VIROLOGICA SINICA 2016, 31 (2): 188–191
DOI: 10.1007/s12250-015-3673-5

LETTER

Narrow distribution of cyanophage *psbA* genes observed in two paddy waters of Northeast China by an incubation experiment

Dear Editor,

Cyanophages are viruses that infect cyanobacteria. They play an important role in shaping the genetic and functional diversity of themselves and their hosts through genetic exchange and shuffling (Lindell et al., 2005; Singh et al., 2012). The *psbA* gene of cyanophages is a "host-derived" gene, and its encoded D1 protein is capable of maintaining photosynthesis in the infected cells when the host's protein synthesis has been shut down by photoinhibition, thus ensuring a source of energy for phage production (Mann et al., 2003; Lindell et al., 2005). High diversity of viral *psbA* sequences was observed in both marine waters (Sharon et al., 2007; Chenard and Suttle, 2008; Sandaa et al., 2008) and lake freshwaters (Chenard and Suttle, 2008; Wilhelm and Matteson, 2008; Ge et al., 2013; Zhong and Jacquet, 2013), but the investigations and analysis of cyanophage *psbA* genes in terrestrial ecosystems have been ignored.

Virus-like particles are more abundant in the paddy water (PW) than those in the ocean (Nakayama et al., 2007). By using culture-independent method, Wang et al. (2009b) firstly obtained 27 different cyanophage *psbA* sequences from Japanese paddy waters (JPWs), and revealed that the distribution of those sequences was more closely related to that from lake freshwaters than that from marine waters. Given that the biomarker genes *g23* of T4-type phages and *g20* of cyanomyoviruses in the paddy fields were different between Northeast (NE) China and Japan (Wang et al., 2009a; Jing et al., 2014), we conjectured that novel sequences or clusters of cyanophage *psbA* genes could be observed in the paddy fields of NE China.

Paddy fields in NE China are irrigated occasionally with river or underground waters, which may impede data interpretation if samplings have been done at inappropriate times. Thus, in this study, an incubation experiment was adopted to survey the *psbA* genes in paddy waters. Briefly, approximate 10 kg of soil samples (0–10 cm depth) were collected from two paddy fields from DaAn (DA) and MuDanJiang (MDJ) in July 2013 (Table 1). The soils were put into different plastic buckets and irrigated with sterilized water to stimulate the cyanobacterial growth. After thorough homogenization of the soil and water, the buckets were placed outside and supplied with sterilized water daily to maintain approximate 10 cm depth of water layer over soils.

After 15 days incubation, approximate 400 mL of surface water was collected from each bucket three times in 15-day intervals. The water samples were centrifuged twice at 5,000 rpm for 30 min at 4 °C followed by filtering the supernatant through 0.4-μm and 0.2-μm filters (Whatman Ltd., UK), and viral particles were collected on 0.03-μm polycarbonate membrane filters (Whatman Ltd., UK). The 0.03-μm filters were put into a tube containing 700 μL 10 mmol/L Tris-HCl buffer (pH 7.5), and treated with DNase and RNase (10 μg/mL each) to digest free DNA and RNA. Viral DNA was extracted and PCR amplification was performed with the degenerate primers *psbA*-F and *psbA*-R according to the methods reported previously (Zeidner et al., 2003; Wang et al., 2009b). PCR products were cloned and sequenced at the Beijing Genomics Institute (BGI, Shenzhen, China). The sequences were deposited in NCBI database with accession numbers KP729651–KP729662, KP729664–KP729667, KP729669–KP729685, KP729687–KP729723 and KP729725–KP729814.

