ARTICLE

Archaeal community profiling of a mixed culture from mud and water slurry from a solfataric mudspring as revealed by a community proteome approach

Andrew D. Montecillo*[1,3], Leopold L. Ilag[2], Ajay Kohli[3], Nacita B. Lantican[1], and Asuncion K. Raymundo[1]

[1]Institute of Biological Sciences, College of Arts and Sciences, University of the Philippines Los Baños, Laguna, Philippines
[2]Department of Environment Science and Analytical Chemistry, Stockholm University, Stockholm, Sweden
[3]International Rice Research Institute, Los Baños, Laguna, Philippines

Analysis of microbial community composition is usually done by DNA-based approaches. Here we used a community proteome approach to study the Archaeal community in carboxy-methyl cellulose-enriched mud and water slurry from Mudspring, a solfataric spring of Mount Makiling in the Philippines. The peptides obtained from proteolytic digestion of total protein extracts were analyzed by using liquid chromatography-electrospray ionization coupled with tandem mass spectrometry (nanoLC-ESI-MS/MS). Majority of the proteins sampled and identified appear to be largely from *Sulfolobus* species, a number of which are known to be cellulase-producing that could be actively engaged in cellulose degradation in the setup. Furthermore, proteins with high similarity to those from *Caldivirga maquilingensis*, *Metallosphaera sedula*, *Desulfurococcus*, *Pyrococcus furiosus*, and *Pyrodictium occultum* were identified. Many of these Archaea were not detected in earlier reports of Mudspring microbial community analysis. A feasibly exhaustive knowledge of the Mudspring microbial community using different approaches will foster an understanding of the microbial diversity in such ecological niches and may enable isolation of thermally stable proteins of commercial value in the future. Metaproteomic technique allows functional and taxonomic survey of community structure.

*Corresponding author
Email Address: admontecillo@up.edu.ph
Date received: February 15, 2019
Date revised: June 19, 2019
Date accepted: June 27, 2019

KEYWORDS

hyperthermophile, community proteomics, cellulose, mudspring, *Sulfolobus*

INTRODUCTION

The Mt. Makiling Mudspring, which lies inside the Makiling Forest Reserve, Mt. Makiling, University of the Philippines Los Baños, Laguna, is considered a solfataric, high-temperature, and acidic environment, which may be harboring a wide range of unidentified microorganisms. Its temperature ranges from 70°C to 90°C, pH from 2.2 to 5.1, and its high salinity is due mostly to iron and sulfates (Bundac et al. 1976; Oles and Houben 1998). Limited information exists on the microbial community of this environment due to the extreme conditions. The need to recreate and simulate such conditions for growing, maintaining, and isolating the microorganisms limits the in vitro techniques available for characterizing the microbial community from such extreme and complex environments.

Itoh et al. in 1999 and 2003 were able to isolate and characterize two novel Crenarchaeote species, namely, *Caldivirga maquilingensis* and *Caldisphaera lagunensis*, respectively. Other archaea species such as *Sulfolobus yangmingensis*, *Metallosphaera sedula*, and *Acidilobus aceticus* were also isolated. Sanchez et al. (1998) and recently Lantican et al. (2011) revealed previously uncharacterized bacterial and archaeal members of the microbial community of the Mt. Makiling Mudspring. They used polymerase chain reaction (PCR) amplification of 16S rDNA using universal primers (519F-1392R) and Archaea-specific primers (23FPL-1391R), followed by cloning and sequencing. Phylogenetic analysis of the cloned 16S rDNA showed that most of the clones were related to *Sulfolobus tokodaii*, *S. islandicus*, *S. solfataricus*, *and S. shibatae.*

In order to distinguish our study of the Mudspring microbial community structure from that of Sanchez et al. (1998) and Lantican et al. (2011) and to test the value of a different method in community analysis, we used the mass spectrometry–based proteomics approach. It had not been widely applied in environmental research because the thinking was that free proteins would not be sufficiently stable. However, based on observations that proteins can readily be detected as a component of dissolved organic carbon in water samples from different environments or communities, a number of environmental proteomics approaches have been developed and applied in ecology and environmental research (Schulze 2004; Ram et al. 2005; Benndorf et al. 2007; Wilmes et al. 2008; Sowell et al. 2009). These show that a metaproteomic approach is capable of distinguishing active sets of organisms in a community, that protein presence can correlate to biogeochemical processes and functions, and that source organisms for specific enzymes can be identified. By contrast, DNA-based approaches' main purpose is to infer the taxonomic profile and functional potential of a microbial community. DNA-based approaches could provide only a glimpse into the functional profile of a microbial community.

We undertook this study to identify mudspring microbes by using the community proteomics approach from a cellulose-enriched mixed culture setup to provide a handle on isolating thermostable cellulases in the future. This setup was used in tandem with another study of isolating and cloning cellulases (manuscript in preparation). Cellulose is the most widespread natural and renewable carbon source on earth as the main component of the plant cell wall. It is also present in some algae, fungi, bacteria, and animals (Templeton 2010). Enzymes from hyperthermophilic organisms can offer solutions as they possess inherent resistance to extreme physicochemical conditions (Blumer-Schuette et al. 2008). We therefore attempted to introduce the community proteomics approach in assessing the microbial community from cellulose-enriched Mt. Makiling Mudspring mud and water slurry samples. The main objective of this study was to identify Archaeal community members based on liquid chromatography coupled with tandem mass spectrometry (nanoLC-ESI-MS/MS) analysis of proteins extracted from the cells present in cellulose-enriched Mudspring mud and water slurry samples.

MATERIALS AND METHODS

Sample Collection

Solfataric mud and solfataric hot spring water were collected from the Mt. Makiling Mudspring, an acid solfatara located at the foot of Mt. Makiling, University of the Philippines Los Baños, Los Baños, Laguna, Philippines. Samples were obtained near the center of the mud pool or mud pots at approximately 1-meter depth and were immediately placed in sterile 3-liter collection bottles. The collection bottles were placed in thermal containers during transport and were processed for total protein extraction immediately or stored in a 4°C refrigerator until further use.

Carboxy-methyl cellulose enrichment of the mud and water slurry

Around 3 liters of the original Mudspring mud and water slurry samples contained in a 4-liter screw-cap amber bottle were enriched in Brock's basal salts medium (Brock et al. 1972) with carboxy-methyl cellulose (CMC) as sole carbon source. Setup was incubated at 70°C to 80°C aerobically for up to four weeks. The bottles were shaken at least once a day to allow mixing of the contents.

