Biotechnology Advances 33 (2015) 633–647

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Thermophiles in the genomic era: Biodiversity, science, and applications

M. Sofía Urbieta [a], Edgardo R. Donati [a], Kok-Gan Chan [b], Saleha Shahar [c], Lee Li Sin [c], Kian Mau Goh [c,*]

[a] CINDEFI (CCT La Plata-CONICET, UNLP), Facultad de Ciencias Exactas, Universidad Nacional de La Plata, Calle 47 y 115, 1900 La Plata, Argentina
[b] Division of Genetics and Molecular Biology, Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia
[c] Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Malaysia

ARTICLE INFO

Article history:
Received 17 September 2014
Received in revised form 18 December 2014
Accepted 14 April 2015
Available online 22 April 2015

Keywords:
Archaea
Biofuels
Bioremediation
Biomining
Extremophile
Hot spring
Microbiome
Metagenome
Thermophilic virus
Thermozyme

ABSTRACT

Thermophiles and hyperthermophiles are present in various regions of the Earth, including volcanic environments, hot springs, mud pots, fumaroles, geysers, coastal thermal springs, and even deep-sea hydrothermal vents. They are also found in man-made environments, such as heated compost facilities, reactors, and spray dryers. Thermophiles, hyperthermophiles, and their bioproducts facilitate various industrial, agricultural, and medicinal applications and offer potential solutions to environmental damages and the demand for biofuels. Intensified efforts to sequence the entire genome of hyperthermophiles and thermophiles are increasing rapidly, as evidenced by the fact that over 120 complete genome sequences of the hyperthermophiles Aquificae, Thermotogae, Crenarchaeota, and Euryarchaeota are now available. In this review, we summarise the major current applications of thermophiles and thermozymes. In addition, emphasis is placed on recent progress in understanding the biodiversity, genomes, transcriptomes, metagenomes, and single-cell sequencing of thermophiles in the genomic era.

© 2015 Elsevier Inc. All rights reserved.

Contents

1. Introduction . . . 634
2. Diversity analysis of thermophile populations . . . 635
   2.1. Culture-dependent studies . . . 635
   2.2. Traditional culture-independent studies . . . 635
   2.3. Partial 16S rRNA gene sequencing and culture-independent studies . . . 636
   2.4. Shotgun metagenome and culture-independent study . . . 636
3. Genomes of prokaryotic thermophiles . . . 638
   3.1. Life under extreme conditions is made possible by special genome features . . . 638
   3.2. Horizontal gene transfer and symbiosis between thermophiles . . . 639
4. Applications of thermophiles and thermozymes . . . 639
   4.1. Bioenergy – biofuels . . . 639
   4.2. Biomining . . . 640
   4.3. Bioremediation . . . 641
   4.4. Thermozymes . . . 641
   4.5. Biosurfactants . . . 641
5. Future prospects on the studies of thermophiles . . . 642
   5.1. Exploring the unculturable thermophiles by culture medium reengineering and single-cell sequencing . . . 642
   5.2. Functional metagenomics for uncultured microorganisms and new applications . . . 643
   5.3. Advanced sequencing and 'omics' techniques . . . 644
   5.4. Discovering archaeoviruses and bacteriophages . . . 644

* Corresponding author.
E-mail address: gohkianmau@utm.my (K.M. Goh).

http://dx.doi.org/10.1016/j.biotechadv.2015.04.007
0734-9750/© 2015 Elsevier Inc. All rights reserved.

6. Conclusions . . . 644
Acknowledgements . . . 645
References . . . 645

1. Introduction

Thermophiles and hyperthermophiles dominate heated environments. The optimum growth temperature (OGT) for thermophiles is generally >55 °C, whereas that for hyperthermophiles is >80 °C. Hot springs are one of the main sites where hyperthermophiles and thermophiles are isolated, although they can also thrive in man-made environments, such as the compost facilities (Rastogi et al., 2010). Fig. 1 shows images of several types of hot springs and biomats taken from around the world. Hot springs adjacent to volcanic environments are usually acidic (Urbieta et al., 2014b), but the pH is neutral or slightly alkaline in regions near limestone. Thermophiles may also live under harsh conditions involving extreme pH or high salt concentrations (Futterer et al., 2004; Giaveno et al., 2013; Ruepp et al., 2000; Urbieta et al., 2014a). Most early biodiversity studies adopted culture-dependent approaches. Findings from these studies demonstrated that typically only 1–10% of the total population from any biosphere is cultivable. Because of this limitation, microbiologists later chose culture-independent approaches involving the direct

Fig. 1. Photographs of hot springs in the form of a pool or basin (A–C), a stream (D), heated mud (E–F), and a pot (G). Various colours of biomats and sediments that can form in hot springs (H–L).

amplification of nearly complete sequences of the 16S rRNA gene from bulk genomes, thereby bypassing the culture and isolation steps. Many examples of biodiversity analysis can be found in the literature, some of which are listed in Table 1.

Thermophilic microorganisms have attracted much interest in the realm of biotechnology, especially in relation to industrial processes (i.e. 'white biotechnology') (Bergquist et al., 2014; Elleuche et al., 2014). White biotechnology can be defined as the use of organisms and enzymes for industrial processing and production of materials, chemicals, and energy. Thermophiles can be used in leaching processes and the removal of heavy metals from waste (Ilyas, 2014). Thermophiles also produce important industrial enzymes. In addition, thermophiles may be used directly and indirectly in renewable energy production (Bhalla et al., 2013; Bhandiwad et al., 2013; Goh et al., 2013; McClendon et al., 2012; Parmar et al., 2011). The purpose of this review is to provide an update regarding recent applications with thermophiles. Progress in understanding the biodiversity, genomes, and transcriptomes of thermophilic prokaryotes is also discussed. This review also describes how advanced genome sequencing technologies can help to broaden our understanding of uncultured thermophiles and viruses and to discover novel applications and genes from thermophiles. This review does not discuss thermophilic eukaryotes or metatranscriptomic aspects of thermophiles or hyperthermophiles, although readers can refer to Murakami et al. (2012) and Liu et al. (2011) for insights into these topics.

