GH105), 1 putative α-$_L$-rhamnosidase (GH106), 1 putative α-$_L$-rhamnosidase (GH145), 1 putative β-$_L$-arabinofuranosidase, 1 putative pectate lyase (PL1), 1 putative rhamnogalacturonan endolyase (PL11), and 2 putative rhamnogalacturonan acetylesterase. Among them, The GH43 family consists of enzymes with activities required for the degradation of cell wall polysaccharides such as hemicellulose and pectin. Inspection of the enzyme commission (EC) annotations for the most up-regulated GH43 genes shows that they encode xylan 1,4-β-xylosidases (EC 3.2.1.37), arabinan endo-1,5-α-$_L$-arabinosidases (EC 3.2.1.99), and α-$_L$-arabinofuranosidases (EC 3.2.1.55). In detail, we could investigate this division: 1 putative α-1,2-$_L$-arabinofuranosidase (GH43_29), 3 putative β-xylosidase/α-$_L$-arabinofuranosidase (GH43_1, GH43_9, and GH43_19), 1 putative arabinoxylan arabinofuranohydrolase (GH43_12), 1 putative endo-α-1,5-$_L$-arabinanase(GH43_17), 1 putative arabinosidase (GH43_18), 1 putative extracellular exo-α-(1 → 5)-$_L$-arabinofuranosidase (GH43_26), and 1 putative intracellular endo-α-(1 → 5)-$_L$-arabinanase (GH43_5). These genes predicted appears to have an arsenal of enzymes for digestion of plant cell wall polysaccharides, which may explain the prevalence of this organism in soil and rhizosphere habitats (Peterson et al. 2006). Many genes of CAZyme from PAMC 28,499 predicted to encode glycohydrolases, esterases, and lyase that could be involved in utilization of plant cell wall polysaccharide, such as xylans (β-1,4-linked polymers of xylose often substituted with acetyl, arabinofuranoside, and glucuronosyl residues), mannans (heteropolysaccharides containing β-1,4-linked mannose residues), and xyloglucans (β-1,4-linked polymers of glucose substituted with xylose and other sugars), were also identified. In the case of pectin among them, the backbone consists mainly of a $_D$-galacturonic acid residue joined by α-1,4-linked that can be replaced by methyl esters or acetyls. Pectin is classified into three general groups: homogalacturonan (linear polymer), xylogalacturonan (branded by β-1,3-linked $_D$-xylose), and rhamnogalacturonan (Caffall and Mohnen 2009; Wong 2008). In particular, RG I consists of repetitive disaccharides ramnose-galacturonic acid. Galacturon residue can be acetylated, and both residues can carry neutral sugar chains made of galactose, arabinose, and xylose (Willats et al. 2006). In the strain PAMC 28,499, most of the CAZyme genes were related to degradation RG I, and it was expected that by disassembling they could be obtained from galactose, arabinose, and xylose, and used as energy sources. All of these genes and clusters showed the character of the cell wall degradation, and also their functions are strongly expected.

Table 1 Activities of pectin degradation enzymes within CAZy families in *Sphingobium* sp. PAMC 28,499

CAZyme family	Enzyme functions
Glycoside hydrolases	
GH2	β-Galactosidase
GH28	Endorhamnogalacturonase
GH35	β-Galactosidase
GH43	α-$_L$-Arabinofuranosidase
	Endo-α-1,5-$_L$-arabinanase
GH51	α-$_L$-Arabinofuranosidase/endoglucanase
GH53	Endo-1,4-β-galactanase
GH105	Unsaturated rhamnogalacturonyl hydrolase
GH106	α-$_L$-Rhamnosidase
GH145	α-$_L$-Rhamnosidase
GH146	β-$_L$-Arabinofuranosidase
Polysaccharide lyases	
PL1	Pectate lyase
PL11	Rhamnogalacturonan endolyase
Carbohydrate esterases	
CE12	Rhamnogalacturonan acetylesterase

Comparative genome analysis of Sphingobium sp.

