Supplementary Information for

Culturing the ubiquitous freshwater actinobacterial acI lineage by supplying a biochemical 'helper' catalase

Suhyun Kim, Ilnam Kang, Ji-Hui Seo, and Jang-Cheon Cho

Department of Biological Sciences, Inha University, Incheon 22212, Republic of Korea

Corresponding author: Jang-Cheon Cho

Email: chojc@inha.ac.kr

This file includes:

- Supplementary information text
- Figs. S1 to S9
- Tables S1 to S5
- References for SI reference citations

Supplementary Information Text

Description of two proposed '*Candidatus*' species. The average nucleotide identity (ANI) value calculated from genome sequences between strain IMCC25003 and '*Ca*. Planktophila sulfonica' MMS-IA-56 was 84% and between strains MCC26103 and '*Ca*. Planktophila lacus' MMS-21-148 was 78%, which were both below the 95–96% cut-off value for bacterial species demarcation (1, 2). Analysis of genomic DNA-DNA relatedness and differential phenotypic characteristics indicated that strains IMCC25003 and IMCC26103 each represent novel species of the genus '*Candidatus* Planktophila'. However, because the two strains did not grow on a defined medium or a synthetic medium but replicated only in complex natural lake water media, limiting the deposition of the acI strains in culture collections, we propose the provisional names '*Candidatus* Planktophila rubra' for strain IMCC25003 and '*Candidatus* Planktophila aquatilis' for strain IMCC26103.

'*Candidatus* Planktophila rubra' (ru'bra. L. fem. adj. *rubra* reddish, pertaining to the reddish color of cells)

Represented by a cultured bacterial strain, IMCC25003. Gram-positive, aerobic, red-pigmented, non-motile, and chemoheterotrophic. Cells are curved rods with biovolume of 0.041 μm^3, 0.46–1.23 μm (average 0.68 μm) long and 0.25–0.37 μm (average 0.30 μm) wide. Grows in FAMV+CM+AA supplemented with >0.5 U mL^{-1} catalase but does not grow in any liquid medium devoid of catalase and on any solid agar medium. Growth occurs at 10–30°C (optimum, 25°C). No single carbon sources enhance cellular growth. Requires sulfur-containing amino acids (methionine and cysteine) but prefers methionine. The major fatty acids (>10%) are summed feature 3 ($C_{16:1}$ $\omega 7c$ and/or $C_{16:1}$ $\omega 7c$, 45.8%), $C_{16:0}$ (23.1%), and $C_{14:0}$ (18.2%). Strain IMCC25003 has a genome size of 1.354 Mbp with DNA G+C content of 49.1%. The complete genome sequence of strain IMCC25003 is available in GenBank (CP029557). Phylogenetically belongs to the acI-A1 tribe.

The representative strain IMCC25003 was isolated from a freshwater lake, Lake Soyang, Republic of Korea, using a dilution-to-extinction culturing.

'*Candidatus* Planktophila aquatilis' (a.qua.ti′lis. L. fem. adj. *aquatilis* living, growing, or found, in or near water, aquatic).

Represented by a cultured bacterial strain, IMCC26103. Gram-positive, aerobic, red-pigmented, non-motile, and chemoheterotrophic. Cells are curved rods with biovolume of 0.061 μm^3, 0.49–1.23 μm (average 0.88 μm) long and 0.22–0.39 μm (average 0.31 μm) wide. Grows in FAMV+CM+AA supplemented with >0.5 U mL^{-1} catalase but does not grow in any liquid medium devoid of catalase and on any solid agar medium. Growth occurs at 10–30°C (optimum, 25°C). D-ribose and D-glucose enhance the cellular growth. Requires sulfur-containing amino acids (methionine and cysteine) but prefers cysteine. The major fatty acids (>10%) are $C_{16:0}$ (28.5%), $C_{18:1}$ $\omega 9c$ (25.8%), summed feature 3 ($C_{16:1}$ $\omega 7c$ and/or $C_{16:1}$ $\omega 7c$, 12.3%), and $C_{18:0}$ (10.5%). Strain IMCC26103 has a genome size of 1.457 Mbp with DNA G+C content of 47.0%. The complete genome sequence of strain IMCC26103 is available in GenBank (CP029558). Phylogenetically belongs to the acI-A4 tribe.

The representative strain IMCC26103 was isolated from a freshwater lake, Lake Soyang, Republic of Korea, using a dilution-to-extinction culturing.

Supplementary Materials and Methods

Measurement of bacterial cell densities. Bacterial cell densities in all growth experiments were determined by flow cytometry (Guava easyCyte Plus, Millipore) as described previously (3, 4). After 200 μL of bacterial cultures were stained with SYBR Green I (5× final concentration, Invitrogen) for 1 h, each stained sample was run for 10 s or until total cell counts reached 5000. To accurately measure the cell counts in the samples with high cell density, the stained sample was diluted to contain less than 200 cells μL^{-1}.

Measurement of *katG* expression by qPCR. Strain IMCC25003 was grown in triplicate in 4 L of FAMV+CM+AA supplemented with 1 U mL^{-1} of catalase until the early stationary phase. Each 4-L bacterial culture was harvested by centrifugation at 20,000 ×*g* for 120 min. To examine the expression of IMCC25003 *katG* and compare the relative level of gene expression according to H_2O_2 concentrations, harvested cells were treated with different concentrations of H_2O_2 (10, 50, and 100 μM) for 30 min and

untreated cells were used as a control. RNA was extracted using TRIzol (Sigma-Aldrich). Reverse transcription was performed with 1 μg of RNA using qPCRBIO cDNA Synthesis Kit (PCRBIO Systems) and real-time qPCR was conducted using qPCRBIO SyGreen Blue Mix Lo-ROX (PCRBIO Systems) in a real-time thermal cycler (Rotor-Gene 3000, Corbett Research). The *katG* primer set (forward, 5′-CATGGCGATGAATGATGAAG-3′; reverse, 5′-GCTGTTCTTCCAGCCAAGTC-3′) for targeting *katG* of IMCC25003 was used to evaluate gene expression and the GAPDH gene of IMCC25003 was employed as a housekeeping gene with the GAPDH primer set (forward, 5′-GTTCAGCGACAGACCTCACA-3′; reverse, 5′-TGGTGAGCTGTGAATCGAAG-3′).