2. Diversity analysis of thermophile populations

2.1. Culture-dependent studies

Traditional microorganism culture methods have been important in characterizing the biochemical and physiological properties of pure cultures, or in developing whole-cell applications. However, the numbers of colonies formed on plates are far below the number of cells that can be visualised by microscopy. Furthermore, easily cultivable thermophiles may not represent those that dominate natural microbiota. In our laboratory, *Anoxybacillus*, *Geobacillus*, and *Meiothermus* have been the genera most commonly isolated from Malaysian hot springs; however, these genera in fact represent the minority of those present. Replacing agar with gelling agents such as Gelrite and nanofibrous cellulose can potentially increase the biodiversity of isolates capable of forming colonies on solid medium, especially with thermophiles (Tsudome et al., 2009). Common manipulations performed to enhance growth conditions, such as modifying media nutrient compositions, pH, incubation temperatures, or the levels of certain gases may increase the number of different isolates obtained; however, fastidious thermophiles will likely be missed. The causes of 'unculturability' have been addressed in several reviews (Pham and Kim, 2012; Stewart, 2012; Vartoukian et al., 2010). Several factors can impede the growth of thermophiles in culture, including a naturally slow growth rate, inhibition caused by antibacterial substances produced by other cells in the population, inhibition caused by growth media components, inhibition caused by agar, a lack of quorum-sensing or beneficial interactions and signals, stringent chemical requirements, or an over-concentration of the supplied nutrients. Certain thermophiles utilise carbon monoxide (CO) (Techtmann et al., 2009), which is infrequently used in laboratory growth conditions because most laboratories are not equipped with a CO supply due to its toxicity to humans. In addition, most hyperthermophiles grow under obligate anaerobic or facultative anaerobic conditions. Culturing anaerobic microorganisms is comparatively more challenging than aerobic microorganisms. *Candidatus* bacteria have yet to be isolated and grown in laboratories. Costas et al. (2012) elucidated several reasons why *Candidatus* bacteria are difficult to cultivate in the laboratory, including the requirement for two other helper strains in addition to its slow doubling time and stringent growth conditions, such as high partial pressure of CO_2.

2.2. Traditional culture-independent studies

Compared to culture-dependent studies, culture-independent studies involving 18S or 16S rRNA gene sequencing represent a more inclusive approach for biodiversity analysis. In culture-independent studies, nearly complete 16S rRNA genes from bulk genomic DNA are amplified, cloned into cloning vectors and transformed into suitable cloning hosts, usually *Escherichia coli*. Subsequently, hundreds or thousands of recombinant clones are then manually differentiated, often by fingerprinting approaches, such as restriction fragment length polymorphism (RFLP) analysis. The process is laborious if performed manually, and the percentage of unique band patterns with RFLP is relatively low. Therefore, it is advantageous to begin with as large a library as possible to obtain a representative overview of the population. However, 16S rRNA gene amplification is biased due to limitations of primers and DNA polymerases, or PCR artefacts (Hongoh et al., 2003), and amplification may be impeded by inhibitors present in hot springs or soils, such as humic acid. In addition, the so-called universal primers are unable to recognize 16S rRNA genes from all genera and thus may not amplify certain genera. In addition, low copy number genes are frequently missed. Some of these limitations can often be overcome using miniprimer PCR and an engineered *Taq* DNA polymerase (Goh et al., 2011; Isenbarger et al., 2008). Using the traditional method of obtaining nearly complete 16S rRNA gene sequences, many unculturable candidate divisions have been discovered.

Table 1
Examples of hot-spring biodiversity using targeted metagenome sequencing.

Hot spring name (Country)	pH	Temp. (°C)	References
Domas and Cibuni (West Java, Indonesia)	2	82–90	Baker et al. (2001)
Boiling Springs Lake (northern California, USA)	2	55	Wilson et al. (2008)
Three hot springs located in Chiang Mai Province (Thailand)	7.7, 8.5, 9	75, 85, 89	Purcell et al. (2007)
Kirishima Natural Park (Japan)	2.0–2.6	66–93	Satoh et al. (2013)
Rehai and Ruidian geothermal fields, located in Tengchong County (Yunnan Province, China)	2.5–9.4	55.1-- 93.6	Hou et al. (2013)
Ulu Slim hot spring (Perak, Malaysia)	7.8	98	(Goh et al., 2011)
Three Heart Lake Geyser Basin (Mount Sheridanin, YNP, USA	8.1–8.6	44, 63, 75	Bowen De León et al. (2013)
Obsidian Pool, Sylvan Spring, and Bison Pool (YNP, USA)	6.5, 5.5, 8.1	80	Meyer-Dombard et al. (2005)
Jim's Black Pool (Mud Volcano, YNP, USA)	6.7	74	Barns et al. (1994)
Crater Hills-Alice Spring, Norris Geyser Basin- Beowulf Spring; Joseph's Coat Hot Springs-Scorodite Spring; Mammoth Hot Springs-Narrow Gauge, Calcite Springs-Scary Spring (YNP, USA)	2.5–7.8	>65	Inskeep et al. (2013); Inskeep et al. (2010)
Hot spring Mutnovsky Volcano (Kamchatka Peninsul, Russia)	3.5–4	70	Wemheuer et al. (2013)
Bison Pool (Lower Geyser Basin, YNP, USA)	7.3–7.8	40	Swingley et al. (2012)
Kapkai and Waramung springs on Ambitle Island (Papua New Guinea)	9.5	>90	Millard et al. (2014)
Three springs located within the Long Valley Caldera (Mammoth Lakes, CA, USA)	7	80	Vick et al. (2010)
Three hydrothermal ponds in Copahue (Argentina)	0.2–1.1	87	Urbieta et al. (2014b)

2.3. Partial 16S rRNA gene sequencing and culture-independent studies

Next-generation sequencing (NGS), also known as second-generation sequencing, has improved the turnaround time of 16S rRNA-gene based biodiversity studies as more samples can be analysed at a lower sequencing cost (Fig. 2). As the read length of most NGS sequencer platforms is relatively short, partial 16S rRNA gene sequences is often used for the analysis. Most 16S rRNA genes consist of nine variable regions (V1-V9) and nine conserved regions (C1-C9). Degenerate or non-degenerate primers are designed based on the selected conserved regions, which collectively are also known as the targeted metagenome.

However, there are some indications that analysing partial sequences may also generate biased results in biodiversity studies, compared to complete 16S rRNA gene analysis. In an assessment of a sludge metagenome, a primer set directed against the V3 and V4 regions was recommended after the V1-V2 primer set was found to significantly underestimate the diversity of the metagenome (Cai et al., 2013). In a separate experiment, the same laboratory later suggested that the V1 and V2 variable regions might actually be better choices and that the V7, V8, and V9 were the worst regions (Guo et al., 2013). The exact variable regions suitable for studying thermophilic microbiota populations require intensive study, but *in silico* analysis conducted using the SILVA 16S/18S rRNA non-redundant reference database may provide some insights (Klindworth et al., 2012). The authors of this study showed that no single primer pair could perfectly amplify all prokaryotic phyla or genera 16S rRNA genes. In addition, shorter read lengths and sequencing incomplete 16S rRNA gene sequences has the unfortunate consequence that many unique and novel environmental microorganisms are likely to be missed.

