Korean Journal of Environmental Biology

Original article

https://doi.org/10.11626/KJEB.2020.38.3.424

Korean J. Environ. Biol.
38(3) : 424-432 (2020)
ISSN 1226-9999 (print)
ISSN 2287-7851 (online)

Metaproteomic analysis of harmful algal bloom in the Daechung reservoir, Korea

Jong-Soon Choi[1,2,*,†], Yun Hwan Park[3], Soo Hyeon Kim[1,2], Ju Seong Park[1,2] and Yoon-E Choi[3,*,†]

[1]Biological Disaster Analysis Team, Korea Basic Science Institute, Daejeon 34133, Republic of Korea
[2]Graduate School of Analytical Science and Technology, Chungnam National University, Daejeon 34134, Republic of Korea
[3]Division of Environmental Science and Ecological Engineering, Korea University, Seoul 02841, Republic of Korea

***Corresponding author**
Jong-Soon Choi
Tel. 042-865-3428
E-mail. jschoi@kbsi.re.kr

Yoon-E Choi
Tel. 02-3290-3040
E-mail. yechoi@korea.ac.kr
† These authors contributed equally to this work as co-corresponding authors.

Received: 29 July 2020
Revised: 10 September 2020
Revision accepted: 11 September 2020

Abstract: The present study aimed to analyze the metaproteome of the microbial community comprising harmful algal bloom (HAB) in the Daechung reservoir, Korea. HAB samples located at GPS coordinates of 36°29′N latitude and 127°28′E longitude were harvested in October 2013. Microscopic observation of the HAB samples revealed red signals that were presumably caused by the autofluorescence of chlorophyll and phycocyanin in viable cyanobacteria. Metaproteomic analysis was performed by a gel-based shotgun proteomic method. Protein identification was conducted through a two-step analysis including a forward search strategy (FSS) (random search with the National Center for Biotechnology Information (NCBI), Cyanobase, and Phytozome), and a subsequent reverse search strategy (RSS) (additional Cyanobase search with a decoy database). The total number of proteins identified by the two-step analysis (FSS and RSS) was 1.8-fold higher than that by one-step analysis (FSS only). A total of 194 proteins were assigned to 12 cyanobacterial species (99 mol%) and one green algae species (1 mol%). Among the species identified, the toxic microcystin-producing *Microcystis aeruginosa* NIES-843 (62.3%) species was the most dominant. The largest functional category was proteins belonging to the energy category (39%), followed by metabolism (15%), and translation (12%). This study will be a good reference for monitoring ecological variations at the meta-protein level of aquatic microalgae for understanding HAB.

Keywords: metaproteome, harmful algal bloom, Daechung reservoir, one & two-step analysis, forward search strategy, reverse search strategy

INTRODUCTION

Harmful algal bloom (HAB) is the outgrowth of microalgae and cyanobacteria colonies that occur in sea and freshwater. Occasionally, HAB produces toxic compounds with harmful effects on marine and freshwater bio-systems, including humans (Heisler *et al.* 2008). Even worse, microbial populations from algal blooms are decomposed in aquatic ecosystems and deplete the oxygen supply, resulting in a hypoxic dead zone where fish and plants can no longer survive. Although many factors contribute to the occurrence of HAB, it is unclear exactly how HAB results from these conditions. Human activities and the resulting water pollution from livestock excretions or excessive fertilizer increase the load of abundant phosphorus and nitrogen (nitrate, ammonia, and urea), allowing HAB to flourish (Oenema 2004). Climate changes such as global warming, extreme rainfall, and drought events exacerbate the prob-

lems of HAB synergistically with human activities (Pearl *et al.* 2018).

HAB occasionally appears bright or deep green in color due to formation of scum, foam, or benthic mat, depending on the pigments the cyanobacterial species possess. Cyanobacterial blooms occur in freshwater lakes, rivers, and coastal areas. In particular, cyanobacteria produce neurotoxins such as microcystins, which destroy the nerve and liver tissues of mammals, including humans, causing neuroparalysis and hepatocarcinoma, respectively (Blaha *et al.* 2009). Thus, HABs are a concern worldwide because of the potential threat they pose to the environment and humans, along with increasing economic loss (McPartlin *et al.* 2017). There have been many attempts to systematically monitor and manage freshwater quality in various ecosystems in order to protect environmental sustainability and public health (Orme-Zavaleta *et al.* 2008). In South Korea, an algae warning system has been established by the Ministry of Environment that has reported survey data from 28 major water resources. Water quality has been forecasted on the basis of collected data regarding water temperature and weather observations for major reservoirs on four major rivers (Lee *et al.* 2018). Here, for metaproteomic analysis, we selected Daechung Reservoir where cyanobacterial blooms occur every summer.