Totally, we obtained 160 different *psbA* sequences in this study. Among them 70, 35, 12, 8, 12 and 23 clones were from water samples PW-DA-I, PW-DA-II, PW-DA-III, PW-MDJ-I, PW-MDJ-II and PW-MDJ-III, respectively. The *psbA* fragments (excluding primers) had lengths of 789–795 bp, encoding 262–264 amino acid residues. BLAST search for the closest relatives at the amino acid level showed that 136 clones were 94%–99% similar to the *psbA* clones from the JPWs; eleven clones were 97%–98% similar to clones of *Synechococcus* myoviruses from the East China Sea; four clones were 97%–98% similar to clones from freshwater Lake Kinneret; five clones were 96%–99% similar to isolates of marine *Synechococcus* phages; four clones (MDJ-I-6, MDJ-I-7, MDJ-III-14, and MDJ-III-19) had the highest similarity to uncultured *Synechococcus* clone and cultured cyanobacteria (Supplementary Table S1).

Phylogenetic relationships among the 160 *psbA* clones obtained in this study were shown in Supplementary Fig-

Table 1. Sampling locations of paddy fields, soil properties and numbers of *psbA* clones obtained in this study.

Sample	Location	Latitude and longtitude	Soil type	Total C (g/kg)	Total N (g/kg)	Total P (g/kg)	pH	Number of *psbA* clones		
								5 August	20 August	4 September
DA	DaAn, Jilin	45°36'N, 123°50'E	Saline-alkaline soil	10.49	0.43	0.37	7.68	70	35	12
MDJ	MuDanJiang, Heilongjiang	44°26'N, 129°29'E	Dark brown soil	14.99	1.07	0.76	6.27	8	12	23

ure S1. All the clones were roughly grouped into at least seven clusters, of which clusters I and V were further divided into six and two subclusters, respectively (Supplementary Figure S1). In addition, the GC contents of all clones were in the range of 44.15~ 59.25%, and the triplet peptides of the D1 protein motif $^{R}/_{K}$ETTXXXS $^{Q}/_{H}$ (Sharon et al., 2007) included EEV, EQE, EVE, ENE, ESE, ETE, and EAE (Supplementary Figure S1).

Given that the *psbA* gene of cyanophage is a "host-derived" gene (Sullivan et al., 2006), problems generally arise on how to distinguish them from their host genes. Phylogenetic tree was a common criterion of distinguishing the environmental *psbA* sequences coming from phage and host (Sullivan et al., 2006; Chenard and Suttle, 2008). Therefore, all the *psbA* clones of this study with cyanophage *psbA* clones from JPWs, marine waters and lake freshwaters, and from isolated cyanophages, as well as *psbA* sequences from several cyanobacteria such as *Synechococcus*, *Anabeana* and *Nostoc* were subjected to building the phylogenetic tree, and the tree was roughly separated into five major clusters (Figure 1).

Cluster α was the largest cluster, consisting of four subclusters. Approximately 86% (138 clones) of the clones from this study, together with environmental cyanophage clones from marine waters, lake freshwaters and JPWs, as well as isolated phages infecting marine *Synechococcus* were grouped into this cluster. Cluster β consisted of two paddy water clones from NE China and Japan with 95% similarity between them (Supplementary Table S1). Cluster δ contained three clones obtained in this study and 27 cultured cyanobacteria. We, therefore, concluded these three clones were not from cyanophages. We also found that none of *psbA* clones from paddy waters fell into cluster ε, which contained isolates of *Synechococcus* and *Prochlorococcus*, *Synechococcus* cyanophage S-SSM1, *Prochlorococcus* phages, and marine clones. Cluster γ was divided into two subclusters. Subcluster γ-1 contained 13 clones from this study, three marine cyanophage clones and one isolated cyanophage S-PM2. All *psbA* sequences from paddy waters in this subcluster had viral-like triplet peptides of EQE and ENE, or virus and host like peptide ETE (Sharon et al., 2007). The findings inferred that the *psbA* sequences in subcluster γ-1 originated from cyanophages. In contrast, subcluster γ-2 contained five clones obtained in this study with three sequences from *Synechococcus* and two clones (CF8; CF(-N)7) from JPWs. It should be noted that the origins of clones CF8 and CF(-N)7 from phages or cyanobacteria were ambiguous (Wang et al., 2009b). Based on the finding of this study, we suggested that the *psbA* clones in this subcluster might not be from cyanophages (Figure 1). Overall, we isolated 152 different sequences of cyanophage *psbA* genes from paddy waters of NE China.