Extraction of cells from enriched acidic mud and water slurry

The method of cell extraction was adopted from Reigstad et al. (2011). Approximately 200 ml of enriched hot spring slurry were obtained, placed in sterile conical tubes, and mixed by hand to obtain a homogeneous sample. Aliquots were pipetted into sterile 2.0 ml microcentrifuge tubes. Most of the mud and clay minerals were removed by briefly spinning the tubes for 5–10 seconds only at full speed (16,000 x *g*). Most of the clay minerals were expected to be pelleted, and the cells suspended in the supernatant were harvested in another centrifugation step (10 minutes at 14,000 rpm). Very fine-grained clay minerals need to be removed first as they inhibit downstream molecular analysis by binding DNA and proteins, enzymatic inhibition, or clogging the filters and pipette tips. The supernatant was carefully and immediately transferred to fresh sterile 2.0 ml tubes and centrifuged for 10 minutes at 14,000 rpm to harvest the cells. Cell extraction was done several times to concentrate the cell pellet from the 200 ml slurry. The cells were resuspended in 200 µl of 1x Tris EDTA buffer and were processed immediately for protein extraction or stored in a 4°C refrigerator until further use.

Total protein extraction and purification

The protocol of Benndorf et al. (2007) was employed, with certain modifications, in protein extraction and purification. The cell pellet from the samples was incubated with 400 µl of 0.1 M NaOH for 30 minutes at 20°C with gentle shaking. The initial extraction step with 0.1 M NaOH was used to extract humic compounds from soil. In addition to extracting associated humic compounds, 0.1 M NaOH has cell lysis effect for subsequent extraction of protein (Guerlava et al. 1998). The sample was centrifuged (10,000 x *g*, 10 minutes, 4°C), and 400 µl supernatant was mixed with 900 µl Tris-saturated liquid phenol and 400 µl sterile Milli-Q water. The mixture was shaken for 1 hour at 20°C. The phases were then separated by centrifugation for 10 minutes at 16,000 x *g* at room temperature. About 1 ml of the lower phenol phase was collected, and the proteins in it were precipitated overnight with 5-fold volume of 0.1 M ammonium acetate in methanol at –18°C. Afterward the sample was centrifuged for 10 minutes at 16,000 x *g*, at 0°C. The pellet was resuspended in 1 ml 0.1 M ammonium acetate in methanol, incubated for 15 minutes at –18°C and centrifuged again. The pellet was successively washed in 1 ml 0.1 M ammonium acetate in methanol, 1 ml 80% acetone, and 1 ml 70% ethanol. In each washing step, 15-minute incubation and centrifugation (10 minutes, 16,000 x *g*, 0°C) were included. Finally, the pellet was resuspended in 200 µl of Tris-Cl buffer (pH 8.0) and stored at –20°C or was lyophilized.

Estimation of concentration of the total protein extract

The concentration of the total protein extracts was estimated by Bradford assay (Bradford, 1976) with slight modification. A volume of 1.0 ml Bradford reagent (Coomassie Brilliant Blue G-

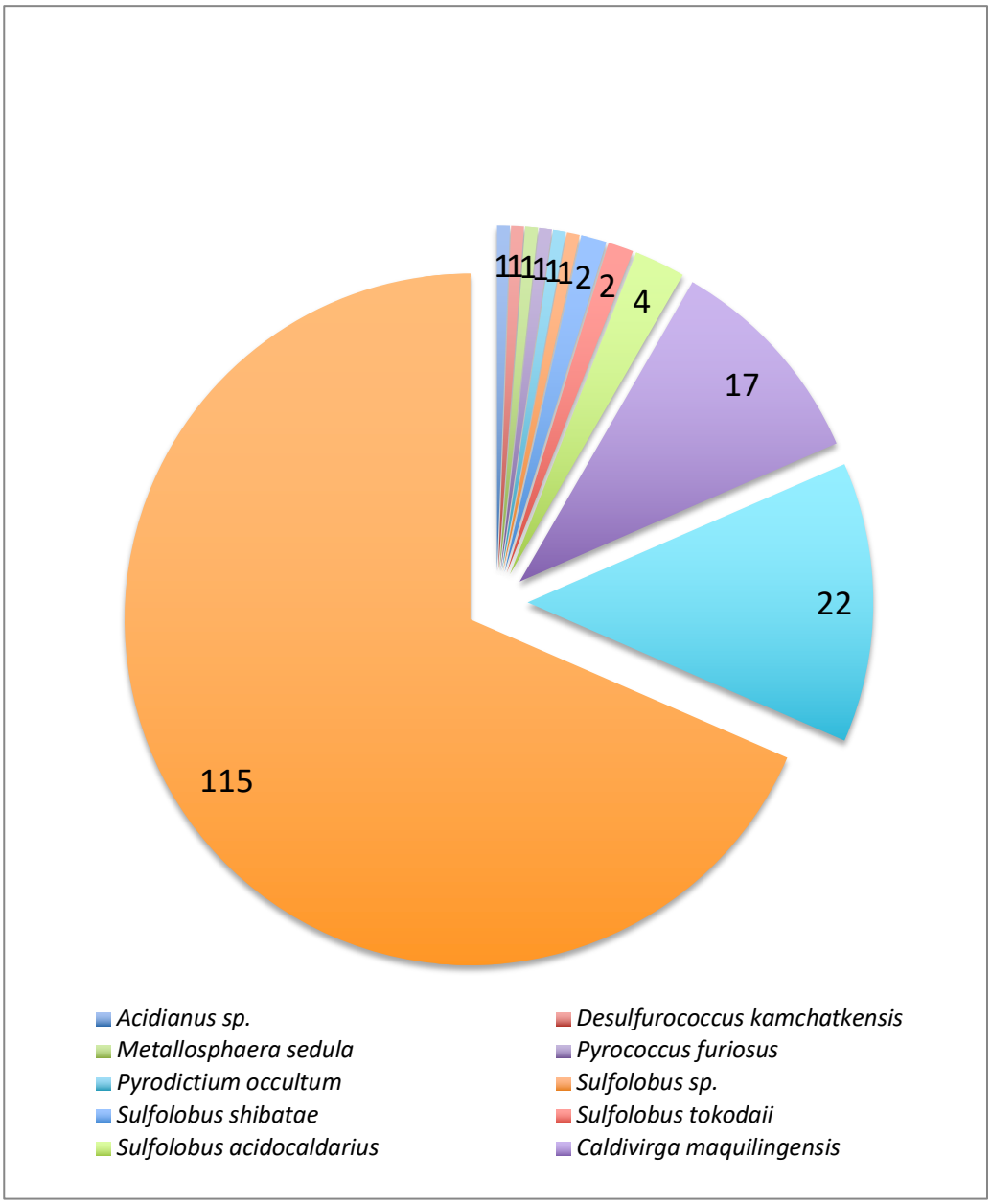

Figure 1: Phylogenetic distribution of proteins identified in cellulose-enriched mud and water slurry samples. Areas of the pie chart represent the number of proteins identified. Total number of proteins identified is 168.