Expression, purification, and characterization of KatG from IMCC25003. The gene encoding catalase-peroxidase (KatG) of IMCC25003 was amplified by PCR using the following primers: forward, 5′-CATATGATGACTCAAGAATCAACTCC-3′; reverse, 5′-CTCGAGTTACTTCTTCTTTGAC-3′. The PCR product was inserted into the pET-15b vector (Novagen) and expressed in *Escherichia coli* BL21 (DE3) using 1 mM isopropyl-*β*-D-thiogalatoside (Sigma-Aldrich). The expressed KatG-His recombinant protein was purified using a gravity-flow Ni^{2+}-nitrilotriacetic acid affinity column (Novagen). The high concentration of salts used for elution in the affinity column was removed using a PD-10 desalting column (GE Healthcare). To increase protein purity, the eluted fractions were applied to a Superose-12 FPLC column (10 × 300 mm, GE Healthcare) equilibrated with 50 mM potassium phosphate buffer (pH 7.0). The purified recombinant proteins were concentrated to approximately 10 mg mL^{-1} in Centricon tubes (MWCO 10,000 Da; Millipore) and stored at 4°C. All purification steps were carried out at 4°C or on ice.

The native molecular weight of IMCC25003 KatG was determined by size exclusion chromatography on a Superose-12 FPLC column (10 × 300 mm, GE Healthcare) equilibrated with 50 mM potassium phosphate buffer (pH 7.0). The molecular weights of the subunits were determined by discontinuous sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) according to the standard Laemmli method. Catalytic activity of recombinant IMCC25003 KatG was determined using a spectrophotometric assay by measuring the decomposition of H_2O_2 at 240 nm. Catalase-specific activity was quantified by allowing varying amounts of enzyme (0–1.0 μg for bovine catalase; 0–5.0 μg for IMCC25003 KatG) to react with 5 mM H_2O_2 in 50 mM potassium phosphate buffer (pH 7.0) at 25°C. The

absorption coefficient at 240 nm, pH 7.0, and 25°C for H_2O_2 was determined to be 49.8 M^{-1} cm^{-1}. Kinetic parameters were determined in triplicate from initial linear reaction rates of H_2O_2 ranging from 1 to 10 mM. The apparent Km (mM) and kcat (s^{-1}) values at these substrate concentrations were determined from a Lineweaver-Burk plots. Catalase from bovine liver was used as a positive control. In SDS-PAGE, staining for catalase activity was performed with the ferricyanide negative staining method using 2% (w/v) ferric chloride and 2% (w/v) potassium ferricyanide solution, and peroxidase activity was detected by double-staining with 3,3′,5,5′-tetramethylbenzidine (5).

Phylogenetic analyses based on 16S rRNA gene, whole genome, and KatG protein. The 16S rRNA gene sequences of the acI genomes (6-8) were downloaded from the IMG database and GenBank, aligned using SINA online aligner, and imported into the ARB-SILVA database (SSURef NR 99, release 123). Multiple alignments of the imported sequences and other reference sequences of the acI lineage were exported with the 'ssuref:bacteria' filter and used to construct a maximum-likelihood tree in RAxML 8.2.7 with the GTRGAMMA model. Phylogenetic assignment of the sequences was performed as described by Newton *et al*. (9) and the recently proposed names for the two *Candidatus* genera (8).

To build phylogenomic trees, protein sequences predicted in 4 completed acI genomes (7) and single-cell genomes (6) and 16 recently published acI genomes (8) were downloaded from the IMG database and NCBI RefSeq database. Downloaded protein sequences were processed using CheckM (10), which produces concatenated alignment of 43 conserved proteins. This concatenated alignment was used to build a maximum-likelihood tree using RAxML 8.2.7 with the PROTGAMMAAUTO model.

KatG proteins in the acI genomes (listed in *SI Appendix*, Table S4) were searched by BLASTp using the KatG sequence of IMCC25003 as a query and acI protein sequences downloaded above as a search database, which revealed a total of 20 acI KatG proteins. Phylogenetic trees of KatG proteins were constructed to identify the phylogenetic positions of the KatGs found in the acI genomes. Sequences collected for tree construction included the following: 20 acI KatG proteins, 16 KatG proteins showing high similarities to acI KatG proteins in BLASTp against the nr database of GenBank, 19 actinobacterial KatG proteins searched from the genomes representing diverse taxonomic groups of the phylum *Actinobacteria* (7), and >300 KatG proteins downloaded from PeroxiBase (11)

(http://peroxibase.toulouse.inra.fr; Category ‘Catalase peroxidase’ of ‘Class I peroxidase superfamily’ under ‘Non Animal peroxidase’). After sequence collection, several rounds of alignment and tree building were performed, and some sequences were excluded because of their short length, poor alignment, or unstable positioning. Finally, 303 KatG proteins were selected, aligned with Muscle (12) implemented in the MEGA 6 program, and used to construct a maximum likelihood tree using RAxML (version 8.2.7), with automatic model selection based on aic criterion (-m PROTGAMMAAUTO --auto-prot=aic). The selected model was LG likelihood with empirical base frequencies. Grouping of the KatG proteins was performed as described by Zamocky *et al*. (13).