The *psbA* sequences of cyanophages obtained from this study were compared with those from JPWs. In detail, the two groups PFW-1a and PFW-2 consisted of cyanophage *psbA* sequences only from JPWs, and seven newly designed paddy water groups (PW-1~PW-7) contained clones exclusively from this study, with an exception of a clone CF(-N)1 from JPWs falling into PW-7. This finding suggested that the *psbA* gene-containing cyanophage communities in paddy fields in NE China were seemingly different from those in Japan. Notably, since the sequence number was largely unbalanced between the two studies (152 versus 27), we speculated that new insights into phage diversity or distribution would be highlighted as more viral *psbA* sequences being obtained from paddy waters in Japan or other countries.

Based on the overall structure of the Figure 1, subclusters α-1 and α-2 could be considered as specific clusters for the lake freshwaters and paddy waters, respectively. In addition, both subclusters were split from the same clade, which was far away from the miscellaneous subcluster α-3 containing isolated cyanophages and cyanophage clones from lake freshwaters, marine waters, paddy waters, and from subcluster α-4 containing isolated cyanomyoviruses infecting marine *Synechococcus*. These findings clearly indicated that the distribution of cyanophage *psbA* sequences from paddy waters was different from those from lake freshwaters and marine waters.

In conclusion, we successfully isolated 152 cyanophage *psbA* clones from two paddy waters of NE China. Although the majority of *psbA* sequences observed in this study had relative high similarity with those from Japanese paddy waters, two and seven specific cyanophage *psbA* groups were constructed in JPWs and this study, respectively, which suggested that cyanophage *psbA*