Only all complete genome for the strains of *Sphingobium* sp. were listed and checked comparative genome analysis to their CAZyme section, because to investigate the characteristics of this strain as a whole. Detailed information on comparative strains was summarized in Table 2, and the total number of genes for CAZyme was also summarized. Based on the results of ANI values (Fig. 1), it was ranged from

Table 2 Genomic information and comparative data on CAZymes for **20** strains, including *Sphingobium* sp. PAMC 28499

	Sphingobium sp.									*Sphingobium yanoikuyae*			*Sphingobium amiense*	*Sphingobium baderi*	*Sphingobium chlorophenolicum*	*Sphingobium cloacae*	*Sphingobium herbicidovorans*	*Sphingobium hydrophobicum*	*Sphingobium indicum*	*Sphingobium japonicum*
		PAMC 28499	EP60837	LF-16	MI1205	RAC03	SYK-6	TKS	YBL2	YG1	SHJ	S72	SJTF8	DSM 16289	DE-13	L-1	JCM 10874	MH	C1	B90A	UT26S
Genome information	Size (Mb)	4.8806	4.2617	4.57255	4.61937	4.36975	4.34813	6.22888	5.42742	5.56389	5.66938	5.53263	5.92959	4.77392	4.58342	4.57322	4.54757	4.2408	4.60264	3.94573	4.42486
	GC%	64.5	61.9	64.6	62.3	62.9	65.6	63.1	64.4	63.1	64.3	64.2	64.1	64.6	62.4	63.8	64.5	62.5	63.2	65.0	64.9
	Scaffolds	1	4	1	6	5	2	11	7	4	3	1	4	7	9	3	7	5	7	4	5
	Genes	4581	4117	4314	4638	4161	4049	6120	5152	5286	5435	5152	5669	4640	4569	4223	4567	4168	4462	3797	4232
	Proteins	4432	3929	4139	4425	4009	3932	5788	4956	4970	5195	5005	5393	4448	4370	4029	4258	3899	4173	3648	4087
CAZyme	AA	7	3	6	5	7	13	3	8	2	6	9	9	10	6	5	4	5	2	3	3
	CBM	2	1	2	2	2	2	2	4	2	3	4	3	4	0	2	0	2	2	2	2
	CE	28	7	26	6	4	13	6	15	5	28	29	27	18	8	6	9	5	8	9	6
	GH	74	18	79	26	23	23	30	57	24	83	107	83	58	20	23	18	24	18	26	20
	GT	49	36	46	37	31	43	37	50	24	49	50	45	38	35	42	34	36	29	37	39
	PL	5	1	5	2	2	0	1	1	1	6	7	5	0	1	1	1	2	0	1	1
	Total	165	66	164	78	69	94	79	135	58	175	206	172	128	70	79	66	74	59	78	71