Metaproteomics is a technique of meta-omics analysis that reveals the metabolic repertoire by linking specific proteins to their corresponding organisms. In addition, metaproteomics helps to elucidate the dominant proteins and organisms that comprise a dynamic bio-community such as HAB species (Hettich *et al.* 2013). Traditional meta-omics refers to metagenomics, which is based on massive 16S and 18S ribosomal DNA sequencing for taxonomic profiling to understand the genetic diversity of a microbial community. However, the genetic diversity resulting from metagenomics provides limited information regarding the landscape of bio-community. In recent years, several metaproteomic analyses have been conducted: analysis of proteorhodopsin in the ubiquitous marine bacterium SAR11 (Giovannoni *et al.* 2005), of microbial plankton in a highly productive coastal upwelling system (Sowell *et al.* 2011), and of microbial populations active in biogeochemical cycling (Hanson *et al.* 2014). Based on a previous report on the metaproteomic analysis of a freshwater microbial community (Russo *et al.* 2014), we attempted to analyze the metaproteome of an algal bloom site in the Daechung reservoir, Korea. Protein identification was performed through a two-step database search analysis (Jagtap *et al.* 2013.). Two-step analysis of protein identification in the present study consisted of a forward search strategy (FSS) using large datasets such as NCBI, Cyanobase, and Phytozome, and a subsequent reverse search strategy (RSS) using the target decoy Cyanobase. Our first attempt to conduct a metaproteomic analysis of HAB in Korea will provide a reference for evaluating and understanding the detrimental microbial community.

MATERIALS AND METHODS

1. Sample preparation

On October 14, 2013, crude HAB samples were harvested at latitude 36°29′N and longitude 127°28′E in GPS coordinates. The protein sample preparation was performed with modification according to a previous protocol (Shin *et al.* 2008). In brief, a total of 2 L of harvested samples were centrifuged at 3,000 *g* for 10 min at room temperature. The precipitate was washed with 10 mM Tris-HCl (pH 8.0) buffer containing a protease inhibitor cocktail. The cell pellet was resuspended in TSD buffer (10 mM Tris-EDTA, 0.1% (w/v) SDS, 1 mM DTT) (Williams *et al.* 2013) or Triton X-114 buffer (10 mM Tris-HCl, pH 7.4, 0.15 M NaCl, 1 mM EDTA and PBS containing 1% (v/v) Triton X-114) (Shevchenko *et al.* 2012). The mixture was disrupted at 20,000 psi using a French Pressure cell press (SLM Aminco 40K, Thermofisher, USA). The crude extract was agitated mildly at 4°C for 1 hr. After centrifugation at 12,000 rpm for 20 min at 4°C, the supernatant was added to 4 volumes of 100% (v/v) acetone and left at −20°C overnight. The mixture was centrifuged at 12,000 rpm for 20 min at 4°C. The pellet was washed with 80% (v/v) acetone three times and subjected to freeze drying (Labconco FreeZone 4.5, Labcono Co., USA) prior to protein analysis. Protein quantification was performed by Peterson's method (Peterson *et al.* 1977).

2. SDS-PAGE and gel slicing

For the gel-based shotgun proteomic analysis, HAB proteins (25 μg per lane) were loaded on a 12% SDS-polyacrylamide gel as previously reported (Lee *et al.* 2015). After electrophoresis, the gels were stained with Coomassie Brilliant Blue (CBB) R250 and cut into 10 slices according to the stained gel band intensity. Each gel slice was transferred to a new Eppendorf tube and subjected to in-gel tryptic digestion. The tryptic digests were extracted with

0.02% (v/v) formic acid and 0.5% (v/v) acetic acid and applied to LC-MS/MS analysis.

3. MS analysis of HAB samples

Peptides were separated using liquid chromatography integrated with electrospray ionization MS (LCQ-DecaXP, Thermofisher, USA). The separated peptides were then eluted from the reverse column in a gradient of 0−65% (v/v) acetonitrile for 80 min at a flow rate of 120 nL min^{-1}. All MS and MS/MS spectra were detected in a data-dependent mode by an LTQ-Velos ESI ion trap MS at the Korea Basic Science Institute (www.kbsi.re.kr). LC-MS/MS analysis was performed in triplicate with different batch samples. Relative quantification of identified protein was performed and the data expressed as mol% of spectral count (Oh *et al.* 2001). Spectral counts imply the total number of MS/MS spectra assigned to a particular protein from a specific database.

4. Bioinformatic analysis

MS/MS spectra were searched in the specified database using MASCOT version 2.4 (www.matrixsciences.com). Database searches were conducted through either one-step or two-step approach. The one-step approach corresponded to either a forward search strategy (FSS) or a reverse search strategy (RSS). FSS is a straightforward search using available large datasets such as NCBI, Cyanobase, and Phytozome. RSS is exclusive protein identification using Cyanobase and its decoy database with a false discovery rate of 5% or less. The two-step approach consists of the sequential search methods of FSS and RSS. Two missed cleavages were allowed to identify proteins with cysteine carbamidomethylation (+57) and methionine oxidation (+16) as fixed and variable modifications, respectively. Peptide identification of LC-MS/MS spectra was performed on the basis of 95% confidence probability with two or more positive detections out of triplicates.