The approach of analysing partial 16S rRNA-based metagenome sequences is currently utilised globally, not only for study microbial diversity in hot springs (http://www.ncbi.nlm.nih.gov/pubmed/25798135) (Inskeep et al., 2013) but also with coastal environments (http://journal.frontiersin.org/article/10.3389/fmicb.2015.00177/abstract) (Somboonna et al., 2012), soil samples (Fierer et al., 2012), municipal wastewater treatment plants (Cai et al., 2013), tongue-coating microbiomes (Jiang et al., 2012), and the mouse gut (Lee et al., 2010). This approach has attracted the attention of ecologists, microbiologists, and bioinformaticians alike, with the bioinformaticians actively developing various user-friendly graphical interfaces to permit easy exploration and data display. Such interfaces include VITCOMIC (Mori et al., 2010), Krona (Ondov et al., 2011), and the Windows-based CLcommunity (Table 2). Estimation of bacterial assemblage in an environmental sample could be inaccurate due to the copy number of these genes varies across the taxonomy tree of life. However, there are good indications that computation simulations are capable of providing insight and enhanced analysis (Kembel et al., 2012).

2.4. Shotgun metagenome and culture-independent study

Phylogenetic analysis of complete or partial 16S rRNA gene sequencing can only answer the question 'who are they?' for organisms in any given location, and such results may be unsatisfactory for inquisitive researchers. Targeted metagenomics (i.e. 16S rRNA biomarker) lacks comprehensive ecological information and cannot answer questions such as 'what are they doing?' or 'how they are interacting'? For answers to these questions, a separate metagenomics approach, known as shotgun metagenome sequencing, can provide a comprehensive ecological view of a given site. To illustrate, with the HiSeq® SBS Kit v4 and the HiSeq Cluster Kit v4, the Illumina HiSeq2500 sequencer is able to generate up to 1 terabyte of data in reads of 2×125 base pairs (bp) using two flow cells in less than 6 days. Shotgun metagenome sequencing data enable metabolic reconstruction that is not possible by 16S rRNA-amplicon NGS (Fig. 2). Bioinformatics interfaces play crucial roles in the analysis. The presence of functional gene sequences in biological systems may then be analysed by KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Clusters of Orthologous Groups), or SEED, where identifying pathogenic or virulence genes is possible using the SEED. Such studies can be found in Inskeep et al. (2010) and Jiménez et al. (2012).

Two drawbacks of shotgun metagenome sequencing are its relatively expensive setup cost and the necessary demands for very high computing power for data storage, retrieval, and analysis. The binning and assembly of reads are very challenging and requires extensive computing costs. Currently, only a handful of programs can handle sufficient amounts of data, among which are the MetaCluster (Wang et al., 2012b), IDBA-UD (Peng et al., 2012), metAMOS (Treangen et al., 2013), MetaVelvet (Namiki et al., 2012), Meta-IDBA (Peng et al., 2011), and SOAPdenovo (Xie et al., 2014) programs. The results can be displayed using programs like MEGAN (Huson et al., 2007), MGRAST (Glass et al., 2010), CAMERA (Seshadri et al., 2007) (discontinued in July 2014), and

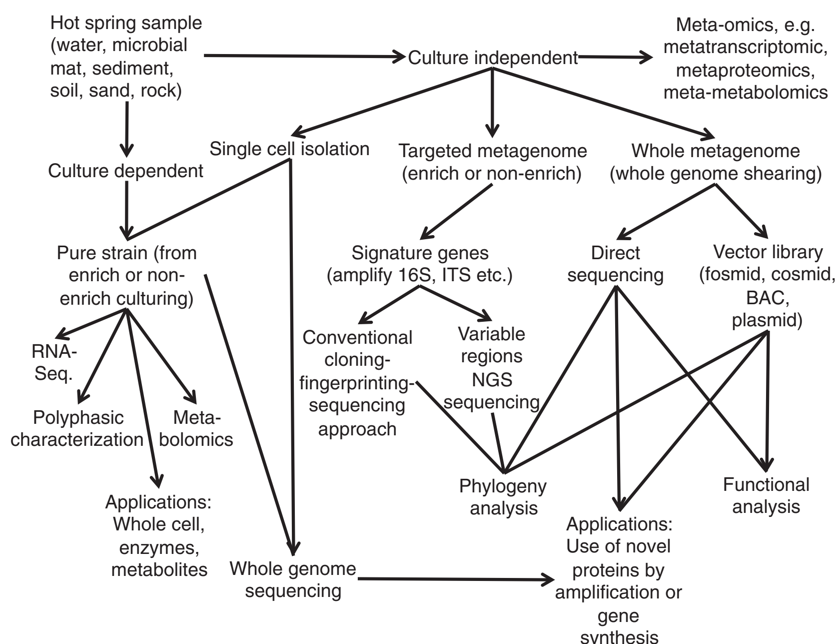

Fig. 2. Overall flow diagram of research related to thermophiles and hyperthermophiles.

Table 2
List of useful websites, software application, or tools related to thermophilic metagenomic and genomic research.