Supplementary Figures

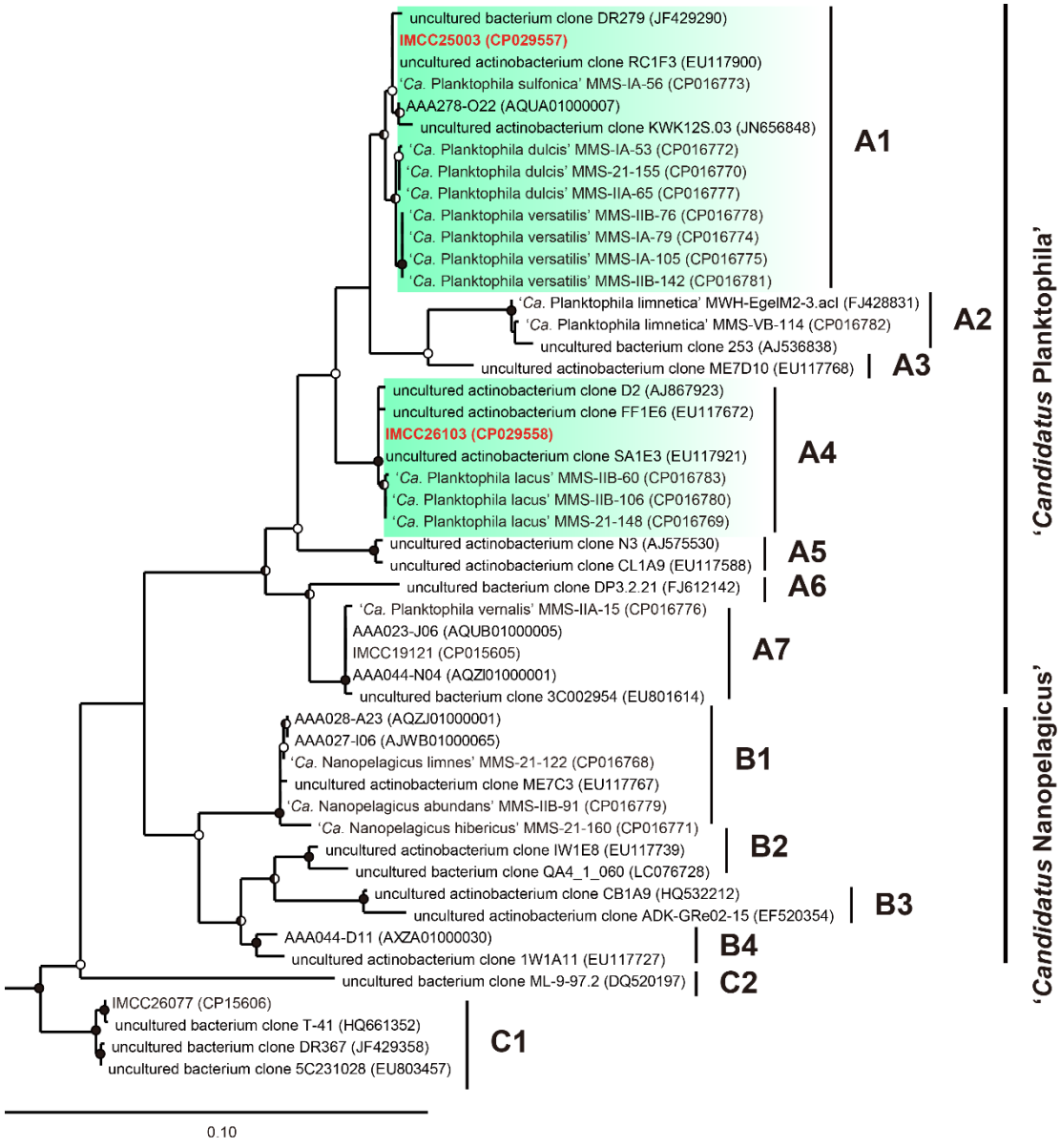


Fig. S1. Maximum-likelihood tree based on 16S rRNA gene sequences showing the phylogenetic position of strains IMCC25003 and IMCC26103. The two strains isolated in this study are marked in red. *Streptomyces sannanensis* (AB184579) and *Streptomyces griseus* (AY999909) were used as outgroup. Bootstrap supporting values (from 600 replicates) are shown at the nodes as filled circles (≥90%), half-filled circles (≥70%), and empty circles (≥50%). Bar, 0.10 substitutions per nucleotide position.

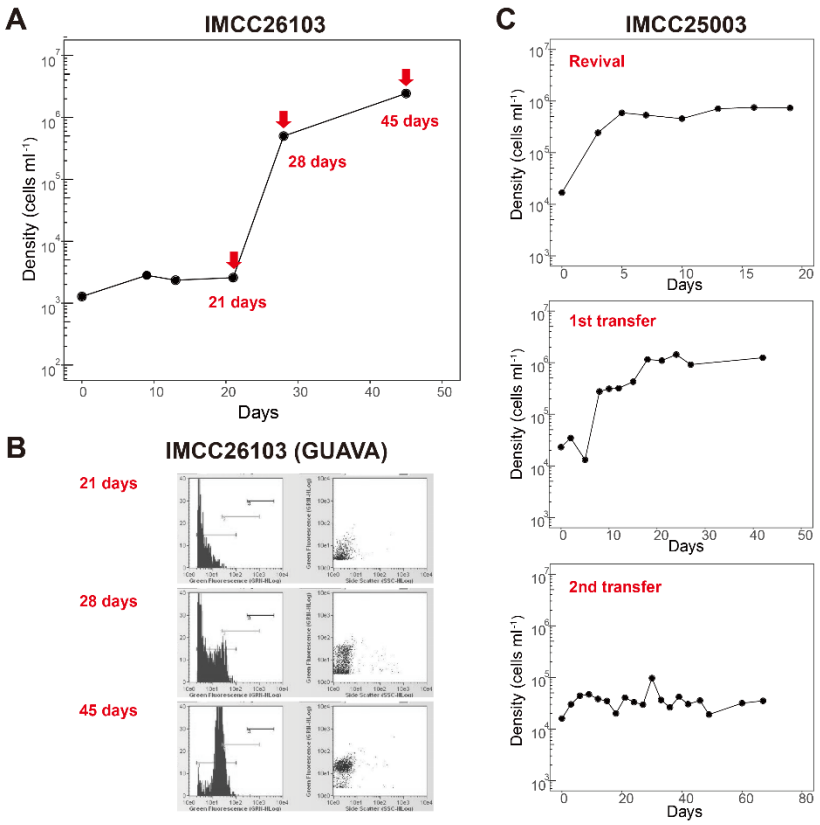
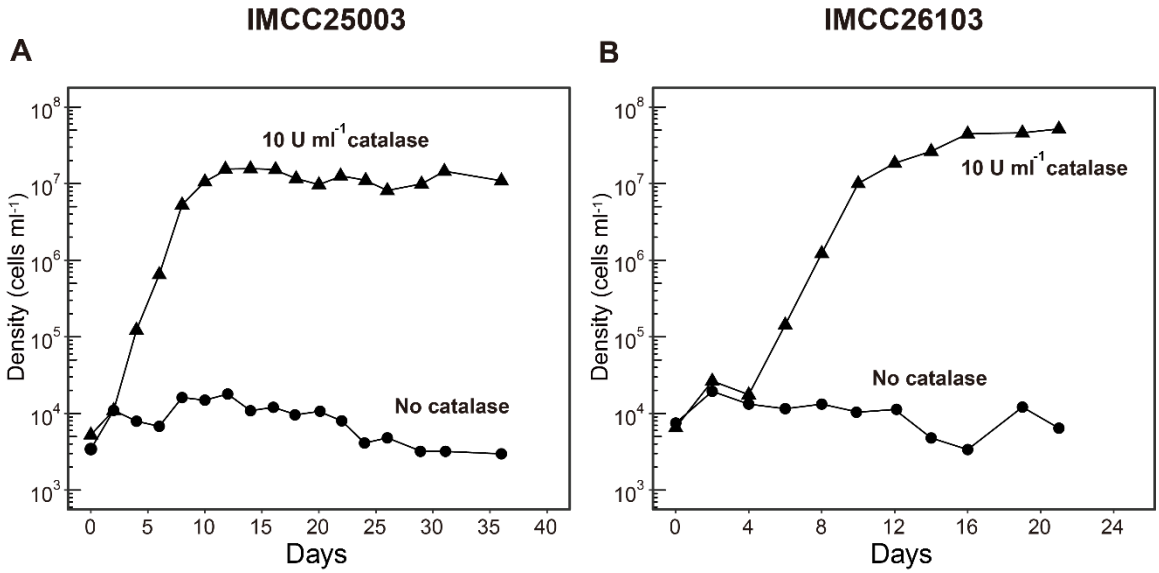


Fig. S2. Revival and transfer cultures of strains IMCC25003 and IMCC26103. (a) Growth curve of a revival culture of strain IMCC26103. (b) Flow cytometry plots of strain IMCC26103 obtained at the time points indicated in (a). Left, histograms showing the distribution of cell counts (y-axis) according to the green fluorescence (x-axis); Right, dot plots showing the distribution of cells according to side scatter (x-axis) and green fluorescence (y-axis). (c) Revival and two subsequent transfer cultures of strain IMCC25003.