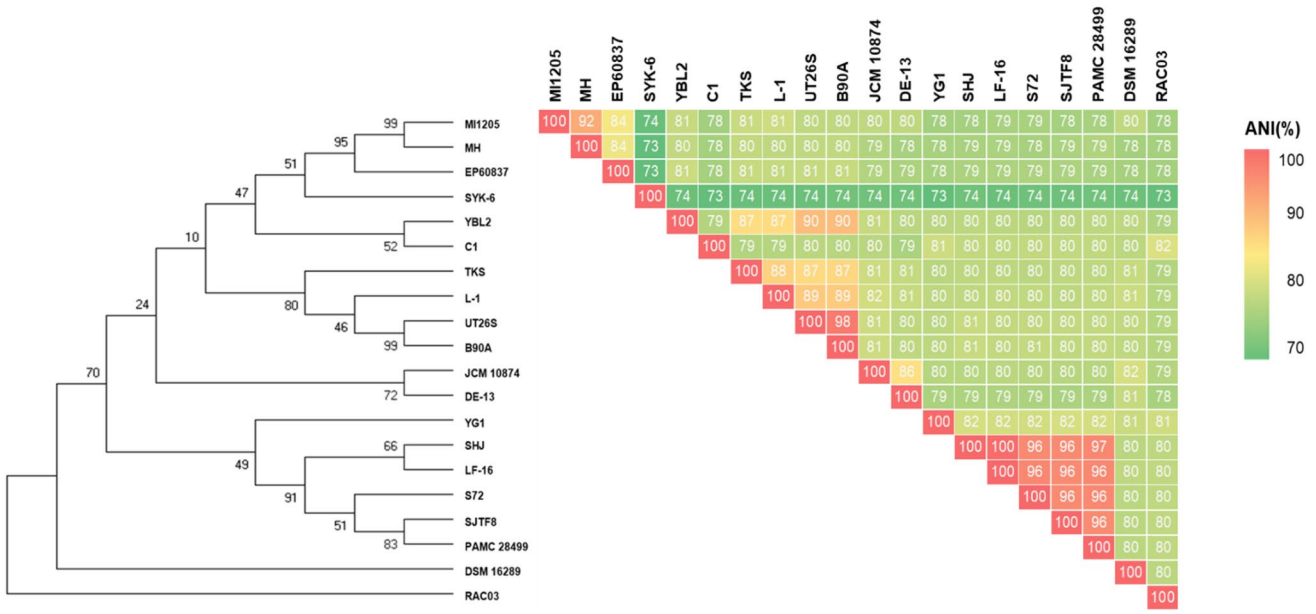

Fig. 1. 16 s rRNA phylogenetic tree and ANI value (%) for 20 strains, including *Sphingobium* sp. PAMC 28499. 16 s rRNA phylogenetic tree showing the relationships for the 19 strains of *Sphingobium* sp. It was prepared using MEGA X based on 16 s rRNA sequences with neighbor-joining. And heat map mean comparative results each strains of ANI value (%). It was calculated using the MUMmer alignment. MI1205, *Sphingobium* sp.; MH, *Sphingobium herbicidovorans*; EP60837, *Sphingobium* sp.; SYK-6, *Sphingobium* sp.; YBL2, *Sphingobium* sp.; C1, *Sphingobium hydrophobicum*; TKS, *Sphingobium* sp.; L-1, *Sphingobium chlorophenolicum*; UT26S, *Sphingobium japonicum*; B90A, *Sphingobium indicum*; JCM 10874, *Sphingobium cloacae*; DE-13, *Sphingobium baderi*; YG1, *Sphingobium* sp.; SHJ, *Sphingobium yanoikuyae*; LF-16, *Sphingobium* sp.; S72, *Sphingobium yanoikuyae*; SJTF8, *Sphingobium yanoikuyae*; PAMC 28499, *Sphingobium* sp. (our strain); DSM 16289, *Sphingobium amiense*; and RAC03, *Sphingobium* sp

97% (*Sphingobium yanoikuyae* SHJ) to 74% (*Sphingobium* sp. SYK-6). Overall comparison shows that all 20 complete genomes for ANI values belong to the same species and the same group was identified by checking 16S rRNA. However, the comparative bacteria *Sphingobium* sp. SYK-6 showed different patterns compared to other strains. Perhaps this strain is a new possibility of species. By using complete genomes of all strains, we have compared the CAZyme each other genes annotation using dbCAN2. At the comparative result Fig. 2, *Sphingobium yanoikuyae* strain S72 was the largest (206) and the *Sphingobium* sp. strain YG1 the least (58) about total CAZyme gene, and the number of result was all different. In conclusion, we found comparative about pectin degradation gene sections containing GH1, GH2, GH3, GH28, GH35, GH43, GH51, GH53, GH78, GH105, GH106, GH 145, GH146, CE12, PL1, and PL11. Interestingly, only the strains in the same group as the PAMC 28499 and 16S rRNA parts were identified by a large number of CAZyme genes and were able to confirm that they had similar pectin degradation genes. Overall, the *Sphingobium* strain, in a group similar to *Sphingobium yanoikuiae*, was shown to have a large number of genes for CAZyme and similarly has a genetic pattern. So our strains are thought to be similar to the *Sphingobium yanoikuiae* strain, and we can also identify the pattern of the pectin degradation gene. *Sphingobium yanoikuiae* strain also had the most gene in the family of GH43 and unlike other *Sphingobium* species, had GH145 and GH146. Its blast results were predicted with genes recently proven to function as α-$_L$-rhamnosidase and β-$_L$-arabinofuranosidase. Based on these results, we were able to predict that there were many genes that could degradation RG I, and that our strains also had the same pattern. This means that even in the *Sphingobium* species, the strains with similar genetic patterns, especially the *Sphingobium yanoikuiae*, can degradation the pectin, and not all *Sphingobium* species can be decomposed.