Name	Description/General function as displayed in the respective website homepages	Link
ALLPATHS	A *de novo* assembler of whole-genome sequencing reads	http://www.broadinstitute.org/software/allpaths-lg/blog/
AmporaNet	Estimates the taxonomic composition of bacterial and archaeal communities from metagenomic shotgun sequencing data. AmporaNet uses 31 bacterial and 104 archaeal protein coding marker genes	http://pitgroup.org/amphoranet/
AMYLOMICS	A website that shows the aims and progress of using metagenomics to find industrial enzymes	http://amylomics.org/
antiSMASH	antiSMASH allows the rapid genome-wide identification, annotation, and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genomes.	http://antismash.secondarymetabolites.org/
BASys	Web server that performs automated, in-depth annotation of bacterial genomic (chromosomal and plasmid) sequences	
Blast2GO	Java-based functional annotation workstation	http://www.blast2go.com/b2ghome
ClaMS	ClaMS is a sequence composition-based classifier for metagenomic sequences. ClaMS works by capturing signatures of each sequence based on the sequence composition.	http://clams.jgi-psf.org/
CRISPRfinder	A web service offering several tools related to clustered regularly interspaced short palindromic repeats	http://crispr.u-psud.fr/Server/
DIYA (Do-It-Yourself Annotator)	Modular and configurable open source pipeline for the rapid annotation of microbial genome sequences	https://sourceforge.net/projects/diyg/
Earth Microbiome Project	A massively multidisciplinary effort to analyse microbial communities across the globe	http://www.earthmicrobiome.org/publications/
EDGAR	Useful for comparative analysis of prokaryotic genomes	https://edgar.computational.bio.uni-giessen.de/cgi-bin/edgar_login.cgi
FGENESB	Annotations of bacterial genome	http://www.softberry.com/
GHOSTM	A homology search tool for huge short reads generated by sequencers. GHOSTM can detect remote homologs like BLAST and is about 40 times more efficient than BLAST.	http://code.google.com/p/ghostm/
GOLD:Genomes Online Database	Resource for comprehensive access to information regarding genome and metagenome sequencing projects	http://www.genomesonline.org/
IDBA-UD	An iterative de Bruijn graph *de novo* assembler for short-read sequencing data with highly uneven sequencing depth	http://i.cs.hku.hk/~alse/hkubrg/projects/idba_ud/
IGS Analysis Engine	Resource for prokaryotic genome annotation, comparative, and transcriptome analysis	http://ae.igs.umaryland.edu/cgi/index.cgi
JGI database	Resource for plant, fungal, metagenome, and microbial programs	http://jgi.doe.gov/our-science/science-programs/metagenomics/search-metagenome-project-list/
JGI genome portal	Resource for genomes and metagenomes	http://genome.jgi-psf.org/programs/metagenomes/tools.jsf
JGI-IMG	Analysis tools for examining publicly available genomes in IMG (Integrated Microbial Genomes and Metagenome)	https://img.jgi.doe.gov/cgi-bin/w/main.cgi
Krona	Krona allows hierarchical data to be explored with zoomable pie charts using input data from Excel or Krona tools.	http://krona.sourceforge.net
MEGAN	MEGAN (MEta Genome ANalyzer) analyses, displays, and compares taxonomic content of single or multiple datasets. MEGAN is able to display GO ontology, KEGG, and SEED information.	http://ab.inf.uni-tuebingen.de/software/megan/
MetaAnnotator MetaCluster	MetaAnnotator is used for binning and annotating short paired-end reads. MetaCluster5.0 is an unsupervised binning method.	http://i.cs.hku.hk/~alse/MetaCluster/
MetaBioMe	A web resource to find novel homologs for known commercially useful enzymes in metagenomic datasets and completed bacterial genomes	http://metasystems.riken.jp/metabiome/
metAMOS	Integrated assembly and analysis pipeline for metagenomic data	http://cbcb.umd.edu/software/metAMOS
MetaPhlAn	MetaPhlAn is used to profile the composition of microbial communities from metagenomic shotgun sequencing data. MetaPhlAn relies on unique clade-specific marker genes identified from 3000 reference genomes.	http://huttenhower.sph.harvard.edu/metaphlan
MetaVelvet	Metagenomic assembly for mixed short reads of multiple species	http://metavelvet.dna.bio.keio.ac.jp/
MGRAST	An automated analysis platform for metagenomes providing quantitative insights into microbial populations	http://metagenomics.anl.gov/
Molbiol-tools	A comprehensive link to other useful tools	http://molbiol-tools.ca/Genomics.htm
mOTUs	Provides species-level profiles that are generated by mapping reads from metagenomes to a database consisting of 10 universal marker genes	http://www.bork.embl.de/software/mOTU/index.html
MycoCosm	Fungal genome resource	www.jgi.doe.gov/fungi
NCBI genome browser	Resource for information on genomes including sequences, maps, chromosomes, assemblies, and annotations	http://www.ncbi.nlm.nih.gov/genome/browse/
NCBI Prokaryotic Genome Annotation Pipeline	Annotates bacterial and archaeal genomes	http://www.ncbi.nlm.nih.gov/genome/annotation_prok/
PHAST	A web server designed to rapidly and accurately identify, annotate, and graphically display prophage sequences within bacterial genomes or plasmids	http://phast.wishartlab.com/
QIIME	A software package for the comparison and analysis of microbial communities, primarily based on high-throughput amplicon sequencing data (such as small subunit ribosomal rRNA) generated on a variety of platforms, but also the supporting analysis of other types of data, such as shotgun metagenomic data	http://qiime.org/
RAST	Fully automated service for annotating complete or nearly complete bacterial and archaeal genomes	http://rast.nmpdr.org/
SOAPdenovo2	Metagenome assembler	http://soap.genomics.org.cn/soapdenovo.html

(continued on next page)

Table 2 (*continued*)

Name	Description/General function as displayed in the respective website homepages	Link
Terragenome (International Soil Metagenome Sequencing Consortium)	Facilitates the international community involved in soil metagenomics by holding periodic meetings to plan strategies and share information, coordinating sequencing and bioinformatics activities, hosting workshops to train students and scientists in metagenomic analysis, and generally enhancing communication and information sharing through the TerraGenome website	http://www.terragenome.org/
VIROME	VIROME informatics pipeline focuses on the classification of predicted open-reading frames (ORFs) from viral metagenomes	http://virome.dbi.udel.edu/
VITCOMIC	VITCOMIC analyses millions of bacterial 16S rRNA gene sequences and calculates the overall taxonomic composition for a microbial community. VITCOMIC plots all sequences in a single figure and indicates relative evolutionary distances.	http://mg.bio.titech.ac.jp/vitcomic/
WebMGA	WebMGA includes over 20 commonly used tools for analysis, such as ORF calling, sequence clustering, quality control of raw reads, removal of sequencing artefacts and contaminations, taxonomic analysis, and functional annotation.	http://weizhong-lab.ucsd.edu/metagenomic-analysis/server/

several other tools that are listed in Table 2. Due to complexities inherent to microbial communities involving the cohabitation of closely related species, the assembler contigs generate build-ups of consensus sequences contributed by slight variations in the genomes of the subspecies tested (Peng et al., 2011). In an excellent review, Thomas et al. (2012) commented that the development of metagenome assemblers is still in an early stage, and it is difficult to access the accuracy of the data. This problem for shotgun metagenomics is also complicated by short read lengths and, hence, only a handful of NGS instruments are suitable for this kind of study. Davenport and Tümmler (2013) have discussed advances in computational analysis and tips in dealing with shotgun sequencing, although it is beyond the scope of this review to address all of these points.