156

157 **Fig. S3.** Growth curves of strains IMCC25003 (a) and IMCC26103 (b) obtained from the revival

158 experiment of frozen glycerol stocks using culture medium (FAMV+CM+AA) supplemented with catalase.

159 The triangle symbol represents growth in the medium amended with 10 U mL^{-1} catalase, while the circle

160 symbol represents growth in the medium without catalase.

161

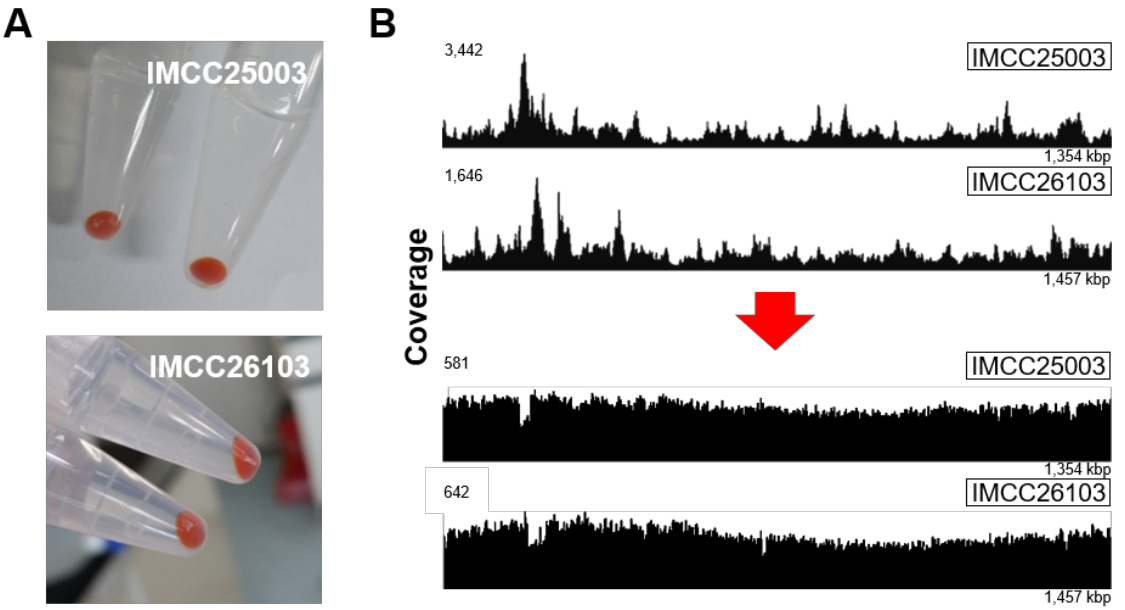


Fig. S4. Genome sequencing using genomic DNA extracted from cultured and harvested cell pellets of the acI strains. (a) Cell pellets obtained by centrifugation from 4-L cultures of IMCC25003 (upper) and IMCC26103 (lower). Genomic DNA extracted from these cell pellets were used for genome sequencing. (b) Coverage variation across the complete genome sequences of the acI strains. The two coverage plots above red arrow were obtained from our previous study using whole genome amplification (WGA) (7). The two coverage plots below the arrow were obtained from this study using large-scale cultures without WGA. Coverage variation was calculated using a 25-bp window based on read mapping. Bar heights were normalized in each plot and the maximum coverages are indicated at the upper left corner of each plot.

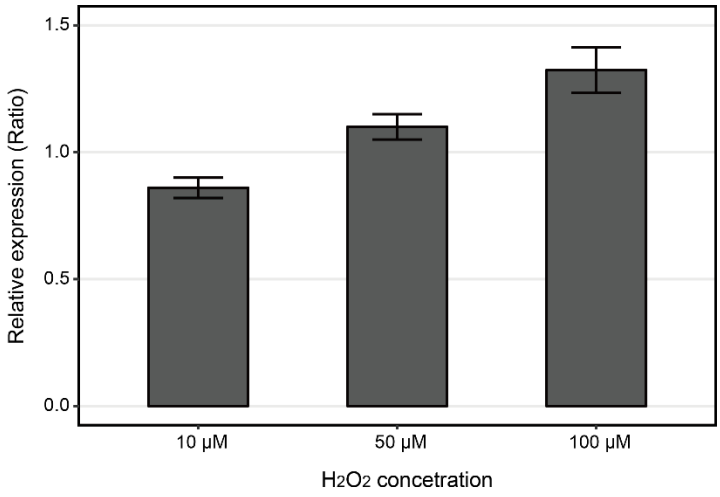


Fig. S5. Increase in IMCC25003 *katG* expression with increasing concentration of H_2O_2. Cells of IMCC25003 were treated with 3 different H_2O_2 concentrations (10, 50, and 100 μM) for 30 min and total RNA was used for the analysis of *katG* expression by qPCR. Expression level of *katG* in H_2O_2-treated cultures was compared with that in the control cultures (no H_2O_2 treatment). FAMV+CM+AA was used as the culture medium. Error bars indicate standard deviations (n = 3).