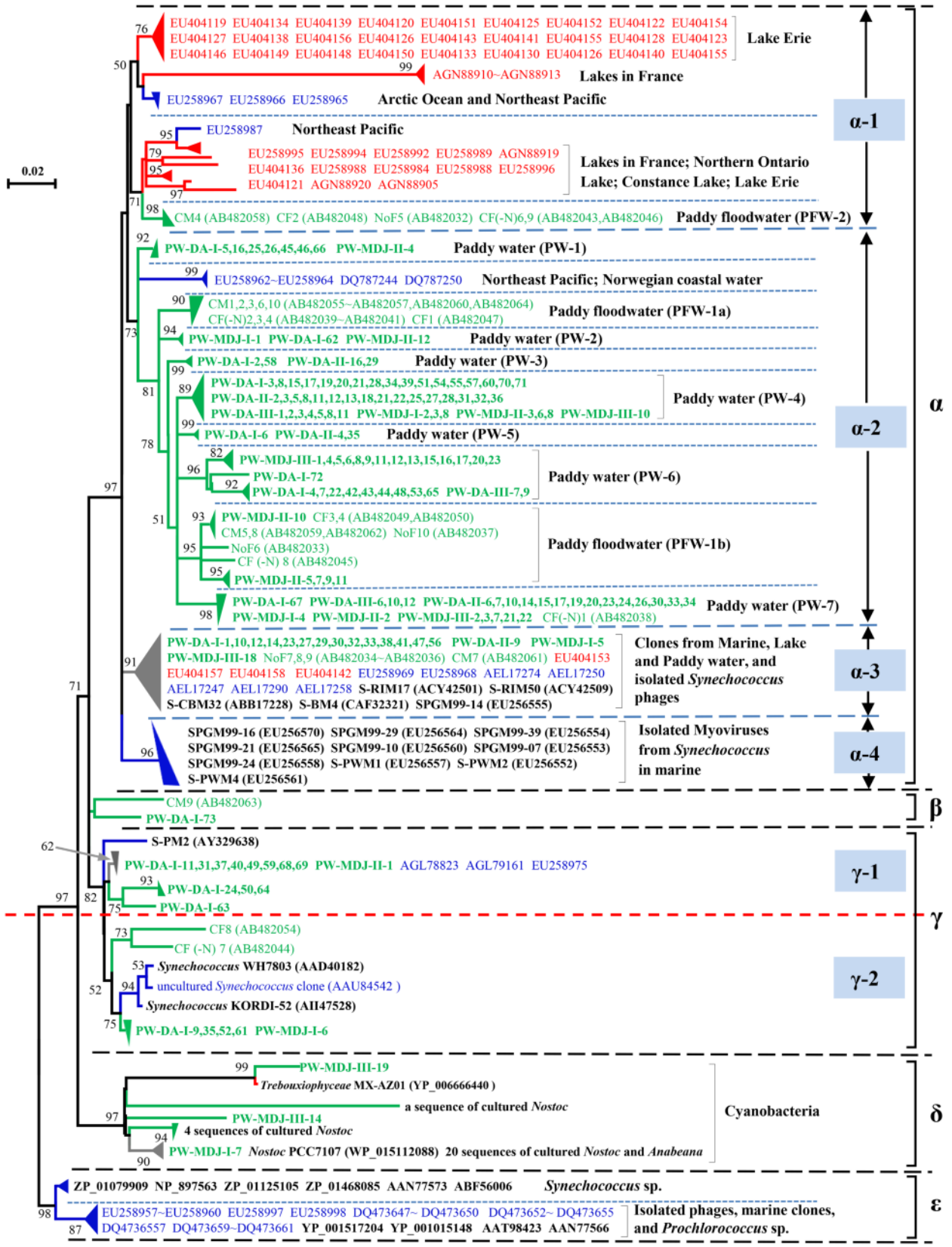
0.02
EU404119 EU404134 EU404139 EU404120 EU404151 EU404125 EU404152 EU404122 EU404154
EU404127 EU404138 EU404156 EU404126 EU404143 EU404141 EU404155 EU404128 EU404123
EU404146 EU404149 EU404148 EU404150 EU404133 EU404130 EU404126 EU404140 EU404155
Lake Erie
AGN88910~AGN88913 Lakes in France
EU258967 EU258966 EU258965 Arctic Ocean and Northeast Pacific
EU258987 Northeast Pacific
EU258995 EU258994 EU258992 EU258989 AGN88919
EU404136 EU258988 EU258984 EU258988 EU258996
EU404121 AGN88920 AGN88905
Lakes in France; Northern Ontario Lake; Constance Lake; Lake Erie
CM4 (AB482058) CF2 (AB482048) NoF5 (AB482032) CF(-N)6,9 (AB482043,AB482046) Paddy floodwater (PFW-2)
PW-DA-I-5,16,25,26,45,46,66 PW-MDJ-II-4 Paddy water (PW-1)
EU258962~EU258964 DQ787244 DQ787250 Northeast Pacific; Norwegian coastal water
CM1,2,3,6,10 (AB482055~AB482057,AB482060,AB482064)
CF(-N)2,3,4 (AB482039~AB482041) CF1 (AB482047) Paddy floodwater (PFW-1a)
PW-MDJ-I-1 PW-DA-I-62 PW-MDJ-II-12 Paddy water (PW-2)
PW-DA-I-2,58 PW-DA-II-16,29 Paddy water (PW-3)
PW-DA-I-3,8,15,17,19,20,21,28,34,39,51,54,55,57,60,70,71
PW-DA-II-2,3,5,8,11,12,13,18,21,22,25,27,28,31,32,36
PW-DA-III-1,2,3,4,5,8,11 PW-MDJ-I-2,3,8 PW-MDJ-II-3,6,8 PW-MDJ-III-10 Paddy water (PW-4)
PW-DA-I-6 PW-DA-II-4,35 Paddy water (PW-5)
PW-MDJ-III-1,4,5,6,8,9,11,12,13,15,16,17,20,23
PW-DA-I-72
PW-DA-I-4,7,22,42,43,44,48,53,65 PW-DA-III-7,9 Paddy water (PW-6)
PW-MDJ-II-10 CF3,4 (AB482049,AB482050)
CM5,8 (AB482059,AB482062) NoF10 (AB482037)
NoF6 (AB482033)
CF (-N) 8 (AB482045)
PW-MDJ-II-5,7,9,11 Paddy floodwater (PFW-1b)
PW-DA-I-67 PW-DA-III-6,10,12 PW-DA-II-6,7,10,14,15,17,19,20,23,24,26,30,33,34