RESULTS AND DISCUSSION

1. Microscopic observation of HAB in the Daechung Reservoir

The Daechung reservoir is a representative lake that is the 4th largest and eutrophic lake to exhibit annual HABs in South Korea. The artificial Lake Daechung was periodically monitored, and seasonal variation of cyanobacterial microcystins was reported (Ishiahara *et al.* 2005). Previously, the intergenic space of *cpc*BA diversity was investigated using 16S rRNA analysis to elucidate the composition and dynamics of cyanobacterial bloom in the Daechung reservoir (Kim *et al.* 2006). In the present study, prior to the metaproteomic analysis of the Daechung HAB, light and fluorescence microscopic analysis was conducted. HAB samples taken from the basin around Daecheong Dam in October 2013 were green (Fig. 1A, B). As shown in the micrographs, clustered microorganisms composed of hundreds of thousands of cell aggregates were observed. Most of the cells appeared red and a few were blue under a fluorescence microscope (Fig. 1C, D). The red color was presumably caused by the autofluorescence of chlorophyll and phycocyanin from cyanobacteria. Red chlorophyll autofluorescence signals are used to assay the viability of cyanobacteria (Schulze *et al.* 2011.). Thus, cell aggregates from the Daechung reservoir HAB collected in October appeared to be mostly viable.

2. Results of metaproteomic analysis by two strategies

In order to efficiently profile and evaluate the metaproteome from the Daechung reservoir HAB, we performed gel-based shotgun proteomics using LTQ-Velos ion-trap MS. Generally, traditional proteomics of human species handles around 7×10^4 sequences or fewer, while metaproteomics of human saliva covers more than 5×10^5 sequences (Jagtap *et al.* 2012). Thus, database searches against large datasets require search engine space and possess a possibility of false positives. To overcome this challenge, an increased stringency of protein identification is applied to metaproteomic analysis; however, a resulting increase of false-negatives occurs, leading to a decreased number of high confidence microbial peptides (Cargile *et al.* 2004; Jagtap *et al.* 2013). Therefore, we applied the two-step method used in human microbiomes to the HAB metaproteome for database search. First, a forward search strategy (FSS) was conducted in large databases such as NCBI, Cyanobase, and Phytozome. Second, a reverse search strategy (RSS) was performed using the exclusive Cyanobase search with a target decoy database with a false discovery rate of 5% or less. The protein numbers identified independently by FSS and RSS were 108 and 158, respectively (Fig. 2A). Unexpectedly, two-step analysis resulted in the identification of 194 proteins, a 1.8-fold and