3. Genomes of prokaryotic thermophiles

Currently, only two bacterial families are categorised as hyperthermophiles (*Aquificaceae* and *Thermotogaceae*), while the majority of hyperthermophiles are *archaea*. At the time of writing, thirteen complete genomes of *Aquificae* and 20 complete genomes *Thermotogae* have been assembled. A total of 38 complete genomes of *Crenarchaeota* and 120 complete genomes of *Euryarchaeota* could be accessed from KEGG (http://www.genome.jp/kegg/catalog/org_list.html). Datasets for the genomes of common thermophiles (e.g. *Geobacillus* and *Thermus*) are abundant. With the availability of thermophilic and hyperthermophilic genome data, many valuable lessons have been learned, some of which are summarised below.

3.1. Life under extreme conditions is made possible by special genome features

A high GC content is not a good indicator of thermophilicity (Hickey and Singer, 2004). For example, *Aquifex aeolicus* has a low GC content of 43.4% (Deckert et al., 1998), despite its high OGT of 95 °C. Moreover, most hyperthermophilic archaea do not have high GC contents either. Two examples include the *Caldisphaera* spp. and *Methanotorris* spp., which have average GC contents of only 30% and 32.3%, respectively, while *E. coli* has a GC content of 50.7%. However, it is possible that the GC content of rRNA/tRNA may be a better indicator of thermophilicity than the GC content of genomic DNA (Trivedi et al., 2005).

There is no simple correlation between the OGT and genomic features because thermophilicity results from a combination of factors. Some of these factors are mentioned herein. Sabath et al. (2013) found that thermophiles have less intergenic DNA, which makes their genomes compact. On average, every 1 kb of a bacterial genome encodes a gene, but the length is slightly shorter in thermophilic archaea. Based on comparisons of 1553 prokaryotes, cells that grow below 45 °C have genomes larger than 6 Mbp, while the average genome length is less than 4 Mbp for thermophiles (Sabath et al., 2013). It seems that if the OGT is higher still, the genome size may be smaller than 2 Mbp, especially for hyperthermophilic archaea and Aquificae. An example of this phenomenon can be seen with *Hyperthermus* spp., which has an average OGT of 100.5 °C and a mean genome size of only 1.67 Mbp.

Thermophiles have adopted several strategies to stabilise their proteins. Thermophilic proteins have shorter amino acid lengths than their counterparts from non-thermophilic organisms, which may reflect the importance of a reduced number of flexible regions in the native protein structures. Based on 204 complete proteomes of bacteria and archaea, the amino acids Ile, Val, Tyr, Trp, Arg, Glu, and Leu have been found to be correlated with OGT (Zeldovich et al., 2007). In an earlier analysis, Singer and Hickey (2003) discussed specific patterns of thermophiles from the viewpoints of codon usage, amino acid composition, and nucleotide content. Thermophiles also use compatible solutes to stabilise cell components (Borges et al., 2010; Empadinhas and da Costa, 2011; Faria et al., 2008; Santos and Da Costa, 2002).

In this genomic era, most researchers search for insights into thermophilicity using the complete annotated genome. It is well established that important ORFs encode heat shock proteins (HSPs), chaperones, chaperonins (assisting the folding of macromolecules); agmatine (thought to stabilize DNA and RNA); spermidines (polyamines from ribosomes that maintain membrane potentials); polyamines (needed for growth, possibly as membrane stabilizers); α- and β-subunit prefoldins (protein folding chaperons), SOS regulons (DNA damage responses); and other DNA repair systems. Reverse gyrase (a heat-protective DNA chaperone) is also believed to play an important role in genome thermostability (Heine and Chandra, 2009). In members of the *Pyrococcus* and *Methanococcus* genera, the chaperone-like protein-folding activity of the FK506-binding protein (FKBP) is involved in stabilising the solubility of proteins (Fukui et al., 2005).

Not all of the aforementioned biomolecules are found exclusively in thermophiles. For instance, DNA-repair systems are also present in mesophilic cells. To elucidate the true mechanisms of thermophilic survival, rational guesswork based on genomic data alone is insufficient. Therefore, DNA microarray, proteome, and transcriptome sequencing (RNA-seq) experiments are also needed. Several transcriptome analyses on various archaea have been performed, and some of the more important findings are summarised in a previous review (Walther et al., 2011). To state a few examples here, the OGT for the methanarchaeon *Methanococcus jannaschii* is 85 °C, and when the cells are grown at 95 °C, upregulation was observed with the prefoldin alpha subunit, small HSP, thermosome subunit, and protease regulatory subunit, CRISPR-associated proteins, and a hypothetical protein

(Boonyaratanakornkit et al., 2005). The functions of hypothetical proteins are, by definition, unknown, but some may eventually be found to serve novel stabilising or protein-repair roles.

Previous experimental data identified HSP60 and elongation factor Tu as factors contributing to the thermostability of *Thermotoga maritima* (Wang et al., 2012c). Unexpectedly, proteins involved in the central carbohydrate metabolism of *Thermotoga maritima* were upregulated at high temperature, although the exact relationship of such upregulation to thermophilicity remains unclear. As more temperature-related RNA-seq data become available in the future, hopefully one day we can determine shared global stabilising mechanisms, as well as pinpoint specific thermophilicity signatures in each thermophile genus or domain.

3.2. Horizontal gene transfer and symbiosis between thermophiles

Previous work suggested that horizontal gene transfer (HGT) is an important process for thermophiles. A few examples are mentioned here. The *Thermomicrobium* genus is a rare thermophilic genus with only two properly described species. The genome of *Thermomicrobium roseum* DSM 5159 contains a circular chromosome (2.0 Mbp) and a megaplasmid (919,596 bp) (Wu et al., 2009). Interestingly, the complete set of genes encoding the flagellar system of this bacterium is found in the megaplasmid and not in the bacterial chromosome, which is unusual. The authors suggested that these genes are the result of HGT activity. In another example, the archaeon *Picrophilus torridus* grows optimally at 60 °C and at a pH of 0.7. Examination of the genome revealed that certain genes supporting the extreme optimal growth conditions of thermoacidophilic archaea were transferred from Crenarchaea and bacteria (Futterer et al., 2004). Twenty-four percent of the genes in the bacterium *Thermotoga maritima* were acquired from archaea (Nelson et al., 1999), thereby enabling *T. maritima* to adapt to high temperatures (Zhaxybayeva et al., 2009). From the above findings, HGT could occur between domains (i.e. hyperthermophilic bacteria and hyperthermophilic archaea) or between close genera (i.e. *Thermotoga* and *Aquifex*; *Anoxybacillus* and *Geobacillus* (Goh et al., 2014)), as part of survival mechanism under harsh conditions.