250, 0.01% w/v; phosphoric acid, 8.5% w/v; methanol, 5% w/v) was added to 25 µl protein extract in sterile 1.5 ml microcentrifuge tube. Absorbance at 595 nm was measured after 15 minutes against Tris-Cl buffer (pH 8.0) as blank. The concentration of proteins was determined based on the standard curve previously prepared using varying concentrations of bovine serum albumin.

In-solution enzymatic digestion of proteins

Lyophilized crude protein samples were resuspended in 0.1% ProteaseMax solution (Promega, Wisconsin, USA) with 50 mM ammonium bicarbonate (AmBic), pH 8, and 10% cerium (IV) ammonium nitrate (CAN). The mixture was sonicated for 10 minutes and then centrifuged. The supernatant was collected and transferred to a fresh microcentrifuge tube. The proteins were reduced in 5 mM dithiothreitol (DTT) for 30 minutes at 56°C and alkylated in 15 mM iodacetamide for 30 minutes at room temperature in dark. Tryptic digestion was done overnight with final concentration of trypsin:protein of 1:30. Finally, tryptic digest was acidified with formic acid.

Identification of proteins from cellulose-enriched mud and water slurry samples via LC-MS/MS and database search

Extracted peptides were cleaned with C18 StageTips according to manufacturer description (Thermo Fisher Scientific Inc.). Eluted peptides were dried and resuspended in 3% acetonitrile (ACN) and 0.2 % formic acid. LC-MS/MS analyses were performed on an Easy-nLC system (Thermo Scientific, Bremen, Germany) directly online coupled to LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific). The peptide separation was performed on a 10-cm long fused silica tip column (SilicaTips™ New Objective Inc.) packed in-house with 3 µm C18-AQ ReproSil-Pur® (Dr. Maisch GmbH, Germany). The chromatographic separation was achieved by using an acetonitrile (ACN)/water solvent system containing 0.2% formic acid. The gradient was set up as follows: 3%−48% ACN in 50 minutes, 48%−80% ACN in 3 minutes and 80% ACN for 7 minutes all at a flow rate of 300 nl/minute. The MS acquisition method was comprised of one survey scan ranging from m/z 300 to m/z 2000 with a resolution of R = 60,000 at m/z 400, followed by 10 data-dependent collision-induced dissociation MS/MS scans from the top 10 precursor ions with a charge state ≥ 2.

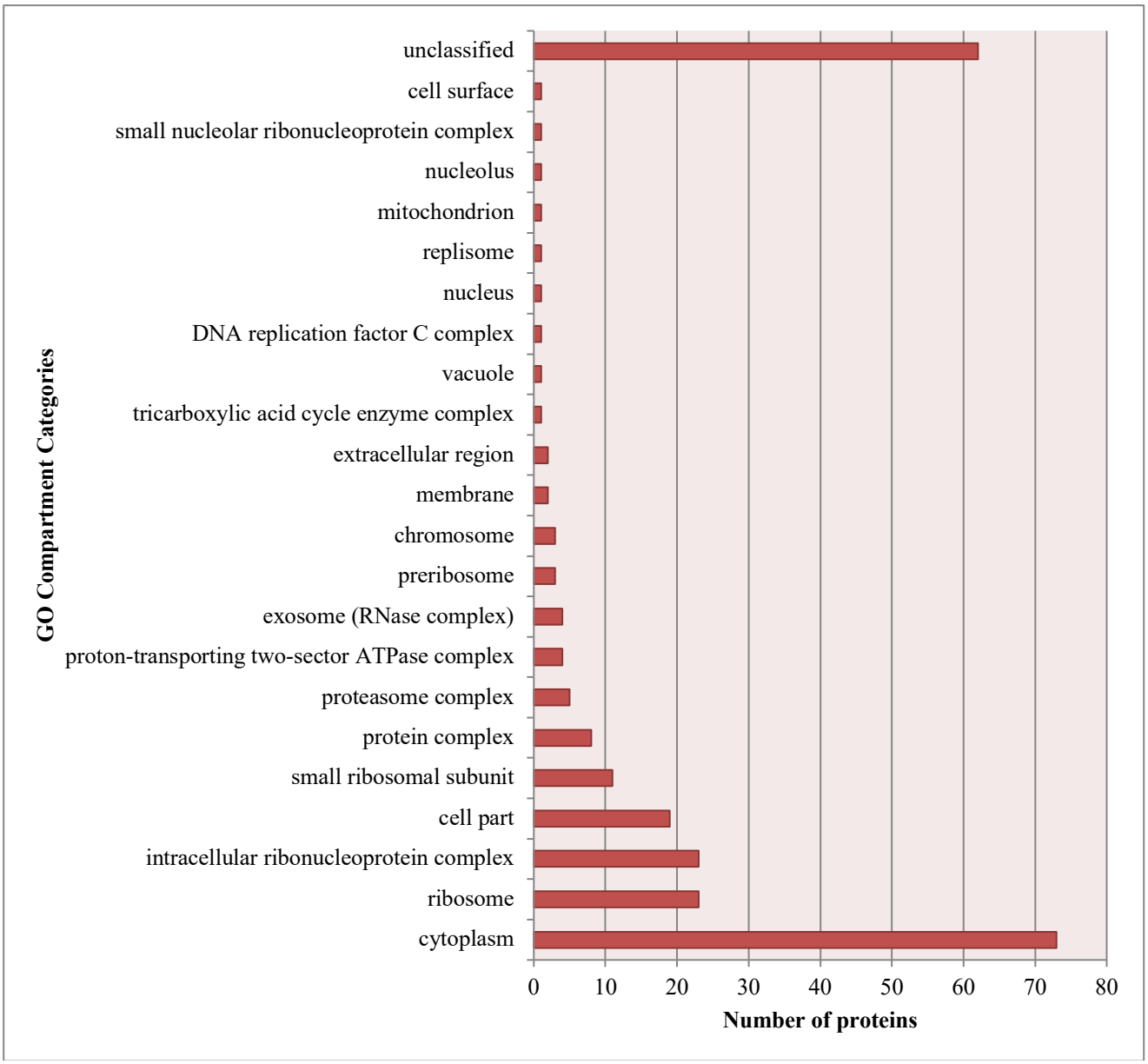

Figure 2: Classification of 168 proteins identified based on gene ontology (GO) compartment categories. The GO numbers were derived from Protein Information Resource (PIR) (http://pir.georgetown.edu/) Batch retrieval tool.