Pectin degradation

Many microorganisms can use these substances as carbon and energy sources by producing enormous amounts of enzymes in different environmental gaps, whose main components are cellulose, starch, lignine, xylan, and pectin. The pectic polysaccharide is a complex and high-molecular glycosidic macromolecules found in higher plants. They exist in the primary cell wall and are the main constituent of the medium lamela, a thin extracellular adhesive formed between the walls of the adjacent young cells (Kaur et al. 2004). Simply put, they are largely responsible for the structural integrity and cohesion of plant tissues. Pectinases is a group of enzymes that catalyze the

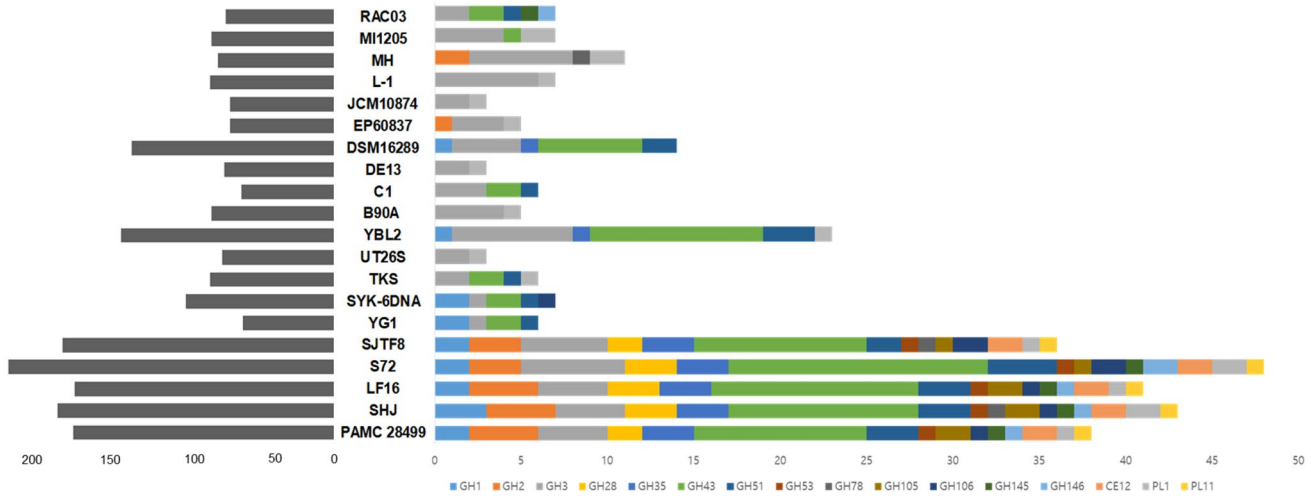

Fig. 2 Comparative total CAZyme genes and pectin degradation enzyme found in 20 genomes. RAC03, *Sphingobium* sp.; MI1205, *Sphingobium* sp.; MH, *Sphingobium herbicidovorans*; L-1, *Sphingobium chlorophenolicum*; JCM 10874, *Sphingobium cloacae*; EP60837, *Sphingobium* sp.; DSM 16289, *Sphingobium amiense*; DE-13, *Sphingobium baderi*; C1, *Sphingobium hydrophobicum*; B90A, *Sphingobium indicum*; YBL2, *Sphingobium* sp.; UT26S, *Sphingobium japonicum*; TKS, *Sphingobium* sp.; SYK-6, *Sphingobium* sp.; YG1, *Sphingobium* sp.; SJTF8, *Sphingobium yanoikuyae*; S72, *Sphingobium yanoikuyae*; LF-16, *Sphingobium* sp.; SHJ, *Sphingobium yanoikuyae*; and PAMC 28499, *Sphingobium* sp. (our strain)

degradation of pectic polysaccharide by means of depolymerization (hydrolysis enzyme and lyase) and well-known pectin degradation enzyme is a rhamnogalacturonan degradation enzyme (Alkorta et al. 1998). Figure 3 shows the degradation RG I pathway that strain PAMC 28499 have. GHs and PLs are two types of hydrolysis enzyme required for pectin backbone degradation. In general, the GHs perform a role in pectin backbone degradation include endo- and exo-polygalacturonases, which break the backbone of smooth regions, while the complex hairy regions are further degraded by endo- and exo-rhamnogalacturonases, xylogalacturonases, α-rhamnosidases, unsaturated glucuronyl hydrolases, and unsaturated rhamnogalacturonan hydrolases (van den Brink and de Vries 2011). Especially, endo- and exo-polygalacturonases are able to degrade α-1,4-glycosidic bonds of α-galacturonic acids, and rhamnogalacturonases degrade α-1,2-glycosidic bonds between $_{D}$-galacturonic acid and $_{L}$-rhamnose residues in the hairy region of the pectin backbone (Suykerbuyk et al. 1995). PLs split up the α-1,4-glycosidic linkages of polygalacturonate, generating oligogalacturonides with C4-C5 unsaturation at the nonreducing end. They have usually specific for the non-methylated polysaccharide or for pectin with a low degree of methyl esterification (Kashyap et al. 2001; Yang et al. 2019).