Symbiosis among thermophiles can enable two or more parties to thrive in extreme conditions. *Candidatus Chloracidobacterium thermophilum* is difficult to culture in a laboratory unless co-cultured with *Anoxybacillus* and *Meiothermus* species. The intercellular communication and symbiosis between these three different thermophilic genera have yet to be studied. The genome of *Candidatus Chloracidobacterium thermophilum* lacks a sulphate reductase gene, and the *Anoxybacillus* and *Meiothermus* species may complement the sulphate reduction function lacked by the *Candidatus* species (Costas et al., 2012). In another example, the Crenarchaeota *Ignicoccus hospitalis* serves as a specific host for *Nanoarchaeum equitans*, and gene exchange between them was found to result in additional biochemical functions (Podar et al., 2008). Exactly how these two cell types recognise and interact with each other is currently unknown.

4. Applications of thermophiles and thermozymes

In general, many thermophiles are polyextremophiles, meaning that they are also well adapted to thrive under other extreme environmental conditions, such as those related to pH, redox potential, salinity, or the presence of a wide spectrum of toxic compounds. The advantages of using thermophilic microorganisms in biotechnological applications are listed in Table 3. Biotechnological applications of thermophiles can be divided between applications using whole cells, either in pure culture or in consortia, and applications using their macromolecules or metabolites. Among the numerous biotechnological applications that use thermophile cells (Table 4), those concerning bioremediation strategies and clean production technologies are among the most interesting. Thermophiles also have well known and potentially highly productive applications in bioenergy, biomining, thermozyme, and biosurfactant production.

4.1. Bioenergy – biofuels

The search for new sources of energy production, especially bioenergy, has become a topic of worldwide interest due to increased concerns over declining fossil fuel reserves and climate change. Biofuels can exist as solids, liquids, and gases. According to European Committee for Standardization, CEN (TC335) solid biofuels may be sub-classified as woody biomass, herbaceous or fruit biomass and blends, and mixtures based on the origin and source utilised for their production (Alakangas et al., 2006).

Liquid biofuels (i.e. bioethanol, biodiesel, biobutanol, and biokerosene) are obtained by fermenting materials such as starch and lignocellulosic biomass, or by extracting lipid fractions from various sources such as plants and microorganisms. Thermophiles are able to produce thermostable enzymes that efficiently degrade lignocellulosic biomass (i.e. xylan, cellulose, and hemicellulose) for liquid biofuel production. For instance, cellulases and xylanases from *Caldicellulosiruptor*, *Caldanaerobius*, and *Clostridium* spp. exhibit efficient lignocellulosic-degrading activity with potential applications in the production of various liquid biofuels (Bhalla et al., 2013; Han et al., 2012; Su et al., 2013). In a recent study involving biobutanol production, Bhandiwad et al. (2013) published very interesting results with *Thermoanaerobacterium thermosaccharolyticum* where they over-expressed specific genes of the *bcs* operon responsible for forming butyryl CoA, and they increased *n*-butanol production by 180% compared to the wild type strain.

Methane and hydrogen are examples of gas biofuels that can be obtained by the anaerobic fermentation of different feedstocks and waste materials. A review published in 2014 gathered the most recent information on anaerobic thermophiles, including the production of gas biofuels (Canganella and Wiegel, 2014). Biohydrogen can be produced by the photosynthetic conversion of sunlight and water facilitated by certain microorganisms, such as cyanobacteria. Cyanobacteria are able to convert up to 10% of the Sun's energy into biomass, compared to only 1% that is converted by conventional energy crops such as corn or sugarcane, or the 5% achieved by algae (Parmar et al., 2011). In addition, biohydrogen production can be achieved by anaerobic dark fermentation of organic wastes (Hawkes et al., 2007; Kongjan et al., 2010; Li and Fang, 2007). In addition, hydrogen production via enzymatic reactions is highly interesting. When oxygen and hydrogen are combined in a fuel cell, the reaction provides electricity and some heat, with water as the only waste product. Biohydrogen can be obtained by transforming cellulose into glucose and then converting the glucose product and its byproduct (gluconic acid) into hydrogen (Woodward et al., 2000). The Woodward team has identified, isolated, purified, and characterised extremozymes that might be useful for energy production. The process of converting glucose to hydrogen is more efficient at higher temperature; thus, it makes sense to replace standard enzymes with those from extremozymes (i.e. hydrogenase). The deep-sea hydrothermal vent *Pyrococcus furiosus* hyperthermophile produces a hydrogenase that works most efficiently at 85 °C and is also one of only two enzymes known to accept electrons from NADPH to produce hydrogen (Raven et al., 1992).

Consolidated bioprocessing (CBP) offers a promising alternative to biohydrogen production using thermophiles (Parisutham et al., 2014). Its most important advantage is in the reduction of costs, especially during the pretreatment stage where hazardous chemicals and expensive enzymes are generally required for the hydrolysis of biomass. CBP combines enzymatic pretreatment, enzymatic saccharification, and sugar fermentation in a single step using specific microorganisms as biorefineries. For example, anaerobic biohydrogen production using *Caldicellulosiruptor saccharolyticus* showed outstanding results in CBP of switchgrass, microcrystalline cellulose, and glucose (Talluri et al., 2013).

Table 3
Advantages of the use of thermophiles in biotechnological applications.

High metabolic activity leads to enhanced product formation rates
Inactivation or elimination of contaminant/pathogenic mesophilic microorganisms
Production of heat stable macromolecules and metabolites
Metabolic reactions occur at the same high temperature that substrates solubilise
No cooling steps required after heating steps
Increased diffusion rates, ionization and solubility of chemicals
Lower density, surface tension, and viscosity of solutions enhance reaction rates
Direct recovery of volatile products
Low bacterial mass formation yields higher ratios of desired product over assimilated substrate and lower production of waste
Express thermostable enzymes

Methane trapped in coal beds has emerged as another alternative energy source. Biomethane is the result of a series of biochemical reactions where coal is converted to methane by certain anaerobic bacteria (Kotelnikova, 2002). The microbial conversion of coal to methane has long been a topic of research (Jones et al., 2008; Thielemann et al., 2004). Currently, research in this area is moving towards thermophilic consortia capable of stimulating methanogenesis from different coal sources at high temperature that characterise coal mining operations. A study using different coal substrates from the Jharia coal mines in India reported 49% methane production at 65 °C using an autochthonous methanogenic microbial community enriched from water in the same coal mines. A phylogenetic study of the community showed that *Methanoculleus thermophilus* was the predominant species involved (Lavania et al., 2014).