The data were searched against the *Archaea* subset (with 19,385 entries) of the SwissProt database (http://www.uniprot.org/) using Mascot 2.5.1 (Matrix Science Ltd., London, UK) and X! Tandem (The GPM, thegpm.org; version CYCLONE (2010.12.01.1)). Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.050 Da and a parent ion tolerance of 10.0 PPM. Carbamidomethyl of cysteine was specified in Mascot and X! Tandem as a fixed modification. Deamidated of asparagine and glutamine and oxidation of methionine were specified in Mascot as variable modifications. Glu->pyro-Glu of the n-terminus, ammonia-loss of the n-terminus, gln->pyro-Glu of the n-terminus, deamidated of asparagine and glutamine, and oxidation of methionine were specified in X! Tandem as variable modifications.

Scaffold (version Scaffold_4.4.5, Proteome Software Inc., Portland, Oregon) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 86.0% probability to achieve a false discovery rate less than 1.0% by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 99.0% probability to achieve an FDR less than 1.0% and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

RESULTS AND DISCUSSION

Taxonomic assignment of identified proteins from cellulose-enriched mud and water slurry samples

Protein sequence or identity derived from MS/MS spectra of tryptic peptides can provide taxonomic information on the origin of the protein (Schulze 2004). In cases where the sequences obtained from tryptic peptides are unique to a specific group of organisms or even to a single species, recovery of even just one protein is sufficient to provide a strong support on the presence of organisms in a certain environment or to identify a species. The very high mass accuracy data obtained by using a linear ion trap (LTQ)-Orbitrap mass spectrometer allowed confident identification of proteins derived from distinct strains within the

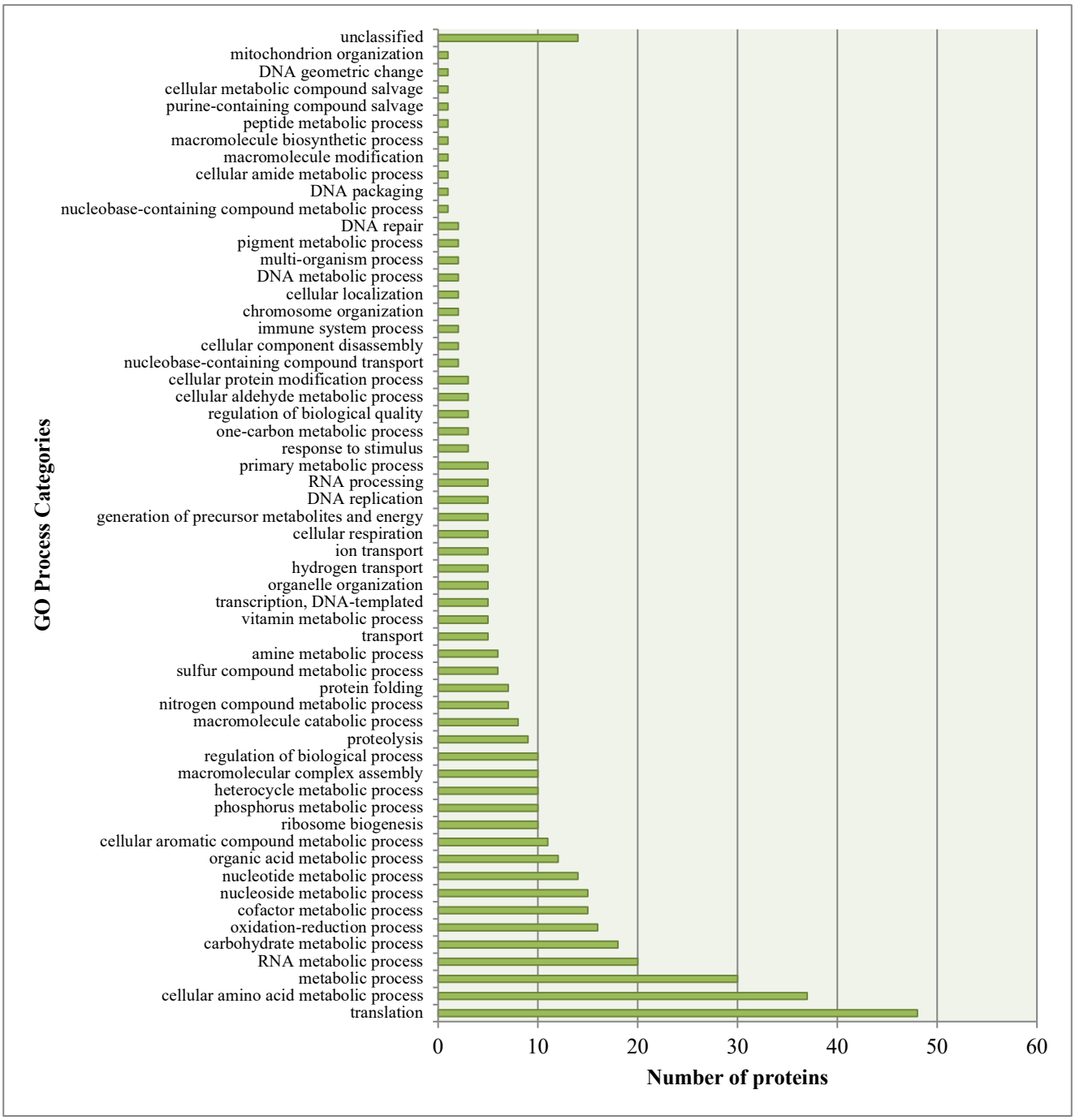

Figure 3: Classification of 168 proteins identified based on gene ontology (GO) process categories. The GO numbers were derived from Protein Information Resource (PIR) (http://pir.georgetown.edu/) Batch retrieval tool.

cellulose-enriched samples. Using the community proteomics method, proteins and peptides from species of *Sulfolobus* such as *S. tokodaii*, *S. acidocaldarius*, *S. Islandicus*, and a certain *Sulfolobus* sp. aside from *S. solfataricus*, *S. shibatae*, were detected, suggesting their presence in the said environment (fig. 1). Furthermore, other species of Archaea such as *Acidianus* sp., *Desulfurococcus kamchatkensis*, *Metallosphaera sedula*, *Pyrococcus furiosus*, and *Pyrodictium occultum*, which were not detected by using the polymerase chain reaction–denaturing gradient gel electrophoresis (PCR-DGGE) method (unpublished data), were detected in the community proteomics approach.