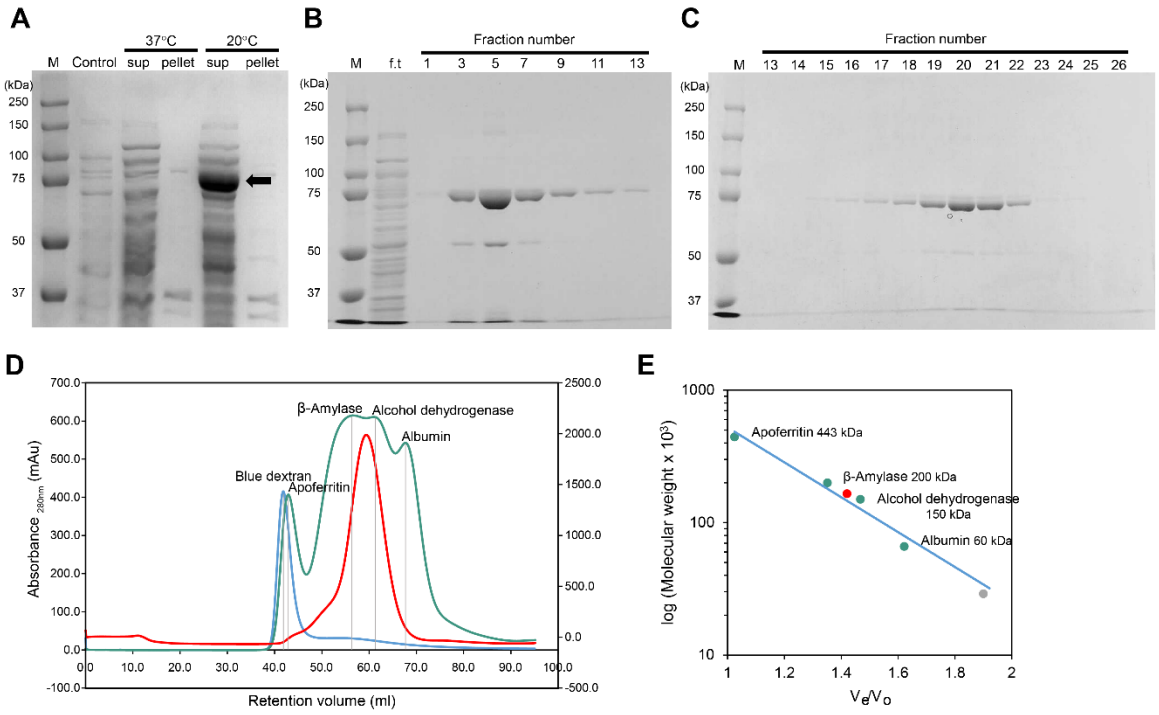


Fig. S6. Expression, purification, and determination of native molecular weight of recombinant IMCC25003 KatG. (a) Expression of IMCC25003 KatG in *E. coli* analyzed by SDS-PAGE. The bold arrow indicates a band of KatG, which is approximately 82.6 kDa. M, molecular weight size marker; control, before induction of expression; sup, supernatant; pellet, cell debris and membrane. Purified IMCC25003 KatG bound to a Ni^{2+}-nitrilotriacetic acid affinity column (b) and the purified protein through a size exclusion superpose-12 column (c), confirmed by SDS-PAGE. M, molecular weight size marker; f.t., unbound flow through fraction. Chromatograms (d) of protein-molecular-weight size markers and IMCC25003 KatG, and the molecular-weight calibration curve (e) obtained from protein-molecular-weight size markers and IMCC25003 KatG. The chromatogram colored in red and the red dot on the calibration curve represent IMCC25003 KatG.

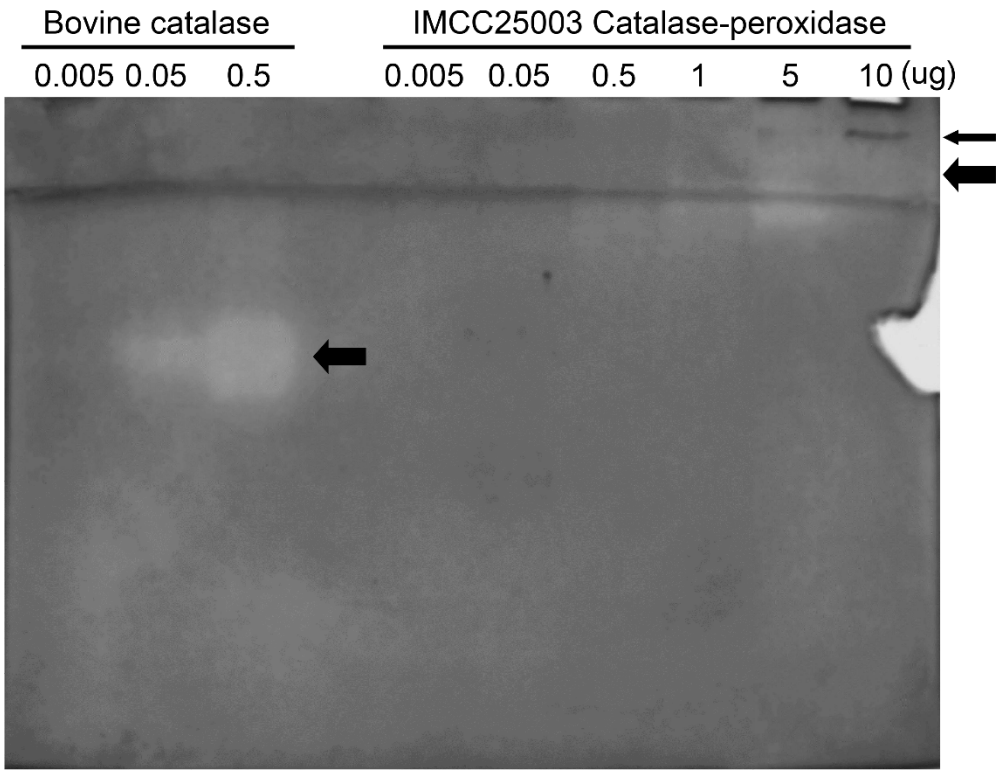
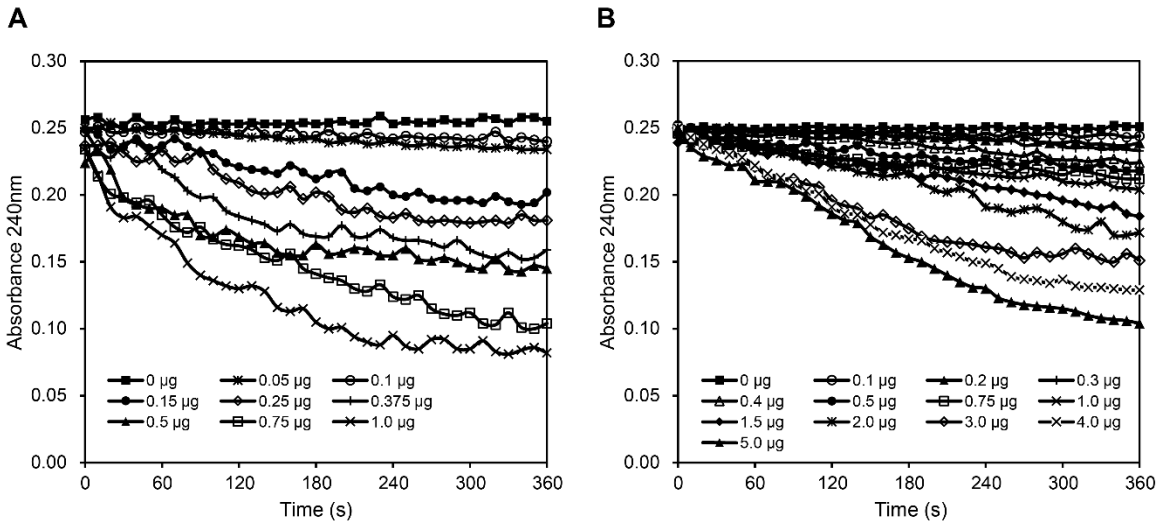


Fig. S7. Catalase and peroxidase activities of purified IMCC25003 KatG and bovine catalase (KatE). Bovine catalase [0.005 (0.01 U), 0.05 (0. 1 U), and 0.5 (1 U) μg] and IMCC25003 KatG (0.005, 0.05, 0.5, 1, 5, and 10 μg) were separated by 8% non-denaturing PAGE. The bold arrows indicate negatively stained catalase activity and the narrow arrow indicates peroxidase activity stained by 3,3′,5,5′-tetramethylbenzidine.