Despite the promising use of thermophiles mentioned above, the use of thermophilic microorganisms in the production of biofuels also presents a series of disadvantages (Lin and Xu, 2013). It is well established that different thermophiles can utilise different carbohydrate sources, including complex polymers. Moreover, some species can co-utilise carbohydrates, a feature that is surprising, especially for monomers, as certain thermophiles seem to lack a mechanism known as 'carbon catabolite repression' that represses sugar co-utilisation in many mesophilic bacteria (Gorke and Stulke, 2008). In spite of the apparently wider metabolic possibilities, the majority of thermophiles present low efficiencies in carbon utilisation, and in some cases, more than one species may be needed for the complete degradation of certain polymers, such as cellulose, with the exception of species such as *Bacillus*, *Geobacillus*, and *Paenibacillus* (Rastogi et al., 2009). Similarly, even though thermophiles have higher enzymatic reaction rates, most cultures of thermophiles achieve low cellular counts. The total yield of biofuels produced by thermophiles is still below industrial needs. Thermophiles present another problem: they are more recalcitrant to genetic manipulation than mesophiles.

4.2. Biomining

Biomining comprises different biological processes with the aim to enhance the recovery of metals from ores. Bioleaching and biooxidation are two bio-extractive processes applied to sulphide minerals included in biomining. Both bioprocesses are performed by the same microorganisms and employ the same mechanisms; however, during bioleaching, the metal of interest is solubilised by biological catalysis, while in biooxidation, the microorganisms dissolve the mineral matrix that occludes the metal to be recovered, which can be later dissolved using other chemical leaching agents. The most suitable biomining microorganisms require the ability to generate oxidizing and acidic conditions that attack metal sulphides and allow the release of the metal of interest to the acidic water solution as soluble sulphates. Acidophilic iron- and sulphur-oxidising prokaryotes are involved in bioleaching and biooxidation processes because they can oxidise Fe (II) to Fe (III) and sulphur compounds to sulphuric acid. Fe (III), a suitable oxidising agent for metal sulphides is continually regenerated by the action of iron-oxidising microorganisms, releasing the metal from the mineral to the solution. In addition, sulphuric acid produced by sulphur-oxidising microorganisms keep the pH at low values to retain metal ions in solution (Donati and Sand, 2007).

Although biomining processes could theoretically be applied to other metals, the main commercial applications are developed for copper bioleaching and the biooxidation of refractory gold ores. Bioleaching of other metals (cobalt, nickel, and zinc) is rarely applied at the commercial scale, except under very specific conditions. Two different technologies

Table 4
Biotechnological applications of whole-cell thermophiles.

Whole cell application	Thermophile	Action	Reference
Biofuel production			
	Caldicellulosiruptor bescii, *Caldanaerobius polysaccharolyticus*	Xylan degrading activity	Han et al. (2012); Su et al. (2013)
	Thermoanaerobacterium thermosaccharolyticum	Butanol production	Bhandiwad et al. (2013)
	Caldicellulosiruptor saccharolyticus	Consolidated bioprocess for anaerobic hydrogen production	Talluri et al. (2013)
	Methanoculleus thermophiles and other autochthonous thermophiles from coal mine	Methane production from coal mine substrates	Lavania et al. (2014)
Bioremediation			
	Geobacillus sp., *Anoxybacillus flavithermus*, *Thermus thermophiles*, *Thermococcus zilligii*	Biosorption of toxic metals	Chatterjee et al. (2010); Sar et al. (2013)
	Thermus scotoductus, *Pyrobaculum. islandicum*, *Thermoanaerobacter* sp., *Thermoterrabacterium ferrireducens*	Immobilisation of radionuclides	Chernyh et al. (2007)
	Aeribacillus sp., *Geobacillus* sp.	Biodegradation of recalcitrant aromatic compounds and hydrocarbons	Mnif et al. (2014)
	Anoxybacillus sp.	Degradation of azo-dyes	Deive et al. (2010)
	Non characterised thermophiles	Construction of biofilters for gas deodoration	Ryu et al. (2009)
Bioleaching			
	Sulfobacillus sp., *Ferroplasma* sp., *Acidianus infernus*	Enhanced Cu extraction from chalcopyrite	Abdollahi et al. (2014); d'Hugues et al. (2002); Qin et al. (2013)
	Acidianus brierleyi, *Acidianus manzaensis*, *Metallosphaera sedula*, *Sulfolobus metallicus*	Metal solubilisation from nickel-copper sulphide	Li et al. (2014)
Food, animal feed, compost for agriculture, cultivation production	*Clostridium* sp., *Caldicellulosiruptor* sp.	Improved compost quality by cellulolytic activity	Basen et al. (2014); Han et al. (2012); Sizova et al. (2011)

are used in biomining: stirred tanks and heaps. Most commercial biomining applications run at temperatures below 40–50 °C, mainly because the first and better-studied biomining microorganisms are mesophiles or moderate thermophiles (Donati and Sand, 2007). However, it is now well known that the use of thermophiles can be advantageous for bioleaching because the possibility of operating at higher temperatures would increase the overall rate of the process. In addition, the use of thermophiles would eliminate the energy input requirements for cooling the system (bioleaching reactions are exothermic, causing a serious increase in temperature in bioreactors and inside the heaps) and would decrease the passivation of mineral surfaces. The use of higher temperatures could cause some problems, such as lowering the solubility of reactant gases (like O_2 consumed by the aerobic thermophiles) and increasing the evaporation rate and increasing the corrosive action of the acidic liquor (which would require changing from steel to ceramic tanks). However, the use of thermophiles could be a solution for the current main problem in copper bioleaching: enhancing copper extraction from chalcopyrite. Chalcopyrite ($CuFeS_2$) is a mineral species that accounts for approximately 70% of the world's copper reserves (Wang, 2005), however it is highly recalcitrant to chemical or mesophilic biological leaching (Johnson et al., 2008). Several studies have proven that thermoacidophilic microorganisms can generate satisfactory copper recovery yields, much higher than those obtained with mesophilic microorganisms (Abdollahi et al., 2014; d'Hugues et al., 2002; Li et al., 2014; Qin et al., 2013). A notable example is the archaeon *Acidianus sulfidivorans* sp. nov., described by Plumb et al. (2007), which is capable of metal extraction under highly extreme conditions. It has optimum growth conditions at a pH of 0.35–3.0 and a temperature of 45–83 °C and is able to grow on various sulphide minerals, including pyrite, arsenopyrite, and chalcopyrite.