The estimated final concentration of total proteins extracted from the cell pellets after resuspending in 200 µl Tris-Cl buffer from the CMC-enriched Mudspring slurry was about 0.36 mg/ml. Separation of humic compounds from the other components of the mixture was indicated by the brown color in the water phase separated by a thin creamy interphase from the clear light yellow phenol phase, which was presumed to contain the proteins and other components. More rigorous extraction conditions were observed to result in a higher yield for extracted proteins but also coupled with more contaminating humic compounds. From the cellulose-enriched samples, 168 proteins were identified (supplementary table). A majority of the proteins (87%, 146 out of 168) were from *Sulfolobus* (figure 1). The *Sulfolobus* species with the greatest number of proteins detected was *S. solfataricus* (68.5%, 115 of 168). The high number of proteins detected belonging to *Sulfolobus* indicates that this species might have dominated in the cellulose-enriched samples.

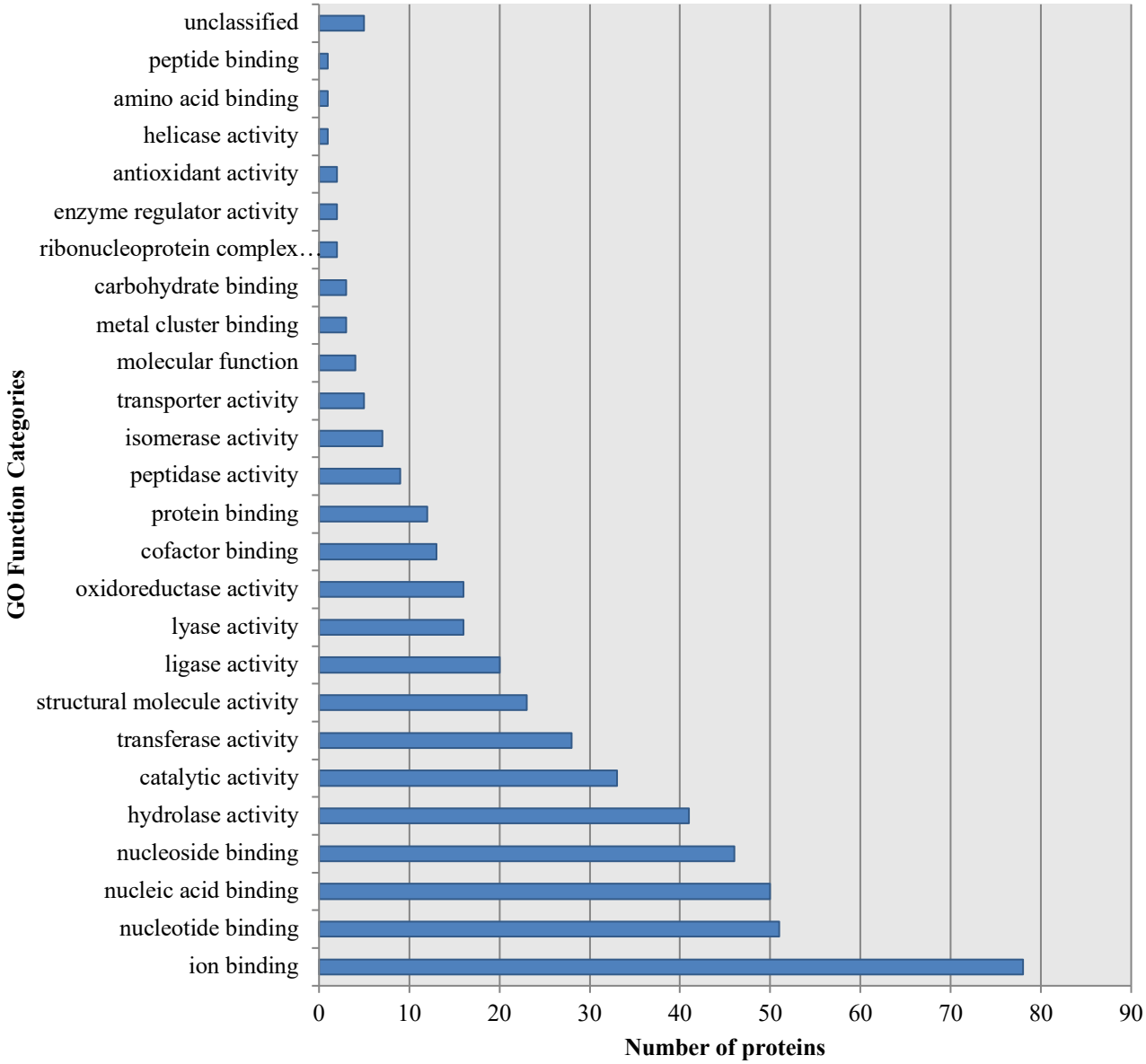

Figure 4: Classification of 168 proteins identified based on gene ontology (GO) function categories. The GO numbers were derived from Protein Information Resource (PIR) (http://pir.georgetown.edu/) Batch retrieval tool.

Most of *Sulfolobus* species are known to produce cellulases, which likely allowed them to exploit the supply of cellulose. However, no cellulase was detected among the identified proteins. Possibly a majority of the cellulases produced by these Archaea species were either secreted in the supernatant during sample preprocessing or were bound to the plasma membrane or anchored to membranes (Yan and Wu 2013; Girfoglio et al. 2012). Our method did not include steps (gradient fractionation and ultracentrifugation) and reagents (such as detergents) that would enrich for membrane fraction, and thus we could have missed out detecting membrane-anchored proteins such as archaeal cellulases. Girfoglio et al. (2012) were able to detect an endoglucanase of *S. solfataricus,* which was found mainly in the plasma membrane fraction. In general, archaeal cellulase genes are said to contain relatively conserved signal sequences followed by a linker region that is rich in serine and threonine residues most likely O-glycosylated in vivo (Brouns et al. 2005). This domain structure and posttranslational modifications are similar to that of S-layer-bound proteins from *Sulfolobus* (Elferink et al. 2001). This suggests that cellulases could be partially bound to the outer surface layer of the cell. By contrast, there are reports of detection of cellulolytic activity in the culture supernatant of *Sulfolobus* (Brouns et al. 2005). However, the extraction method we followed was limited to proteins from cell pellets obtained and concentrated from mud and water slurry. A broader proteomic profile could have been obtained had the secreted proteins in the supernatant and from the membrane fraction been included in the extraction process.