Discussion

Pectin is structurally the most complex component of the cell walls of the dicotyledonous plant, and pectic oligosaccharides are present in the primary side walls and middle lamella. This are most abundant in the cell walls of the soft part of the plant and surrounding growing cells. The structural complexity of pectin presents considerable difficulties to the human digestive system, and humans must rely on the coherent action of CAZymes produced by symbiotic gut bacteria to reduce pectin. Pectin-degrading enzyme is a have prerequisite for initiating plant cell wall degradation and reduce pectin.

In this study we presented that the complete genome sequence of *Sphingobium* sp. PAMC 28499 isolated from glaciers of Mt. Rwenzori show a genome size 4,880,615 bp and GC content of 64.5%. Through complete genome analysis, we suggested that it has pectin degradation character with diverse CAZyme genes. Overall, we confirmed that PAMC 28,499 have 74 GH family, 49 GT family, 28 CE family, 5 PL family, 2 CBM family, and 7 AA family in complete genome. In addition, by comparative genome of the *Sphingobium* strain, totally this strain species itself showed that the pectin degradation is a bacterium, and this reason is probably because cold environment. The change in cell wall remodeling in plant growth nourishes the stress resistance, and cell wall adjustment is an important determinant of plant adaptation. The bacteria that live there survive by deposition and metamorphosis in cell wall polysaccharides. A deep explanation of how to detect and transmit this signal is needed, but the function of the pectin-degradation gene will be identified, affecting plant cell walls such as pectin disassembly and secondary wall formation, and an important aspect of how bacteria can survive in cold areas. In addition, the strain is isolated strain from the disappearing the glaciers