PW-MDJ-I-4 PW-MDJ-II-2 PW-MDJ-III-2,3,7,21,22 CF(-N)1 (AB482038) Paddy water (PW-7)
PW-DA-I-1,10,12,14,23,27,29,30,32,33,38,41,47,56 PW-DA-II-9 PW-MDJ-I-5
PW-MDJ-III-18 NoF7,8,9 (AB482034~AB482036) CM7 (AB482061) EU404153
EU404157 EU404158 EU404142 EU258969 EU258968 AEL17274 AEL17250
AEL17247 AEL17290 AEL17258 S-RIM17 (ACY42501) S-RIM50 (ACY42509)
S-CBM32 (ABB17228) S-BM4 (CAF32321) SPGM99-14 (EU256555)
Clones from Marine, Lake and Paddy water, and isolated *Synechococcus* phages
SPGM99-16 (EU256570) SPGM99-29 (EU256564) SPGM99-39 (EU256554)
SPGM99-21 (EU256565) SPGM99-10 (EU256560) SPGM99-07 (EU256553)
SPGM99-24 (EU256558) S-PWM1 (EU256557) S-PWM2 (EU256552)
S-PWM4 (EU256561)
Isolated Myoviruses from *Synechococcus* in marine
CM9 (AB482063)
PW-DA-I-73
S-PM2 (AY329638)
PW-DA-I-11,31,37,40,49,59,68,69 PW-MDJ-II-1 AGL78823 AGL79161 EU258975
PW-DA-I-24,50,64
PW-DA-I-63
CF8 (AB482054)
CF (-N) 7 (AB482044)
Synechococcus WH7803 (AAD40182)
uncultured *Synechococcus* clone (AAU84542)
Synechococcus KORDI-52 (AII47528)
PW-DA-I-9,35,52,61 PW-MDJ-I-6
PW-MDJ-III-19
Trebouxiophyceae MX-AZ01 (YP_006666440)
a sequence of cultured *Nostoc*
PW-MDJ-III-14
4 sequences of cultured *Nostoc*
PW-MDJ-I-7 *Nostoc* PCC7107 (WP_015112088) 20 sequences of cultured *Nostoc* and *Anabeana*
Cyanobacteria
ZP_01079909 NP_897563 ZP_01125105 ZP_01468085 AAN77573 ABF56006 *Synechococcus* sp.
EU258957~EU258960 EU258997 EU258998 DQ473647~ DQ473650 DQ473652~ DQ473655
DQ4736557 DQ473659~DQ473661 YP_001517204 YP_001015148 AAT98423 AAN77566
Isolated phages, marine clones, and *Prochlorococcus* sp.
α-1
α-2
α-3
α-4
α
β
γ-1
γ
γ-2
δ
ε