Sulfolobus species (from Sulfolobaceae family belonging to Crenarchaeota phylum) are generally aerobic, but heterotrophic growth has also been reported, including capabilities to oxidize a range of carbohydrates, yeast extract, and peptide mixtures producing carbon dioxide (Grogan 1989; Schonheit and Schafer 1995). *Sulfolobus* became one of model organisms for Archaea, particularly Crenarchaea due to their ability to grow easily aerobically on a wide variety of heterotrophic substrates such as peptides and sugars, in both liquid and solid media (Grogan 1989). They are also abundant globally (DeLong and Pace 2001) and possess mobile genetic elements, such as insertion sequences (IS elements), viruses, small plasmids, and large conjugative plasmids (Zillig et al. 1998). Included in the identified proteins were clustered regularly interspaced short palindromic repeats-associated proteins from *Sulfolobus* (supplementary table). Many of the unusual physiological characteristics of *Sulfolobus* have been investigated including

its thermostable enzymes such as cellulase. Cellulase produced by *S. solfataricus* is suggested to be highly acid-stable and thermostable at the same time (Huang et al. 2005). This species has also been shown to produce xylanases and β-glucosidases/xylosidases involved in the degradation of plant-derived complex polysaccharides in tandem with cellulases. Hence the potential of the *Sulfolobus* species to utilize cellulose could explain their apparent dominance in the CMC-enriched setup. These results showed that the Archaeal community composition detected by the community proteomics approach strongly reflected what was expected after cellulose enrichment. However, since we did not measure cellulase activity in our setups as this was not the goal of the study, we could not directly relate cellulolytic activity to the identified members of the microbial community.

Aside from *Sulfolobus* species, several proteins (17 out of 168) from *Caldivirga maquilingensis* were also detected. *C. maquilingensis*, an anaerobic rod-shaped Crenarchaeote within Thermoproteaceae family with optimum growth temperature at 85°C and at a pH range of 3.7 to 4.2, was first isolated and identified by Itoh et al. (1999) from the Mt. Makiling Mudspring, the same source of our mud and water slurry. It was, however, not detected in the PCR and molecular cloning techniques used by Lantican et al. (2011) and by our PCR-DGGE method (unpublished data). *C. maquilingensis* was predicted to possess an endo-1,4-beta-glucanase B (NCBI GenBank: ABW02444) based on its whole genome sequence, which could have allowed it to proliferate in the cellulose-enriched setup.

Metallosphaera sedula is a Crenarcheote (from Sulfolobaceae family) that has unusually high tolerance to heavy metals (Huber et al. 1989). It can successfully thrive in sulfur-rich hot springs (e.g., Mt. Makiling Mudspring), volcanic fields, and acid mine drainage (AMD) communities (Baker and Banfield 2003). The environments are characterized by high amounts of metal ions, low pH, and high temperature. *M. sedula*, an obligate aerobe that has growth temperature optimum at 75°C and at pH 2.0, is also physiologically versatile capable of heterotrophic, autotrophic, or mixotrophic growth (Auernik and Kelly 2010). This archaeon also has the potential to be used for coal depyritization because of its ability to oxidize pyrite (Clark et al. 1993). *M. sedula* was also successfully isolated and identified by Itoh et al. (2003) from Mt. Makiling Mudspring.

Interestingly, some proteins assigned to those of previously undetected Archaea in the Mt. Makiling Mudspring such as *Acidianus* sp., *Pyrococcus furiosus*, *Desulfurococcus kamchatkensis*, and *Pyrodictium occultum* were also identified. *Acidianus* is a genus within the Sulfolobaceae family whose members are thermoacidophilic archaebacteria. Members of this genus could grow facultatively aerobically by either oxidation or reduction of elemental sulfur (Segerer et al. 1986). Several strains of *Acidianus* have been isolated from solfataric environments similar to those in the Mt. Makiling Mudspring. *Pyrococcus furiosus* is an aquatic anaerobic hyperthermophilic Euryarchaeote (under Thermococcaceae family), which was first isolated in a hydrothermal vent near Vulcano Island, Italy. Its optimal growth temperature is 100ºC (Fiala and Stetter 1986). It is the source of the highly thermostable *Pfu* polymerase and is known to produce an endoglucanase, EgIA, which hydrolyzes beta-1,4 bonds (Bauer et al. 1999). *Desulfurococcus kamchatkensis* is an anaerobic organotrophic hyperthermophilic Crenarchaeon (under Desulfurococcaceae family), which thrives in solfataric hot springs of temperatures up to 87°C and at pH 5.5–7.5 (Kublanov et al. 2009). First discovered in 1979, *Pyrodictium occultum* is a hyperthermophilic Crenarcheaon from Pyrodictiaceae family found in deep-sea hydrothermal vents (Stetter et al. 1984). Its optimum temperature for growth is said to be between 80°C and 105°C. Interestingly, *P. occultum* is known to form a unique cell structure involving a network of cannulae and flat, disk-shaped cells.

The presence of the above-mentioned Archaea, which were previously unreported in the Mt. Makiling Mudspring, needs to be ascertained complementarily through a metagenomics approach. The role of metagenome sequences in profiling the strain composition of microbial consortia cannot be overstated. The metaproteome data are usually searched against metagenomic results serving as reference libraries, and missing genome segments will translate into unidentified peptides in the metaproteome datasets (Keller and Hettich 2009). Protein identification is easier with the availability of extensive metagenomic sequences. Many protein databases were derived from *in silico* predictions of gene products and open reading frames of type strains with sequenced genomes. The depth and quality of the metagenome sequences, however, directly control and somewhat limit the success of the metaproteome investigations (Keller and Hettich 2009). By contrast, the de novo sequencing of peptides from environmental samples offers an alternative tool in identifying proteins without prior complete environmental sequence data available. It identifies peptides by searching against the current databases using the MS basic local alignment search tool (MS-BLAST) algorithm (Shevchenko et al. 2001; Standing 2003). Protein analysis has the advantage of rapid identification of taxonomic units over a broad range of the phylogenetic tree. However, proteomic analysis of a community is not without its bias. Specifically, it may be biased toward emphasis of metabolically active organisms with a high cellular turnover rate (Schulze 2004), resulting in overrepresentation of proteins from dominant active members of the community. In our study, *Sulfolobus* strains appeared to have dominated the setup based on the number of proteins identified from this species.

Gene ontology classification of identified proteins

The identified archaeal proteins were classified by using the gene ontology (GO) (http://www.geneontology.org/) classification for localization (fig. 2), biological process (fig. 3), and molecular function (fig. 4). Molecular function ontology describes the functions or abilities of a certain gene product. By contrast, cellular component ontology describes locations, at the levels of subcellular structures and macromolecular complexes. It includes multi-subunit enzymes and other protein complexes, but not individual proteins or nucleic acids. Biological process ontology includes terms that represent collections of processes as well as terms that represent a specific, entire process (Ashburner et al. 2000).