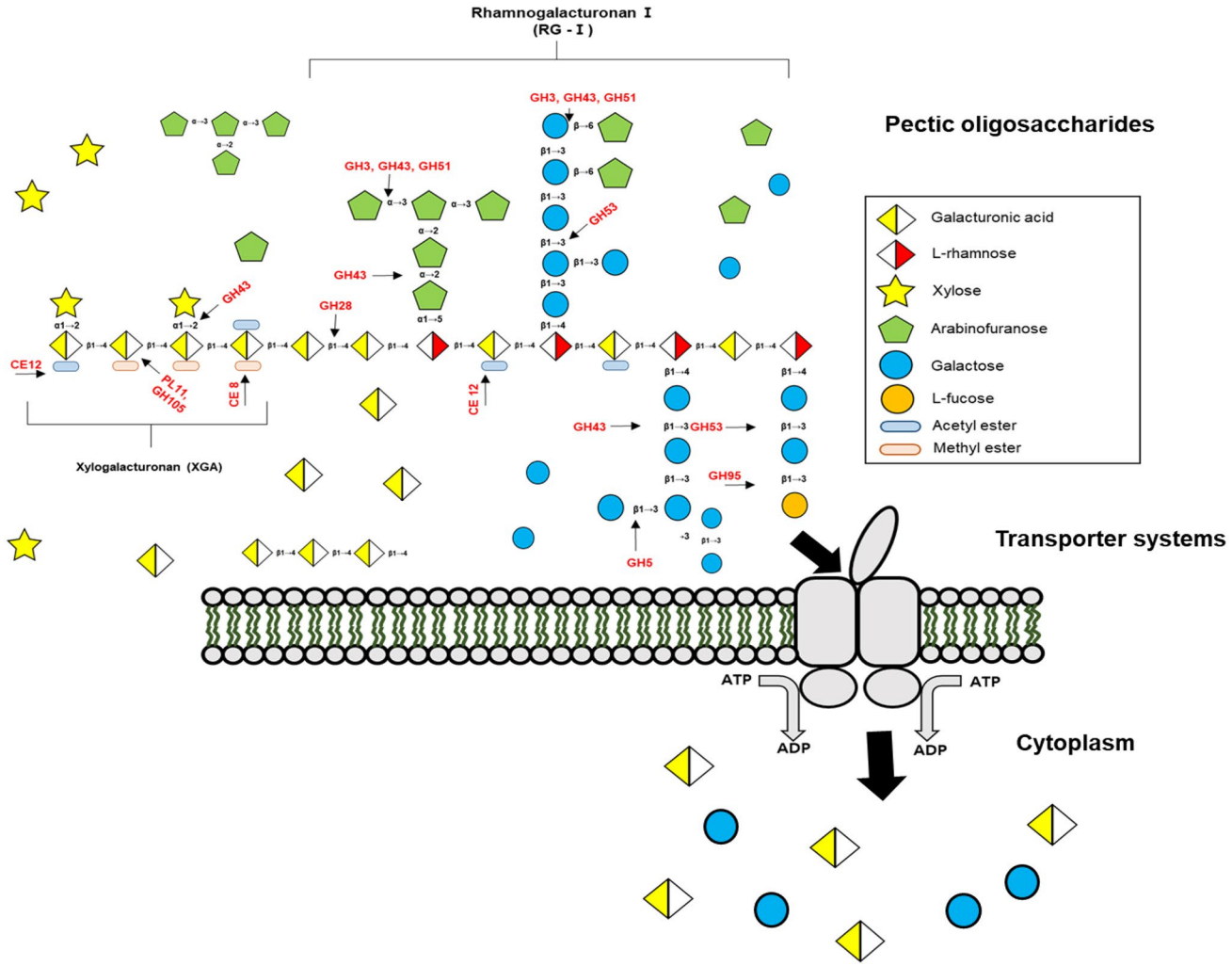

Fig. 3 Activities of pectin-degrading pathway on RG I. GH2, β-galactosidase and endorhamnogalacturonase; GH35, β-galactosidase; GH43, xylan 1,4-β-xylosidases, arabinan endo-1,5-α-$_L$-arabinosidases, and α-$_L$-arabinofuranosidases; GH51, α-$_L$-arabinofuranosidase; GH105, endo-1,4-β-galactanase and rhamnogalacturonyl hydrolase; GH106, α-$_L$-rhamnosidase; GH145, α-$_L$-rhamnosidase; PL1, β-$_L$-arabinofuranosidase and putative pectate lyase; and PL11, rhamnogalacturonan endolyase and rhamnogalacturonan acetylesterase

of Uganda, which is worth protecting and can provide an understanding of their survival environment through complete genome sequence.

Conclusion

In summary, its genome encodes a simple set of metabolic pathways relevant to pectin and its predicted degradation protein an unusual distribution of CAZymes with extracellular esterases and pectate lyases etc., about pectin degradation enzyme. The genome sequence provides valuable information for functional elucidations of novel enzymes for both biotechnological application and fundamental research purposes. Through the complete genome we can predict how this strain can store and produce energy in extreme environment. It can also provide bioengineered data by finding new genes that degradation the pectin.

Nucleotide sequence accession numbers

The complete genome sequence has been deposited at GenBank/EMBL/DDBJ under the Accession Number CP039248.

Acknowledgements This work was supported by the Korea Polar Research Institute (KOPRI Grant Number PM20030).

Compliance with ethical standard

Conflict of interest The authors declare that there are no conflicts of interest.

References

Adapa V, Ramya LN, Pulicherla KK, Rao KR (2014) Cold active pectinases: advancing the food industry to the next generation. Appl Biochem Biotechnol 172:2324–2337

Alkorta I, Gabirsu C, Lhama MJ, Serra JL (1998) Industrial applications of pectic enzymes: a review. Proc Biochem 33:21–28

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008) The RAST server: rapid annotations using subsystems technology. BMC Genom 9:75

Benen JAE, Vincken JP, van Alebeek GJWM (2002) Microbial pectinases. In: Seymour GB, Knox JP (eds) Pectins and their manipulation. Blackwell Publishing Ltd, Oxford, pp 174–221

Benoit I, Coutinho PM, Schols HA, Gerlach JP, Henrissat B, de Vries RP (2012) Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. BMC Genom 13:321

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60

Busk PK, Pilgaard B, Lezyk MJ, Meyer AS, Lange L (2017) Homology to peptide pattern for annotation of carbohydrate-active enzymes and prediction of function. BMC Bioinform 18:214