Figure 1. Neighbour-joining phylogenetic tree of 330 *psbA* sequences, showing the relationships of *psbA* amino acid sequences from paddy waters in Northeast (NE) China in this study (green words in bold) with those cloned from marine waters (blue words), lake freshwaters (red words), Japanese paddy floodwaters (green words in normal) and obtained from isolated cyanophages and bacteria (black words in bold). Clones in this study were named to reflect sampling sites, sampling data and clone number. In detail, PW stands for paddy water; DA and MDJ respectively stands for DaAn and MuDanJiang; I, II, III stands for sampling date on 5 August, 20 August and 4 September, 2013, respectively. The blue, red and green triangles indicate *psbA* clusters from marine waters, lake freshwaters and paddy waters, respectively, and the grey triangles represent *psbA* clusters from more than one environment. The two groups PFW-1 (being further divided into PFW-1a and PFW-1b based on the data of this study) and PFW-2 were previously observed in Japanese paddy floodwaters (Wang et al., 2009b) and seven new paddy water groups (PW-1~PW-7) were designed in this study. Numbers in parentheses are the accession numbers of the *psbA* sequences from the NCBI website. The clones obtained in this study above the red dotted line were confirmed as phage origins.

assemblages in paddy waters, to some extent were different between two countries. In addition, we also found that the majority of *psbA* clones from paddy waters of both countries fell into the subcluster α-2, suggesting that the distribution of cyanophage *psbA* genes in paddy waters is narrow, but far away from those from environmental freshwater and seawater.

FOOTNOTES

This study was financially supported by grants from National Nature Science Foundation of China (41271262), and the Strategic Priority Research Program of Chinese Academy of Sciences (XDB15010103). The authors declare that they have no conflict of interest. This article does not contain any studies with human or animal subjects performed by any of the authors.

Suppementary figures/tables are available on the website of *Virologica Sinica*: www.virosin.org; link.springer.com/journal/12250.

Xinzhen Wang[1,2], Ruiyong Jing[1,3], Junjie Liu[1], Zhenhua Yu[1], Jian Jin[1], Xiaobing Liu[1], Xiaojuan Wang[4], Guanghua Wang[1]✉

1. Key Laboratory of Mollisols Agroecology, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China
2. University of Chinese Academy of Sciences, Beijing 100049, China
3. College of Life and Sci-technology, Heilongjiang BaYi Agricultural University, Daqing 163319, China
4. Centre for AgriBioscience, La Trobe University, Melbourne Campus, Bundoora, Vic 3086, Australia
✉Correspondence:
Phone: +86-451-86602745, Fax: +86-451-86603736,
Email: wanggh@iga.ac.cn, guanghuawang@hotmail.com
ORCID: 0000-0003-4737-8126

Published online: 26 January 2016

REFERENCES

Chenard C, Suttle C. 2008. Appl Environ Microbiol, 74: 5317–5324.
Ge XY, Wu YQ, Wang MN, et al. 2013. Virol Sin, 28: 280–290.
Jing RY, Liu JJ, Yu ZH, et al. 2014. PloS One, 9: e88634.
Lindell D, Jaffe JD, Johnson ZI, et al. 2005. Nature, 438: 86–89.
Mann NH, Cook A, Millard A, et al. 2003. Nature, 424: 741–741.
Nakayama N, Okumura M, Inoue K, et al. 2007. Soil Sci Plant Nutr, 53: 420–429.
Sandaa RA, Clokie M, Mann NH, et al. 2008. FEMS Microbiol Ecol, 63: 2–11.
Sharon I, Tzahor S, Williamson S, et al. 2007. ISME J, 1: 492–501.
Singh P, Singh SS, Srivastava A, et al. 2012. Afr J Biotechnol, 11: 2591–2608.
Sullivan MB, Lindell D, Lee JA, et al. 2006. PLoS Biol, 4: e234.
Wang GH, Jin J, Asakawa S, et al. 2009a. Soil Biol Biochem, 41: 423–427.
Wang GH, Murase J, Asakawa S, et al. 2009b. FEMS Microbiol Ecol, 70: 79–86.
Wilhelm SW, Matteson AR. 2008. Freshwater Biol, 53: 1076–1089.
Zeidner G, Preston CM, Delong EF, et al. 2003. Environ Microbiol, 5: 212–216.
Zhong X, Jacquet S. 2013. Appl Environ Microbiol, 79: 7169–7178.