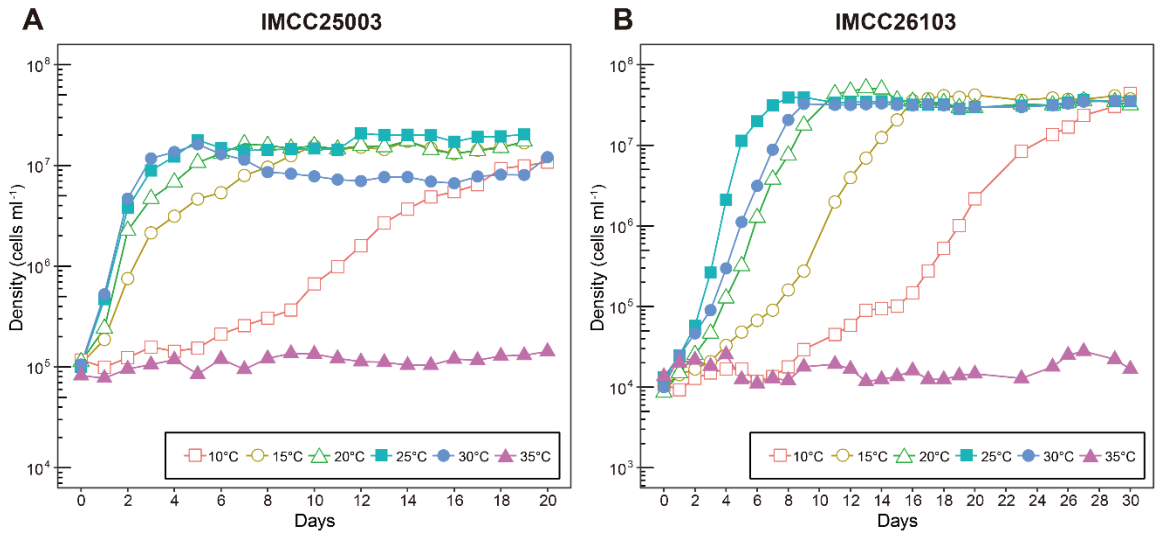
198

199 **Fig. S8.** Kinetic curves of H_2O_2 decomposition by IMCC25003 KatG and bovine catalase. The curves of

200 absorbance at 240 nm over time were generated using varying quantities of (a) bovine catalase (0–1.0 μg)

201 and (b) IMCC25003 KatG (0–5.0 μg).

202



203

204 **Fig. S9.** Growth curves of strain IMCC25003 (a) and strain IMCC26103 (b) at different temperatures.

205

Supplementary Tables

Table S1. Media used in this study and their composition.

Components of media		
Components (abbreviation)	Compound(s)	Final concentration
Ammonium (N)	NH_4Cl	10 μM
Phosphate (P)	KH_2PO_4	10 μM
Trace metals (TM)	$FeCl_3 \cdot 6H_2O$	117 nM
	$MnCl_2 \cdot 4H_2O$	9 nM
	$ZnSO_4 \cdot 7H_2O$	800 pM
	$CoCl_2 \cdot 6H_2O$	500 pM
	$Na_2MoO_4 \cdot 2H_2O$	300 pM
	Na_2SeO_3	1 nM
	$NiCl_2 \cdot 6H_2O$	1 nM
Vitamin mixture (V)	Thiamine·HCl	59 nM
	Niacin	81 nM
	Ca-Pantothenate	84 nM
	Pyridoxine	59 nM
	Biotin	409 pM
	Folic acid	453 pM
	Vitamin B12	70 pM
	Myo-inositol	555 nM
	p-Aminobenzoic Acid	7 nM
Carbon mixture (CM)	Pyruvate	50 μM
	D-Glucose	5 μM
	N-Acetyl-D-glucosamine	5 μM
	D-Ribose	5 μM
	Methyl alcohol	5 μM
20 proteinogenic amino acid mixture (AA)	Each amino acid	100 nM, each
Media definition		
Media	Definition	
FAM	0.2 μm-filtered and autoclaved freshwater medium supplemented with N, P, and TM	
FAMV	FAM supplemented with V	
FAMV+CM	FAMV supplemented with CM	
FAMV+AA	FAMV supplemented with AA	
FAMV+CM+AA	FAMV supplemented with CM and AA	

Table S2. Trials to establish pure culture of strain IMCC25003.

Trial	Media composition	Additional substrate	Reference
1st attempt	FAMV		
	FAMV+CM		
	FAMV+AA	0.5×, 1×, 5×, and 10× of CM	
	AFM[a]+V+CM+AA		
	FM[b]+V+CM+AA		
2nd attempt	FAMV+CM+AA	20 μM acetate	[(14)]
		20 μM oxaloacetate	[(6)]
		20 μM putrescine	[(6, 15)]
		20 μM glycerol	[(6)]
		20 μM xylose	[(15, 16)]
		1 mg L^{-1} proteose peptone No. 3	
		1 mg L^{-1} yeast extract	
3rd attempt	FAMV+CM+AA	1:20 diluted spent medium[c]	[(15, 17)], This study
4th attempt	FAMV+CM+AA	10 U mL^{-1} catalase	

[a]AFM, Artificial freshwater medium (18). [b]FM, 0.1 μm-filtered but non-autoclaved freshwater medium. [c]Spent medium, a spent medium of the genus *Limnohabitans* filtrated through 0.1 μm pore-size membrane after cultivation of *Limnohabitans* sp. IMCC26003. For the media abbreviations, refer to Supplementary Table S1.

Table S3. Kinetic parameters of various catalase-peroxidases and bovine catalase.