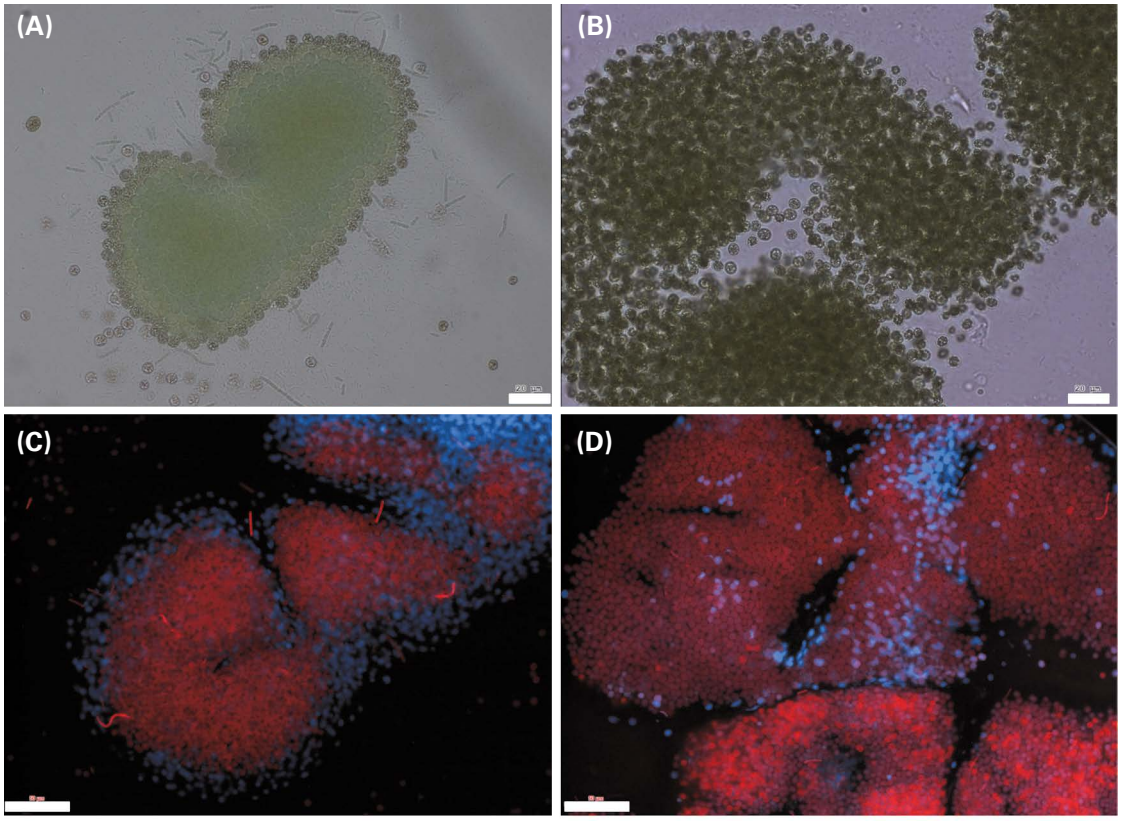

Fig. 1. Microscopy pictures of HAB samples. Light microscopy images. Most of the clustered cells are *Microcyctsis* sp. (A, B). Fluorescence microscopy images were taken with a Zeiss LSM 900 (C, D). Scale bar, 20 μm.

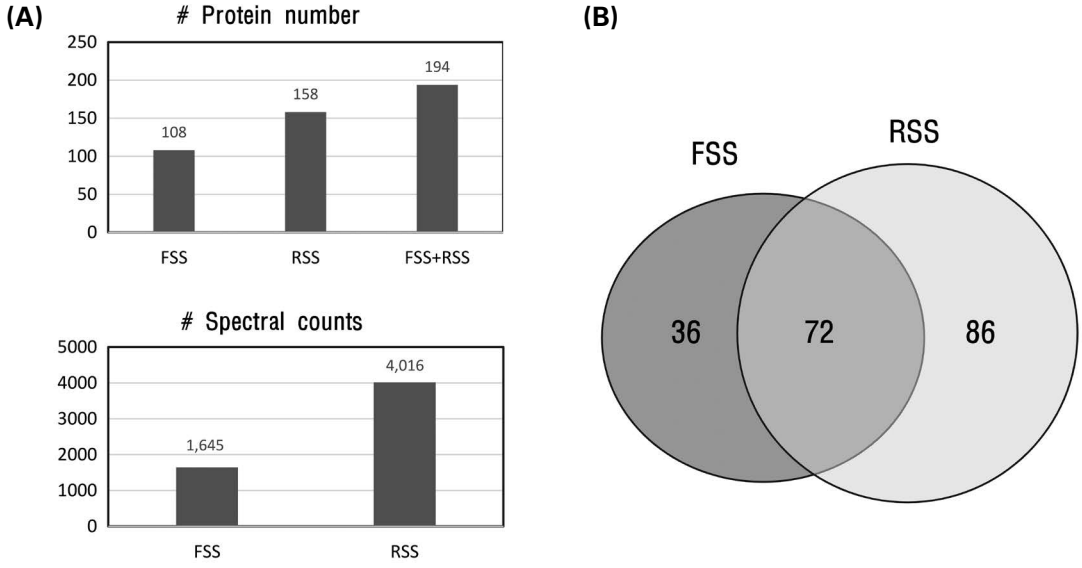

Fig. 2. Comparison between the FSS (forward search strategy) and RSS (reverse search strategy). (A) Number of proteins identified using one-step (FSS) and two-step analysis (FSS + RSS) (upper). Spectral counts obtained using FSS and RSS (lower). (B) Venn diagram of proteins identified using FSS and RSS.

Table 1. Summary of proteins identified by one-step (forward search strategy) and two-step (forward search strategy + reverse search strategy) analyses

Microorganism	One-step analysis		Two-step analysis	
	# of Proteins	Sum of Mol%	# of Proteins	Sum of Mol%
Cyanobase	76	79%	162	83%
Anabaena sp. PCC7120	1	0.3%	16	6.2%
Gloeobacter violaceus PCC7421			7	3.1%
Microcystis aeruginosa NIES-843	62	70.2%	111	62.3%
Synechocystis sp. PCC6803	8	4.6%	18	6.9%
Thermosynechococcus elongates BP-1	5	3.8%	10	4.5%
NCBI	29	20%	29	16%
Anabaena sp. PCC90*	7	3.2%	7	2.6%
Anabaena cylindrica PCC7122	3	2.0%	3	1.6%
Cyanobacterium aponinum PCC10605	8	8.3%	8	6.8%
Cyanobacterium stanieri PCC7202	2	1.0%	2	0.8%
Cyanobacterium thalassa UCYN-A	1	0.3%	1	0.3%
Oscillatoria acuminate PCC6304	4	3.1%	4	2.5%
Oscillatoria nigro-viridis PCC7112	4	2.1%	4	1.7%
Phytozome	3	1%	3	1%
Clamydomonas reinhardtii 236	3	1.1%		
Total	108	100%	194	100%