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res 344:1879–1900

Chaudhary DK, Jeong SW, Kim JS (2017) *Sphingobium naphthae* sp. nov., with the ability to degrade aliphatic hydrocarbons, isolated from oil-contaminated soil. Int J Syst Evol Microbiol 67:2986–2993

Cheng KJ, Dinsdale D, Stewart CS (1979) Maceration of clover and grass leaves by *Lachnospira multiparus*. Appl Environ Microbiol 38:723–729

Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569

Chung D, Pattathil S, Biswal AK, Hahn MG, Mohnen D, Westpheling J (2014) Deletion of a gene cluster encoding pectin degrading enzymes in *Caldicellulosiruptor bescii* reveals an important role for pectin in plant biomass recalcitrance. Biotechnol Biofuels 7:147

Dinu D, Nechifor MT, Stoian G, Costache M, Dinischiotu A (2007) Enzymes with new biochemical properties in the pectinolytic complex produced by *Aspergillus niger* MIUG 16. J Biotechnol 131:128–137

Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195

Gadre RV, Van Driessche G, Van Beeumen J, Bhat MK (2003) Purification, characterisation and mode of action of an endo-polygalacturonase from the psychrophilic fungus *Mucor flavus*. Enzyme Microb Technol 32:321–330

Grant JR, Arantes AS, Stothard P (2012) Comparing thousands of circular genomes using the CGView Comparison Tool. BMC Genom 13:202

Hope GS, Peterson JA, Radok U, Allison I (1976) The Equatorial glaciers of New Guinea (Results of the 1971–1973 Australian universities' expeditions to Irian Jaya: survey, glaciology, meteorology, biology and palaeoenvironments). A. A. Balkema, Rotterdam, pp 81–92

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M, Jensen LJ, von Mering C, Bork P (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44:D286–D293

Kashyap DR, Vohra PK, Chopra S, Tewari R (2001) Applications of pectinases in the commercial sector: a review. Bioresour Technol 77:215–227

Kaur G, Kumar S, Satyanarayana T (2004) Production, characterization and application of a thermostable polygalacturonase of a thermophilic mould Sporotrichum thermophile Apinis. Bioresour Technol 94:239–243

Kumari H, Gupta SK, Jindal S, Katoch P, Lal R (2009) *Sphingobium lactosutens* sp. nov., isolated from a hexachlorocyclohexane dump site and *Sphingobium abikonense* sp. nov., isolated from oil-contaminated soil. Int J Syst Evol Microbiol 59:2291–2296

Maeda AH, Kunihiro M, Ozeki Y, Nogi Y, Kanaly RA (2015) *Sphingobium barthaii* sp. nov., a high molecular weight polycyclic aromatic hydrocarbon-degrading bacterium isolated from cattle pasture soil. Int J Syst Evol Microbiol 65:2919–2924

Maughan R (2009) Carbohydrate metabolism. Surg (Oxf) 27:6–10

Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277

Mölg T, Rott H, Kaser G, Fischer A, Cullen NJ (2006) Comment on ''Recent glacial recession in the Rwenzori Mountains of East Africa due to rising air temperature'' by Richard G. Taylor, Lucinda Mileham, Callist Tindimugaya, Abushen Majugu, Andrew Muwanga, and Bob Nakileza. Geophys Res Lett 33:L20404

Peterson SB, Dunn AK, Klimowicz AK, Handelsman J (2006) Peptidoglycan from *Bacillus cereus* mediates commensalism with rhizosphere bacteria from the Cytophaga-Flavobacterium group. Appl Environ Microbiol 72:5421–5427

Poveda G, Gil-Durán C, Vaca I, Levicán G, Chávez R (2018) Cold-active pectinolytic activity produced by filamentous fungi associated with Antarctic marine sponges. Biol Res 51:28

Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106:19126–19131

Saito K, Takakuwa N, Oda Y (2004) Purification of the extracellular pectinolytic enzyme from the fungus *Rhizopus oryzae* NBRC 4707. Microbiol Res 159:83–86

Schols HA, Voragen AGJ (1996) Complex pectins: structure elucidation using enzymes. Prog Biotechnol 14:3–19

Silva IR, Jers C, Meyer AS, Mikkelsen JD (2016) Rhamnogalacturonan I modifying enzymes: an update. New Biotechnol 33:41–54

Suykerbuyk ME, Schaap PJ, Stam H, Musters W, Visser J (1995) Cloning, sequence and expression of the gene coding for rhamnogalacturonase of *Aspergillus aculeatus*; a novel pectinolytic enzyme. Appl Microbiol Biotechnol 43:861–870