In terms of localization (fig. 2), most of the proteins are found in the cytoplasm, supporting our hypothesis on why no cellulase was detected due to the biases of the protein extraction method toward cytosolic proteins. Only 1 protein was found on the cell surface, and only 2 out of 168 proteins were membrane-associated. This warrant use of other extraction methods or further protein extraction method optimizations to be able to obtain broader proteomic information (i.e., extracellular and membrane-associated proteins) necessary for a more comprehensive proteomic fingerprinting and profiling of the cellulose-enriched Mt. Makiling Mudspring mud and water slurry.

Several proteins were classified in many GO categories. A majority of the proteins (48 out of 168) identified are involved in protein synthesis (translation) based on biological process classification, such as 30S and 50S ribosomal proteins (fig. 3). Moreover, most proteins are involved in various metabolic processes (e.g., amino acid, nucleotide, nucleoside,

carbohydrate, protein metabolism, etc.). Thermostable proteases (amido- and carboxypeptidases) and proteasomes, which are involved in protein metabolism, are likewise detected.

In terms of molecular functions (fig. 4), proteins that are involved in correct folding, such as thermosomes, were detected in the samples. Thermosomes are chaperonins that are essential multi-subunit assemblies promoting facilitated protein folding in concert with ATP hydrolysis (Ditzel et al. 1998). These proteins, which aid proteins refolding to their native form and in function restoration are known to be produced by *Sulfolobus*. They are linked to survival capacities in extreme environments, such as those in low pH and high temperature. Also, many of the proteins detected have various molecular binding abilities (e.g., ion binding, nucleotide, nucleoside, nucleic acid, amino acid, protein, etc.). Collectively, detection of these proteins suggests an active set of organisms present in the CMC-enriched samples. Specifically, a majority of these proteins were from *S. solfataricus*, which, most probably, is the most active species in the enrichment.

CONCLUSION AND RECOMMENDATIONS

The application of metaproteomics to more complex microbial communities such as those in Mt. Makiling Mudspring still presents a considerable challenge. For complex environments (e.g., seawater, soil, and other water environments) that are of great interest to microbial ecologists, one can expect to resolve only a minute fraction (less than 1%) of the metaproteome with the methods that are currently available (Wilmes and Bond 2006). These limitations necessitate improvements in the protein extraction, separation, and identification technologies. Via database analysis using the level of homology to other species, the proteomics method has the potential to reveal the identity of the active microorganisms in an environment. It can also be used as a decision tool in prioritizing the investigation of elusive functional genes from large metagenomic datasets (Wilmes and Bond 2006).

Our community proteomics approach was able to identify dominant Archaeal community members from the CMC-enriched mud and water slurry fraction based on proteins and peptides analyses. The dominant Archaeal species identified were *Sulfolobus* and *Caldivirga maquilingensis* and other species of Archaea such as *Acidianus* sp., *Desulfurococcus kamchatkensis*, *Metallosphaera sedula*, *Pyrococcus furiosus*, and *Pyrodictium occultum* suggesting their presence in the said environment. This proteomic approach can allow a survey of the functional and taxonomic structure of a microbial community from various sources.

ACKNOWLEDGMENTS

Andrew D. Montecillo was supported by the Philippines' Department of Science and Technology–Science Education Institute through an ASTHRDP (Accelerated Science and Technology Human Resource Development Program) scholarship grant, by the International Rice Research Institute, and by the University of the Philippines Los Baños. We would like to thank Alexey Chernobrovkin of the Karolinska Institutet, Stockholm, Sweden, for the mass spectrometry data preprocessing.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHORS' CONTRIBUTIONS

ADM, LLI, and AK were directly involved in the proteomic experiments, data analysis, and drafting the manuscript. AKR and NBL supervised the experiments and helped draft the manuscript. All authors read and approved the final manuscript.

REFERENCES

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA. Gene ontology: tool for the unification of biology. Nature Genetics 2000; 25(1): 25.

Auernik, KS, Kelly, RM. Physiological versatility of the extremely thermoacidophilic archaeon *Metallosphaera sedula* supported by transcriptomic analysis of heterotrophic, autotrophic, and mixotrophic growth. Appl Environ Microbiol 2010; 76: 931-935.

Baker, BJ, Banfield, JF. Microbial communities in acid mine drainage. FEMS Microbiol Ecol 2003; 44: 139-152.

Bauer, MW, Driskill, LE, Callen, W, Snead, MA, Mathur, EJ, Kelly, RM. An endoglucanase, EglA, from the hyperthermophilic archaeon *Pyrococcus furiosus* hydrolyzes β-1,4 bonds in mixed-linkage (1→3),(1→4)-β-d-glucans and cellulose. J Bacteriol 1999; 181(1): 284-290.

Benndorf, D, Balcke, GU, Harms, H, von Bergen, M. Functional metaproteome analysis of protein extracts from contaminated soil and ground water. ISME J 2007; 1: 224-234.

Blumer-Schuette, SE, Kataeva, I, Westpheling, J, Adams, MW, Kelly, RM. Extremely thermophilic microorganisms for biomass conversion: Status and prospects. Curr Opin Biotechnol 2008; 19: 210-217.

Bradford, MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976; 72: 248-254.

Brock, TD, Brock, KM, Belly, RT, Weiss, RL. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. Archiv fur Mikrobiologie 1972; 84: 54-68.

Brouns, SJJ, Etterna, TJG, Stedman, KM, Walther, J, Smidt, H, Snijders, APL. The hyperthermophilic archaeon *Sulfolobus*: from exploration to exploitation. In Inskeep, W, McDermott, T, eds. Geothermal Biology and Geochemistry in Yellowstone National Park. Bozeman, Montana: Montana State University Publications 2005: 261-276.

Bundac, AE, Carreon, VR, Fernandez, A. Report on the thermal and geochemical survey of the geothermal areas in Los Baños and vicinity. Tech, Rep. Quezon City: Commission on Volcanology 1976.

Clark, TR, Baldi, F, Olson, GJ. Coal depyritization by the thermophilic archaeon *Metallosphaera sedula*. Appl Environ Microbiol 1993; 59: 2375-2379.

DeLong, EF, Pace, NR. Environmental diversity of Bacteria and Archaea. Syst Biol 2001; 50:470-478.

Ditzel, L, We, JL, Stock, D, Stetter, KO, Huber, H, Huber, R, Steinbacher, S. Crystal structure of the thermosome, the archaeal chaperonin and homolog of CCT. Cell 1998; 93:125-138.