* The *Anabaena* genus is also known as *Nostoc*

Table 2. Most abundant proteins in the Daechung reservoir

Accession	Species	Description	Category	Mol%
MAE47890	NIES-843*	Ribulose bisphosphate carboxylase large subunit (RbcL)	Energy	4.36%
MAE46080	NIES-843	60 kDa chaperonin (GroL)	Protein Destination	3.93%
MAE06090	NIES-843	Porin type major outer membrane protein (OmpA)	Cell Structure	3.88%
MAE00920	NIES-843	ATP synthase beta subunit (AtpB)	Energy	3.83%
MAE49370	NIES-843	Phycobilisome core-membrane linker polypeptide (ApcE)	Energy	3.81%
MAE27990	NIES-843	Porin type major outer membrane protein (OmpA)	Cell Structure	3.30%
MAE37620	NIES-843	Gas vesicle structural protein (GvpC)	Cell Structure	2.99%
MAE24450	NIES-843	Phycocyanin beta-subunit (CpcB)	Energy	2.66%
MAE03410	NIES-843	60 kDa molecular chaperonin 2 (GroL2)	Protein Destination	2.60%
Slr0009	PCC6803**	Ribulose bisphosphate carboxylase large subunit (RbcL)	Energy	2.42%

* *Microcystis aeruginosa* NIES-843, ** *Synechocystis* sp. PCC6803

1.2-fold increase compared to FSS and RSS, respectively. The number of spectral counts resulting from FSS and RSS were 1,645 and 4,016, respectively. A Venn diagram of proteins identified by FSS and RSS showed 72 shared proteins (Fig. 2B). FSS and RSS exclusively contained 36 and 86 proteins, respectively. The protein lists identified through either one-step or two-step analysis are attached (Supplementary Table 1; Supplementary Table 2). Our results show that two-step analysis is suitable for efficient identification of larger metaproteomes in which RSS greatly contributes to the identification of more cyanobacterial proteins. Recently, dedicated algorithms and software have been developed to handle to the growing data within metaproteomics (Heyer *et al.* 2017).

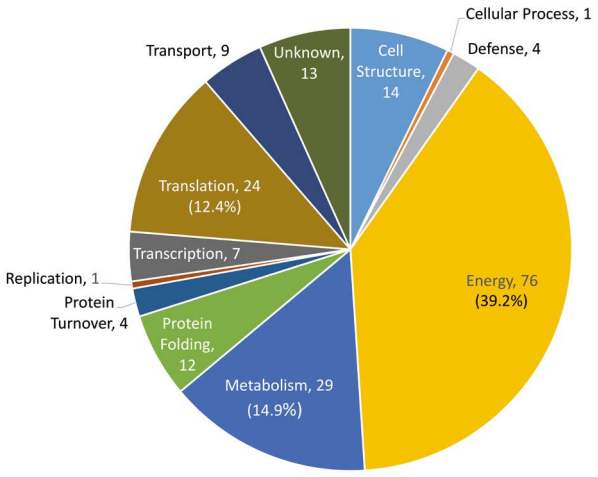

Fig. 3. Gene ontology of identified whole proteins based on functional categories. Percentages are shown in parentheses only for categories with more than 10%.

3. Microalgae profiles of HAB in the Daechung Reservoir

According to two-step analysis of HAB samples in the Daechung reservoir, a total of 194 proteins were assigned to 12 cyanobacterial species (99 mol%) and 1 green algae species (1 mol%). Using one-step analysis by FSS, 76 proteins (subtotal 79 mol%), 29 proteins (subtotal 20%), and 3 proteins (1%) were identified in Cyanobase, NCBI, and Phytozome, respectively (Table 1). However, during two-step analysis using sequential searches of FSS and RSS, 162 proteins (subtotal 83 mol%) were collected using Cyanobase, approximately double the number identified during one-step analysis, while the identification numbers from the NCBI and Phytozome datasets remain unchanged. This increase in identification during two-step analysis was presumably caused by a reduction of the false negative peptide sequence matches. Among the species assigned, the toxic microcystin-producing *Microcystis aeruginosa* NIES-843 (62.3%) (hereafter, NIES-843) was the most dominant species in the Daechung reservoir. NIES-843 is known to produce toxic cyanobacterial blooms in freshwater ecosystems (Steffen *et al.* 2012; Zhao *et al.* 2018). *Synechocystis* sp. PCC6803 (hereafter, PCC6803) was found to be the second most dominant microorganism in the HAB metaproteome. Greater protein identification of cyanobacteria resulted from application of the systematic database Cyanobase (Nakao *et al.* 2010). RSS assigned 5 cyanobacterial species out of the 39 species contained in the current Cyanobase. Conclusively, two-step analysis of the HAB metaproteome is an effective method that increases the number of peptide sequence matches with high confidence.