Source	Molecular weight	Structure	Specific activity (Units mg^{-1})	*Km* (mM)	*kcat* (s^{-1})	*k*cat/*K*m (M^{-1} s^{-1})	pI	Reference
IMCC25003[a]	165,000	A2	179.3	11.7	9.05×10^2	8.01×10^4	7.6[b]	This study
Archaeoglobus fulgidus[a]	NA	NA	5,280	3.8	7.77×10^3	2.04×10^6	5.6[b]	[(19)]
Bacillus selenatarsenatis SF-1	165,000	A2	3,375	2.6	1.15×10^4	4.41×10^6	6.0	[(20)]
Burkholderia pseudomallei[a]	NA	NA	3,630	4.5	5.68×10^3	1.26×10^6	5.9[b]	[(19)]
Escherichia coli K10	337,000	A4	1,486.5	3.9	1.63×10^4	4.19×10^6	5.1[b]	[(21)]
Escherichia coli O157:H7	NA	NA	NA	4.0	1.40×10^4	3.50×10^6	5.1[b]	[(22)]
Geobacillus stearothermophilus[a]	NA	NA	3,120	4.4	1.40×10^3	3.18×10^5	5.2	[(23)]
Halobacterium salinarum	240,000	A4	43.2	3.7	NA	NA	3.8	[(24)]
Mycobacterium smegmatis	NA	NA	NA	1.4	2.38×10^3	1.70×10^6	5.0[b]	[(25)]
Mycobacterium tuberculosis[a]	175,000	A2	2,420	5.2	1.01×10^4	1.94×10^6	5.1	[(26)]
Rhodobacter capsulatus	236,000	A4	7,800	4.2	NA	NA	4.5	[(27)]
Rhodobacter capsulatus[a]	NA	NA	4,830	3.7	6.64×10^3	1.79×10^6	5.1[b]	[(19)]
Synechococcus elongatus PCC 6301[a]	165,000	A2	1,491	4.8	8.85×10^3	1.84×10^6	4.6	[(28)]
Synechococcus elongatus PCC 6301	165,000	A2	NA	4.3	7.20×10^3	1.67×10^6	5.1[b]	[(29)]
Synechococcus elongates PCC 7942[a]	NA	NA	NA	4.2	2.60×10^4	6.19×10^6	5.1[b]	[(30)]
Synechocystis sp. PCC 6803[a]	170,000	A2	5,420	4.9	3.50×10^3	7.14×10^2	5.4	[(31)]
Thermoascus aurantiacus	330,000	A4	NA	48.0	1.07×10^5	2.22×10^6	4.5	[(32)]
Thermus brockianus	178,000	A4	5,300	35.5	6.00×10^3	1.69×10^5	4.7	[(33)]
Bos taurus[c]	240,000	A4	1980.3	20.6	9.05×10^4	8.01×10^4	5.4	This study

[a]Biochemical properties were determined using recombinant catalase-peroxidase. [b]Theoretical pI values were estimated based on amino acids sequences. [c]The monofunctional bovine catalase which was amended to culture media of IMCC25003 was used as an experimental positive control. NA, not available.

Table S4. List of acI genomes used in this study and the presence or absence of *katG* gene.

Tribe	Organism name	Genome ID	Isolation site	Complete	No. of Scaffolds	Genome size (bp)	Length of KatG (aa)
A1	Actinobacteria bacterium IMCC25003	2602042019[a]	Lake Soyang	O	1	1,353,947	746
	actinobacterium SCGC AAA278-O22	2236661007[a]	Lake Mendota	X	43	1,138,490	X
	actinobacterium SCGC AAA027-M14	2236661003[a]	Lake Mendota	X	22	822,296	725
	'Ca. Planktophila dulcis' MMS-IIA-65	CP016777[b]	Lake Zurich	O	1	1,348,019	732
	'Ca. Planktophila dulcis' MMS-IA-53	CP016772[b]	Lake Zurich	O	1	1,365,934	732
	'Ca. Planktophila dulcis' MMS-21-155	CP016770[b]	Lake Zurich	O	1	1,361,776	732
	'Ca. Planktophila sulfonica' MMS-IA-56	CP016773[b]	Lake Zurich	O	1	1,344,614	747
	'Ca. Planktophila versatilis' MMS-IIB-76	CP016778[b]	Lake Zurich	O	1	1,325,420	733
	'Ca. Planktophila versatilis' MMS-IA-79	CP016774[b]	Lake Zurich	O	1	1,331,009	733
	'Ca. Planktophila versatilis' MMS-IA-105	CP016775[b]	Lake Zurich	O	1	1,326,591	733
	'Ca. Planktophila versatilis' MMS-IIB-142	CP016781[b]	Lake Zurich	O	1	1,266,983	733
A2	'Ca. Planktophila limnetica' MMS-VB-114	CP016782[b]	Lake Zurich	O	1	1,328,793	722
A4	Actinobacteria bacterium IMCC26103	2602042020[a]	Lake Soyang	O	1	1,456,516	X
	'Ca. Planktophila lacus' MMS-IIB-106	CP016780[b]	Lake Zurich	O	1	1,384,812	721
	'Ca. Planktophila lacus' MMS-IIB-60	CP016783[b]	Lake Zurich	O	1	1,410,107	721
	'Ca. Planktophila lacus' MMS-21-148	CP016769[b]	Lake Zurich	O	1	1,460,061	721
A5	actinobacterium SCGC AAA044-O16	2606217200[a]	NA	X	17	1,313,698	718
	actinobacterium SCGC AAA028-G02	2606217191[a]	NA	X	18	1,231,401	718
A6	actinobacterium SCGC AAA028-E20	2602042080[a]	NA	X	19	727,714	X
	actinobacterium SCGC AAA028-I14	2619618809[a]	NA	X	11	623,569	717
A7	Actinobacteria bacterium IMCC19121	2606217181[a]	Lake Soyang	O	1	1,506,415	X
	actinobacterium SCGC AAA044-N04	2236661005[a]	Damariscotta Lake	X	23	1,286,658	718
	actinobacterium SCGC AAA024-D14	2264265190[a]	Sparkling Lake	X	82	778,696	X
	actinobacterium SCGC AAA023-J06	2236661001[a]	Sparkling Lake	X	98	695,943	X
	'Ca. Planktophila vernalis' MMS-IIA-15	CP016776[b]	Lake Zurich	O	1	1,364,004	718
B1	actinobacterium SCGC AAA027-L06	2505679121[a]	Lake Mendota	X	75	1,163,583	X
	actinobacterium SCGC AAA027-J17	2236661002[a]	Lake Mendota	X	81	966,755	X
	actinobacterium SCGC AAA278-I18	2236661006[a]	Damariscotta Lake	X	54	944,397	X

	actinobacterium SCGC AAA028-A23	2236661004[a]	Lake Mendota	X	64	833,294	X
	actinobacterium SCGC AAA023-D18	2236661009[a]	Sparkling Lake	X	67	753,259	X
	actinobacterium SCGC AB141-P03	2236876028[a]	Lake Stechlin	X	66	660,403	X
	'Ca. Nanopelagicus limnes' MMS-21-122	CP016768[b]	Lake Zurich	O	1	1,238,108	X
	'Ca. Nanopelagicus hibericus' MMS-21-160	CP016771[b]	Lake Zurich	O	1	1,223,088	X
	'Ca. Nanopelagicus abundans' MMS-IIB-91	CP016779[b]	Lake Zurich	O	1	1,161,863	X
B4	actinobacterium SCGC AAA044-D11	2619618811[a]	NA	X	18	1,095,756	719
C1	Actinobacteria bacterium IMCC26077	2602042021[a]	Lake Soyang	O	1	1,551,612	X

[a]IMG Genome ID (IMG Taxon ID). [b]GenBank accession number. NA, not available.

Table S5. Fatty acids composition (%) of two acI strains.