Takasawa T, Sagisaka K, Yagi K, Uchiyama K, Aoki A, Takaoka K, Yamamato K (1997) Polygalacturonase isolated from the culture of the psychrophilic fungus *Sclerotinia borealis*. Can J Microbiol 43:417–424

Takeuchi M, Hamana K, Hiraishi A (2001) Proposal of the genus *Sphingomonas sensu* stricto and three new genera, Sphingobium, Novosphingobium and Sphingopyxis, on the basis of phylogenetic and chemotaxonomic analyses. Int J Syst Evol Microbiol 51:1405–1417

Taylor RG, Mileham L, Tindimugaya C, Majugu A, Muwanga A, Nakileza B (2006) Recent glacial recession in the Rwenzori Mountains of East Africa due to rising air temperature. Geophys Res Lett 33:L10402

Uetake J, Tanaka S, Hara K, Tanabe Y, Samyn D, Motoyama H, Imura S, Kohshima S (2015) Novel biogenic aggregation of moss gemmae on a disappearing African glacier. PLoS One 10:e0123375

Ushiba Y, Takahara Y, Ohta H (2003) *Sphingobium amiense* sp. nov., a novel nonylphenol-degrading bacterium isolated from a river sediment. Int J Syst Evol Microbiol 53:2045–2048

van den Brink J, de Vries RP (2011) Fungal enzyme sets for plant polysaccharide degradation. Appl Microbiol Biotechnol 91:1477–1492

Voragen AGJ, Beldman G, Schols H (2001) Chemistry and enzymology of pectins. In: McCleary BV, Prosky L (eds) Advanced dietary fibre technology. Blackwell Publishing Ltd., Oxford, pp 379–397

Willats WGT, Knox P, Mikkelsen JD (2006) Pectin: new insights into an old polymer are starting to gel. Trends Food Sci Technol 17:97–104

Wong D (2008) Enzymatic deconstruction of backbone structures of the ramified regions in pectins. Protein J 27:30–42

Yabuuchi E, Yano I, Oyaizu H, Hashimoto Y, Ezaki T, Yamamoto H (1990) Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulate* comb. nov., and two genospecies of the genus Sphingomonas. Microbiol Immunol 34:99–119

Yang YJ, Lin Wei, Singh RP et al (2019) Genomic, Transcriptomic and Enzymatic Insight into Lignocellulolytic System of a Plant Pathogen Dickeya sp. WS52 to Digest Sweet Pepper and Tomato Stalk. Biomolecules 9:75. https://doi.org/10.3390/biom9120753

Yang Y, Qin S, Zhao F, Chi X, Zhang X (2007) Comparison of envelope-related genes in unicellular and filamentous cyanobacteria. Comp Funct Genom 2007:25751

Young CC, Ho MJ, Arun AB, Chen WM, Lai WA, Shen FT, Rekha PD, Yassin AF (2007) *Sphingobium olei* sp. nov., isolated from oil-contaminated soil. Int J Syst Evol Microbiol 57:2613–2617

Affiliations

So-Ra Han[1] · Sung-Min Jang[2] · Young Min Chi[2] · Byeollee Kim[1] · Sang-Hee Jung[3] · Yung Mi Lee[4] · Jun Uetake[5] · Jun Hyuck Lee[6,7] · Hyun Park[2] · Tae-Jin Oh[1,8,9]

[1] Department of Life Science and Biochemical Engineering, Graduate School, SunMoon University, Asan, Korea

[2] Division of Biotechnology, Korea University, Seoul 02841, Korea

[3] Department of Dental Hygiene, Gangneung Yeongdong University, Gangneung, Korea

[4] Department of Polar Sciences, Korea Polar Research Institute, Incheon, Korea

[5] Department of Atmospheric Science, Colorado State University, 200 West Lake Street, 1371 Campus Delivery, Fort Collins, CO 80523-1371, USA

[6] Unit of Research for Practical Application, Korea Polar Research Institute, Incheon, Korea

[7] Department of Polar Sciences, Korea University of Science and Technology, Incheon, Korea

[8] Genome-Based BioIT Convergence Institute, Asan, Korea

[9] Department of Pharmaceutical Engineering and Biotechnology, SunMoon University, Asan 31460, Korea