4. Characteristics of the HAB metaproteome in the Daechung Reservoir

A total of 194 proteins were classified into 12 different categories on the basis of putative physiological functions. The largest functional category included proteins in the energy category (39.2%), comprised of photosynthesis and carbon fixation pathways. Two prevalent proteins identified were ribulose bisphosphate carboxylase large subunit (RbcL) from NIES-843 and PCC6803 and ATP synthase beta subunit (AtpB) from NIES-843 (Supplementary Table 2). RbcL has been used to monitor algal bloom in coastal environments in North America, Japan, and Korea (Ki *et al.* 2007). AtpB was previously identified as a novel target, in addition to phosphatases (PP1 and PP2A), of microcystins observed during algal bloom (Mikhailov *et al.* 2003). Harmful blooms of *Microcystis* sp. were reported to be a driver of non-nitrogen fixing cyanobacteria in Lake Erie (Newell *et al.* 2019). Moreover, the dormant *Microcystis aeruginosa* was reported to initiate recruitment from the benthic habitat during algal blooming in Lake Chongtian in China (Zou *et al.* 2018). Thus, regulation of the *Microcystis* population is very important for monitoring HAB in aquatic ecosystems. Unfortunately, microcystin-producing gene products (Mcy) were not found in our current analysis. However, several metabolic enzymes linked to algal bloom were detected in the HAB metaproteome in the Daechung reservoir. For example, glutamate-ammonia ligase, called glutamine synthetase (GS), was identified in the Daechung HAB. GS is an essential enzyme for nitrogen metabolism in NIES-843 and *Gloeobacter violaceus* PCC7421 (Supplementary Table 2). GS has been used as an indicator of nitrogen-replete algal bloom in the sub-tropical coastal waters of Key West, Florida (Hoch *et al.* 2008). In addition to boosting the profiles of metagenomes, predictions of metabolic potential to influence HABs will be crucial for understanding the aquatic cyanobacterial community (Zhang *et al.* 2018).

5. Interpretation of HAB cyanobacteria and perspectives of metaproteomics

The protein components belonging to the energy category were assigned to a wide range of cyanobacterial species, such as NIES-843, PCC7120, UCYN-A, PCC90,

PCC6304, PCC7112, PCC10605, PCC7202, PCC7122, PCC7421, PCC6803, and BP-1. The second major category of proteins in the HAB metaproteome was metabolism (15%) and was primarily related to carbon metabolism. The third largest category was translation (12%), consisting of 50S and 30S ribosomal proteins. This suggests that the Daechung HAB possesses very active and viable cyanobacterial protein synthesis. The top ten most abundant proteins in the HAB of Lake Daechung comprised 33.3 mol% out of the 194 proteins identified by two-step analysis (Table 2). These proteins mainly belonged to the following categories: energy, protein folding, and cell structure. Due to high complexity in the microbial community, metaproteome data analysis requires greater computing power and more efficient algorithms than other analyses, i.e., the discrimination of homologous proteins caused by redundant protein identification (Herbst *et al.* 2016). Compared to the approximately one trillion estimated species on Earth, the currently available UniProt/TrEMBL database contains 1,961,734 metagenomes (www.ebi.ac.uk/uniprot/TrEMBL_status, 06.2019 version), covering only 0.0002% of the estimated species. Thus, such an approach for exact protein identification produces great ambiguity. Recently, publicly available meta-omics datasets have been integrated in order to assemble uncertainties and genomic variants so that we can utilize metagenomes and/or metatranscriptomes for metaproteomic interpretation (Li *et al.* 2019). Our present study describes the early stages of a straightforward metaproteomic analysis of HAB. However, systematic metaproteomics requires access to gigantic genomic/transcriptomic databases for greater protein identification. More practical metaproteomic analysis methods will be further developed in the integrated search database platforms linked to functional meta-omics. Our attempts to conduct an HAB metaproteomic analysis will be a good reference for monitoring ecological variation of aquatic microalgae in the detrimental microbial community at a meta-protein level.