Elferink, MG, Albers, SV, Konings, WN, Driessen, AJ. Sugar transport in *Sulfolobus solfataricus* is mediated by two families of binding protein-dependent ABC transporters. Mol Microbiol 2001; 39:1494-1503.

Fiala, G, Stetter, KO. *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100°C. Arch Microbiol 1986; 145: 56-61.

Girfoglio, M, Rossi, M, Cannio, R. Cellulose degradation by *Sulfolobus solfataricus* requires a cell-anchored endo-B-1-4-glucanase. J Bacteriol 2012; 194:5091-5100.

Grogan, DW. Phenotypic characterization of the archaebacterial genus *Sulfolobus*: comparison of five wild-type strains. J Bacteriol 1989; 171:6710-6719.

Guerlava, P, Izac, V, Tholozan, JL. Comparison of different methods of cell lysis and protein measurements in *Clostridium perfringens*: Application to the cell volume determination. Curr Microbiol 1998; 36:131-135.

Hiras J, Wu YW, Deng K, Nicora CD, Aldrich JT, Frey D, Kolinko S, Robinson EW, Jacobs JM, Adams PD, Northen TR. Comparative community proteomics demonstrates the unexpected importance of actinobacterial glycoside hydrolase family 12 protein for crystalline cellulose hydrolysis. Mbio. 2016 Sep 7;7(4):e01106-16.

Huang, Y, Krauss G, Cottaz, S, Driguez H, Lipps G. A highly acid-stable and thermostable endo-beta-glucanase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. Biochem J 2005; 385:581-588.

Huber, G, Spinnler, C, Gambacorta, A, Stetter, KO. *Metallosphaera sedula* gen. nov. and sp. nov. represents a new genus of aerobic, metal-mobilizing, thermoacidophilic archaebacteria. Syst Appl Microbiol 1989; 12:38-47.

Itoh, T, Suzuki, K, Sanchez, PC, Nakase, T. *Caldivirga maquilingensis* gen nov-, sp. nov., a new genus of rod-shape crenarchaeote isolated from hot spring in the Philippines. Int J Sys Bacteriol 1999; 49:1157-1163.

Itoh, T, Suzuki, K, Sanchez, PC, Nakase, T. *Caldisphaera lagunensis* gen nov., sp. nov., a novel thermoacidophilic crenarchaeote isolated from a hot spring at Mt. Maquiling, Philippines. Int J Sys Evol Microbiol 2003; 53:1149-1154.

Keller, M, Hettich, R. Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. Microbiol Mol Biol Rev 2009; 73(1):62-70.

Kublanov, IV, Bidjieva, SK, Mardanov, AV, Bonch-Osmolovskaya, EA. *Desulfurococcus kamchatkensis* sp. nov., a novel hyperthermophilic protein-degrading archaeon isolated from a Kamchatka hot spring. Int J Syst Evol Microbiol 2009; 59:1743-1747.

Lantican, NB, Diaz, MGQ, Cantera, JJL, de los Reyes III, FL, Raymundo, AK. Microbial community of a volcanic mudspring in the Philippines as revealed by 16S rDNA sequence analysis and flourescence in situ hybridization. World J Microbiol Biotechnol 2011; 27:859-867.

Nesvizhskii, AI, Keller, A, Kolker, E, Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003; 75(17):4646–4658.

Oles, D, Houben, G. Greigite (Fe_3S_4) in an acid mudpool at Makiling volcano, the Philippines. J Southeast Asian Earth Sci 1998; 16:513-517.

Ram, RJ, Verberkmoes, NC, Thelen, MP, Tyson, GW, Baker, BJ, Blake RC, Shah M, Hettich RL, Banfield JF. Community proteomics of a natural microbial biofilm. Science 2005; 308:1915-1920.

Reigstad, LJ, Bartossek, R, Schleper, C. Preparation of high-molecular weight DNA and metagenomic libraries from soils and hot springs. In: Klotz, MG, Stein, LY, eds. Methods in Enzymology: Research on Nitrification and Related Processes. London, UK: Academic Press, 2011: B319-344.

Sanchez, PC, Nakase, T, Patricio, VV. Isolation of protease-producing thermophiles in mudspring of Mt. Makiling, Philippines. Proc. International Conference on Asian Network on Microbial Researches 1998.

Schonheit P, Shafer, T. Metabolism of hyperthermophiles. World J Microbiol Biotechnol 1995; 11(1):26-57.

Schulze, WX. Environmental proteomics—what proteins from soil and surface water can tell us: A perspective. Biogeosciences Discussions 2004; 1:195-218.

Segerer, A, Neuner, AM, Kristjansson, JK, Stetter, KO. *Acidianus infernus* gen. nov., sp. nov., and *Acidianus brierleyi* comb. nov.: facultatively aerobic, extremely acidophilic thermophilic sulfur-metabolizing archaebacteria. Int J Syst Bacteriol 1986; 36(4):559–564.

Shevchenko, A, Sunyaev, S, Loboda, A, Shevchenko, A, Bork, P, Ens, W, Standing, KG. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. Anal Chem 2001; 73:1917-1926.

Sowell, SM, Wilhelm, LJ, Norbeck, AD, Lipton, MS, Nicora, CD, Barofsky, CA, Carlson CA, Smith RD, Giovanonni SJ. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. ISME J 2009; 3:93-105.

Standing, KG. Peptide and protein de novo sequencing by mass spectrometry. Curr Opin Struct Biol 2003; 13(5):595-601.

Stetter, KO, Konig, H, Stackebrandt, E. *Pyrodictium* gen. nov., a new genus of submarine disc-shaped sulphure reducing archaebacteria growing optimally at 105°C. Syst Appl Microbiol 1984; 4:535-551.

Templeton, DW. Compositional analysis of lignocellulosic feedstocks: 2. Method uncertainties. J Agric Food Chem 2010; 58:9054-9062.

Wilmes P, Bond PL. Metaproteomics: studying functional gene expression in microbial ecosystems. Trends Microbiol 2006; 14(2):92-7.

Wilmes, P, Wexler, M, Bond, PL. Metaproteomics provides functional insight into activated sludge wastewater treatment. PLoS ONE 2008; 3(3):e1778.

Yan, S, Wu, G. Secretory pathway of cellulase: a mini-review. Biotech. Biofuels 2013; 6:177.

Zillig, W, Arnold, HP, Holz, I, Prangishvili, D, Schweier, A, Stedman, K, She Q, Phan H, Garrett R, Kristjansson JK. Genetic elements in the extremely thermophilic archaeon *Sulfolobus*. Extremophiles 1998; 2:131-140.