Fatty acid	IMCC25003	IMCC26103
Saturated fatty acids		
C10:0		0.46
C12:0	1.64	7.93
C14:0	18.22	7.85
C16:0	23.11	28.45
C17:0		1.05
C18:0	2.14	10.49
Unsaturated fatty acids		
C15:1 ω6*c*	1.11	
C17:1 ω8*c*	2.31	1.35
C18:1 ω9*c*	2.10	25.80
summed feature 3 (16:1 ω7*c*/16:1 ω6*c*)	45.79	12.28
summed feature 5 (18:2 ω6,9*c*/18:0 ante)		0.99
summed feature 8 (18:1 ω7*c*, 18:1 ω6*c*)	3.58	3.36

References

1. Konstantinidis KT, Tiedje JM (2005) Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* 102(7):2567-2572.
2. Kim M, Oh HS, Park SC, Chun J (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 64(Pt 2):346-351.
3. Tripp HJ, et al. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452(7188):741-744.
4. Carini P, Steindler L, Beszteri S, Giovannoni SJ (2013) Nutrient requirements for growth of the extreme oligotroph '*Candidatus* Pelagibacter ubique' HTCC1062 on a defined medium. *ISME J* 7(3):592-602.
5. Thomas PE, Ryan D, Levin W (1976) An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Anal Biochem* 75(1):168-176.
6. Ghylin TW, et al. (2014) Comparative single-cell genomics reveals potential ecological niches for the freshwater acI *Actinobacteria* lineage. *ISME J* 8(12):2503-2516.
7. Kang I, Kim S, Islam MR, Cho J-C (2017) The first complete genome sequences of the acI lineage, the most abundant freshwater Actinobacteria, obtained by whole-genome-amplification of dilution-to-extinction cultures. *Sci Rep* 7:42252.
8. Neuenschwander SM, Ghai R, Pernthaler J, Salcher MM (2018) Microdiversification in genome-streamlined ubiquitous freshwater *Actinobacteria*. *ISME J* 12(1):185-198.
9. Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S (2011) A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol Rev* 75(1):14-49.
10. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043-1055
11. Fawal N, et al. (2013) PeroxiBase: a database for large-scale evolutionary analysis of peroxidases. *Nucleic Acids Res* 41:D441-444.

12. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.

13. Zamocky M, Gasselhuber B, Furtmuller PG, Obinger C (2012) Molecular evolution of hydrogen peroxide degrading enzymes. *Arch Biochem Biophys* 525(2):131-144.

14. Buck U, Grossart H-P, Amann R, Pernthaler J (2009) Substrate incorporation patterns of bacterioplankton populations in stratified and mixed waters of a humic lake. *Environ Microbiol* 11(7):1854-1865.

15. Garcia SL, McMahon KD, Grossart H-P, Warnecke F (2014) Successful enrichment of the ubiquitous freshwater acI *Actinobacteria*. *Environ Microbiol Rep* 6(1):21-27.

16. Garcia SL, et al. (2013) Metabolic potential of a single cell belonging to one of the most abundant lineages in freshwater bacterioplankton. *ISME J* 7(1):137-147.

17. Jezbera J, Sharma AK, Brandt U, Doolittle WF, Hahn MW (2009) '*Candidatus* Planktophila limnetica', an actinobacterium representing one of the most numerically important taxa in freshwater bacterioplankton. *Int J Syst Evol Microbiol* 59(11):2864-2869.

18. Kim S, Kang I, Cho JC (2017) Genomic analysis of a freshwater actinobacterium, "*Candidatus* Limnosphaera aquatica" Strain IMCC26207, isolated from Lake Soyang. *J Microbiol Biotechnol* 27(4):825-833.

19. Singh R, et al. (2008) Comparative study of catalase-peroxidases (KatGs). *Arch Biochem Biophys* 471(2):207-214.

20. Gudelj M, et al. (2001) A catalase-peroxidase from a newly isolated thermoalkaliphilic *Bacillus* sp with potential for the treatment of textile bleaching effluents. *Extremophiles* 5(6):423-429.

21. Claiborne A, Fridovich I (1979) Purification of the o-dianisidine peroxidase from *Escherichia coli* B. Physicochemical characterization and analysis of its dual catalatic and peroxidatic activities. *J Biol Chem* 254(10):4245-4252.

22. Varnado CL, Hertwig KM, Thomas R, Roberts JK, Goodwin DC (2004) Properties of a novel periplasmic catalase-peroxidase from *Escherichia coli* O157 : H7. *Arch Biochem Biophys* 421(1):166-174.

23. Kobayashi C, et al. (1997) Thermal conversion from low- to high-activity forms of catalase I from *Bacillus stearothermophilus*. *J Biol Chem* 272(37):23011-23016.

24. Brownpeterson NJ, Salin ML (1993) Purification of a catalase-peroxidase from Halobacterium-Halobium - characterization of some unique properties of the halophilic ezyme. *J Bacteriol* 175(13):4197-4202.

25. Marcinkeviciene JA, Magliozzo RS, Blanchard JS (1995) Purification and characterization of the Mycobacterium-smegmatis catalase-peroxidase involved in isoniazid activation. *J Biol Chem* 270(38):22290-22295.

26. Johnsson K, Froland WA, Schultz PG (1997) Overexpression, purification, and characterization of the catalase-peroxidase KatG from *Mycobacterium tuberculosis*. *J Biol Chem* 272(5):2834-2840.

27. Hochman A, Shemesh A (1987) Purification and characterization of a catalase-peroxidase from the photosynthetic bacterium *Rhodopseudomonas capsulata*. *J Biol Chem* 262(14):6871-6876.

28. Engleder M, et al. (2000) Nucleotide sequence analysis, overexpression in *Escherichia coli* and kinetic characterization of *Anacystis nidulans* catalase-peroxidase. *Biochimie* 82(3):211-219.

29. Obinger C, Regelsberger G, Strasser G, Burner U, Peschek GA (1997) Purification and characterization of a homodimeric catalase-peroxidase from the cyanobacterium *Anacystis nidulans*. *Biochem Biophys Res Commun* 235(3):545-552.

30. Mutsuda M, Ishikawa T, Takeda T, Shigeoka S (1996) The catalase-peroxidase of *Synechococcus* PCC 7942: purification, nucleotide sequence analysis and expression in *Escherichia coli*. *Biochem J* 316:251-257.

31. Jakopitsch C, et al. (1999) Catalase-peroxidase from the cyanobacterium *Synechocystis* PCC 6803: cloning, overexpression in *Escherichia coli*, and kinetic characterization. *Biol Chem* 380(9):1087-1096.

32. Wang HX, Tokusige Y, Shinoyama H, Fujii T, Urakami T (1998) Purification and characterization of a thermostable catalase from culture broth of *Thermoascus aurantiacus*. *J Ferment Bioeng* 85(2):169-173.

33. Thompson VS, Schaller KD, Apel WA (2003) Purification and characterization of a novel thermo-alkali-stable catalase from *Thermus brockianus*. *Biotechnol Prog* 19(4):1292-1299.