CONCLUSIONS

In the present study, a gel-based shotgun proteomics method was used to analyze the metaproteome of the microbial community comprising harmful algal bloom (HAB) in Daechung reservoir, Korea. Also, microscopic observation of HAB samples showed red signals, presumably caused by the autofluorescence of chlorophyll and phycocyanin in viable cyanobacteria. Proteomic analysis performed by two-step analysis (FSS and RSS) showed that this analysis of the metaproteome was 1.8 times higher than that of the one-step analysis (FSS only). As a result of the analysis, 12 species (99 mol%) of cyanobacteria and 1 species (1 mol%) of green algae were found, and the most dominant species was *Microcystis aeruginosa* NIES-843 (62.3%), which produce microcystin. These results will propose a better denotation or monitoring ecological variation on the point of a meta-protein level of aquatic microalgae for understanding HAB.

ACKNOWLEDGEMENTS

This research was a part of the project titled 'The foundation on the commercial natural products using functional sulfated polysaccharides derived from marine microalga *Porphyridium* sp. (20190013)', funded by the Ministry of Oceans and Fisheries, Korea. This research was supported by the Korea university research grant. The present study was also supported by the Korea Basic Science Institute research grant (C39712) awarded to J.S. Choi. We would like to thank former KBSI members Dr. D.G. Lee and Dr. C.W. Choi for technical help.

REFERENCES

Blaha L, P Babica and B Marsalek. 2009. Toxins produced in cyanobacterial water blooms - toxicity and risks. Interdiscip. Toxicol. 2:36–41.

Cargile BJ, JL Bundy and JL Stephenson. 2004. Potential for false positive identifications from large databases through tandem mass spectrometry. J. Proteome Res. 3:1082–1085.

Giovannoni SJ, L Bibbs, JC Cho, MD Stapels, R Desiderio, KL Vergin, MS Rappe, S Laney, LJ Wilhelm, HJ Tripp, EJ Mathur and DF Barofsky. 2005. Proteorhodopsin in the ubiquitous marine bacterium SAR11. Nature 438:82–85.

Hanson BT, I Hewson and EL Madsen. 2014. Metaproteomic survey of six aquatic habitats: discovering the identities of microbial populations active in biogeochemical cycling. Microb. Ecol. 67:520–539.

Heisler J, PM Gilbert, JM Burkholder, DM Anderson, W Cochlan, WC Dennison, Q Dortch, CJ Gobler, CA Heil, E Humphries, A Lewitus, R Magnien, HG Marshallm, K Sellner, DA Stockwell, DK Stoecker and M Suddleson. 2008. Eutrophication and harmful algal blooms: A scientific consensus. Harmful Algae 8:3–13.

Herbst FA, V Lunsmann, H Kjeldal, N Jehnlich, A Tholey, M von Bergen, JL Nielsen, RL Hettich, J Steifert and PH Nielsen. 2016. Enhancing metaproteomics - the value of models and defined environmental microbial systems. Proteomics 16: 783–798.

Hettich RL, C Pan, K Chourey and RJ Giannone. 2013. Metaproteomics: harnessing the power of high performance mass spectrometry to identify the suite of proteins that control metabolic activities in microbial communities. Anal. Chem. 85:4203–4214.

Heyer R, K Schallert, R Zoun, B Becher, G Saake and D Benndorf. 2017. Challenges and perspectives of metaproteomic data analysis. J. Bacteriol. 261:24–36.

Hoch MP, KS Dillon, RB Coffin and LA Cifuentes. 2008. Sensitivity of bacterioplankton nitrogen metabolism to eutrophication in sub-tropical coastal waters of Key West, Florida. Mar. Pollut. Bull. 56:913–926.

Ishihara Y, Y Oda, T Tabata, T Sato, T Nagasu, J Rappsilber and M Mann. 2005. Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequence peptides per protein. Mol. Cell. Proteomics 4:1265–1272.

Jagtap P, J Goslinga, JA Kooren, T McGowan, MS Wroblewski, SL Seymour and TJ Griffin. 2013. A two-step database search method improves sensitivity in peptide sequence matches for metaproteomics and proteogenomics studies. Proteomics 13:1352–1357.

Jagtap P, T McGowan, S Bandhakavi, ZJ Tu, S Seymour, TJ Griffin and JD Rudney. 2012. Deep metaproteomic analysis of human salivary supernatant. Proteomics 12:992–1001.

Ki JS and MS Han. 2007. Nuclear rDNA and chloroplast rbcL, rbcS and IGS sequence data, and their implication from the Japanese, Korean, and North American harmful algae, *Heterosigma akashiwo* (Raphidophyceae). Environ. Res. 103:299–304.

Kim SG, SK Rhee, CY Ahn, SR Ko, GG Choi, JW Bae, YH Park and HM Oh. 2006. Determination of cyanobacterial diversity during algal blooms in Daechung reservoir, Korea, on the basis of cpcBA intergenic spacer region analysis. Appl. Environ. Microbiol. 72:3252–3258.

Lee DG, J Kwon, CY Eom, YM Kang, SW Rho, KB Kee and JS Choi. 2015. Directed analysis of cyanobacterial membrane phosphoproteome using stained phosphoproteins and titanium-enriched phosphopeptides. J. Microbiol. 53:279–287.

Lee S and D Lee. 2018. Improved prediction of harmful algal blooms in four major South Korea's rivers using deep learning models. Int. J. Environ. Res. Pub. Health 15:1322.

Li S, H Tang and Y Ye. 2019. A meta-proteogenomic approach to peptide identification incorporating assembly uncertainty and genomic variation. Mol. Cell. Proteomics 18:S183–192.

McPartlin DA, JH Loftus, AS Crawley, J Silke, CS Murphy and RJ O'Kennedy. 2017. Biosensors for the monitoring of harmful algal bloom. Curr. Opin. Biotechnol. 43:164–169.

Mikhailov A, AS Härmälä-Brasken, J Hellman, J Meriluoto and JE Eriksson. 2003. Identification of ATP-synthase as a novel intracellular target for microcystin-LR. Chem. Biol. Interact. 142:223–137.

Nakao M, S Okamoto, M Kohara, T Fujishiro, T Fujisawa, S Sato, S Tabata, T Kaneko and Y Nakamura. 2010. CyanoBase: the cyanobacteria genome database update 2010. Nucleic Acids Res. 38:D379–D381.

Newell SE, TW Davis, TH Johengen, D Gossiaux, A Burtner, D Palladino and MJ McCarthy. 2019. Reduced forms of nitrogen are a driver of non-nitrogen-fixing harmful cyanobacterial blooms and toxicity in Lake Erie. Harmful Algae 81:86–93.

Oenema O. 2004. Governmental policies and measures regulating nitrogen and phosphorus from animal manure in European agriculture. J. Anim. Sci. 82:E196–E206.

Oh HM, SJ Lee, JH Kim, HS Kim and BD Yoon. 2001. Seasonal variation and indirect monitoring of microcystin concentrations in Daechung reservoir, Korea. Appl. Environ. Microbiol. 67:1484–1489.

Orme-Zavaleta J and WR Munns Jr. 2008. Integrating human and ecological risk assessment: application to the cyanobacterial harmful algal bloom problem. Adv. Exp. Med. Biol. 619:867–883.

Paerl HW. 2018. Mitigating toxic planktonic cyanobacterial blooms in aquatic ecosystems facing increasing anthropogenic and climate pressure. Toxins (Basel) 10:E76.

Peterson GL. 1977. A simplification of the protein assay method of Lowry *et al.* which is more generally applicable. Anal. Biochem. 83:346–356.

Russo DA, N Couto, AP Beckerman and J Pandhal. 2016. A metaproteomic analysis of the response of a freshwater microbial community under nutrient enrichment. Front. Microbiol. 7:1172.

Schulze K, DA Lopez, UM Tilich and M Frohme. 2011. A simple viability analysis of unicellular cyanobacteria using a new autofluorescence assay, automated microscopy, and ImageJ. BMC Biotechnol. 11:118.

Shevchenko G, S Musunuri, M Wetterhall and J Bergquist. 2012. Comparison of extraction methods for the comprehensive analysis of mouse brain proteome using shotgun-based mass spectrometry. J. Proteome Res. 11:2441–2451.

Shin BJ, J Oh, S Kang, YH Chung, YM Park, YH Kim, S Kim, J Bhak and JS Choi. 2008. Cyanobacterial hybrid kinase Sll0043 regulates phototaxis by suppressing pilin and twitching motility protein. J. Microbiol. 46:300–308.

Sowell SM, PE Abraham, M Shah, NC Verberkmoes, DP Smith, DF Barofsky and SJ Giovannoni. 2011. Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. ISME J. 5:856–865.

Steffen MM, Z Li, TC Effler, LJ Hauser, GL Boyer and SW Wilhem. 2012. Comparative metagenomics of toxic freshwater cyanobacteria bloom communities on two continents. PloS One 7:e44002.

Williams TJ, D Wilkins, E Long, F Evans, MZ DeMaere, MJ Raftery and R Cavicchioli. 2013. The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. Environ. Microbiol. 15:1302–1317.

Zhang H, K Wang, L Shen, H Chen, F Hou, X Zhou, D Zhang and X Zhu. 2018. Microbial community dynamics and assembly follow trajectories of an early-spring diatom bloom in a semienclosed bay. Appl. Environ. Microbiol. 84:e01000–e01018.

Zhao L, Y Song, L Li, N Gan, JJ Brand and L Song. 2018. The highly heterogeneous methylated genomes and diverse restriction-modification systems of bloom-forming *Microcystis*. Harmful Algae 75:87–93.

Zou W, Z Wang, Q Song, S Tang and Y Peng. 2018. Recruitment-promoting of dormant *Microcystis aeruginosa* by three benthic bacterial species. Harmful Algae 77:18–28.