4.3. Bioremediation

Bioremediation can be defined as an eco-friendly pollution treatment technology that exploits microorganisms (autochthonous or artificially introduced), generally at low costs to reduce, eliminate, contain, and/or transform environmental organic or inorganic contaminants to benign products (Tabak et al., 2005). The use of extremophilic microorganisms has become a promising alternative to treat metal-contaminated sites (Sen et al., 2014). This is because thermophiles exhibit enhanced metal solubilisation through sulphur and/or iron metabolic oxidation. Thermophilic microbial communities are able to couple metal reduction with the oxidation of different organic and inorganic substrates. They are also able to reduce a wide spectrum of metals, including Mn (IV), Cr (VI), U (VI), Tc (VII), Co (III), Mo (VI), Au (I, III), and Hg (II), which can be used for the immobilisation of toxic metals during bioremediation, and even of hot wastewater in disposal sites containing radioactive wastes.

The biosorption of different metals, such as Cd, Cu, Ni, Zn, and Mn has been evaluated with *Geobacillus* species, *Anoxybacillus flavithermus*, *Thermus thermophiles*, and the archaeon *Thermococcus zilligii* (Chatterjee et al., 2010; Sar et al., 2013). Thermophiles present great advantages for the precipitation and/or immobilisation of toxic metals and radionuclides (U, Cr, Tc, Co) in the treatment of heated nuclear waste streams. Enzymatic uranium and technetium reduction has been shown in *Thermus scotoductus*, *Pyrobaculum islandicum*, *Thermoanaerobacter* sp., and *Thermoterrabacterium ferrireducens* (Chernyh et al., 2007). The reduction of Cr (VI), Se (IV), and Te (IV) by thermophilic species has also been reported (Sar et al., 2013).

The use of thermophiles for the biodegradation of hydrocarbons with low water solubility is of interest, as the solubility of hydrocarbons and thus their bioavailability are enhanced at elevated temperatures. Indigenous thermophilic bacteria from the genus *Bacillus* were used to degrade crude oil spills left in the Persian Gulf after oil extraction procedures and the Gulf war (Al-Maghrabi et al., 1999). Other examples of biodegradation and bioremediation of hydrocarbons were discussed by Margesin and Schinner (2001). Regarding the bioremediation of haloorganic compounds, Guerrero-Barajas et al. (2014) reported 74% biodegradation of trichloroethylene (TCE) biodegradation in an upflow anaerobic sludge blanket inoculated with sulphidogenic sludge from hydrothermal vents sediments. The microorganisms identified in the sludge during TCE biodegradation were related to the genera *Dehalobacter*, *Desulfotomaculum*, *Sulfospirillum*, *Desulfitobacterium*, *Desulfovibrio*, and *Clostridium*.

Thermophiles may also serve in an increasing number of potential applications in the bioremediation of highly pollutive organic compounds. Thermophilic bacteria related to the *Aeribacillus* and *Geobacillus* genera have shown biodegradation capacities for several recalcitrant aromatic compounds and hydrocarbons (Mnif et al., 2014). There is one report on the use of *Anoxybacillus* sp. in degrading azo dyes, which represent approximately 50% of the industrially used dyes and are extremely hazardous compounds (Deive et al., 2010). In the past few years, thermophiles have been tested for the construction of biofilters for deodorising gases emitted from various industries. Thermophilic biofiltration has the advantage of not requiring cooling of the gases, thus diminishing the economic cost of the process. Promising results have been reported for the removal of volatile organic compounds, such as ethanol, benzene, ethyl acetate, and sulphur-containing gases (Ryu et al., 2009). Although the bioremediation of inorganic and organic pollutants has many advantages, i.e. more environmentally friendly with lower operating costs than that of traditional clean-up methods, its slow rate is the main disadvantage and the probable reason this technology has not been adopted. Considering the results of these studies, the use of thermophilic species in bioremediation seems to be a promising tool to increase the efficiency of the processes.

4.4. Thermozymes

The outstanding properties of thermozymes are suited to particular industries that employ elevated temperatures, such as the pulp and paper, food, brewing, and feed processing industries. Thermophiles are often highly resistant to harsh conditions such as chemical denaturing agents, wide pH ranges, and/or non-aqueous solvents. Examples of such enzymes are cellulases, xylanases, pectinases, chitinases, amylases, pullulanases, proteases, lipases, glucose isomerases, alcohol dehydrogenases, and esterases (Bergquist et al., 2014; Chai et al., 2012; Elleuche et al., 2014; Kahar et al., 2013). Thermophilic enzymes have played important roles not only at the industrial level (Table 5), but also in pharmaceutical applications requiring use of specific aldolases for the synthesis of enantiopure compounds (Falcicchio et al., 2014).

Thermozymes are most commonly used in the production of biofuels from starch and lignocellulosic material. The conversion of starch to single fermentable glucose molecules requires high temperature and low pH conditions, with the catalytic action of amylases, glucoamylases, and/or pullulanases. For the production of second-generation biofuels, the complex structure of lignocellulosic materials must be degraded by chemical and physical means (high temperature, low pH, high pressure) to obtain cellulose, hemicellulose, and lignin, which then must be treated with the corresponding hydrolytic enzymes (Elleuche et al., 2014). Considering the pretreatment conditions, extremophilic hydrolytic enzymes are better suited for the task than mesophilic enzymes.

4.5. Biosurfactants

Biosurfactants are amphiphilic compounds that are produced by microorganisms and help to enhance the emulsification and dispersal of water-insoluble compounds. Biosurfactants can be glycolipids, lipopolysaccharides, lipoproteins, phospholipids, hydroxylated and cross-linked fatty acids, and/or complete cells. The use of biosurfactants has increased notably in industries related to food, agriculture, pharmaceuticals, petroleum, and paper/pulp due to their lower toxicity; higher biodegradability; better environmental compatibility; higher foam formation; higher specific activity at extreme temperature, pH, and

Table 5
Recent reports on macromolecules produced by thermophiles with known or potential biotechnological applications.

Macromolecule	Thermophile/Source	Activity	Industry/Bioprocess	Reference
Thermozymes				
Aldolase	*Sulfolobus solfataricus, S. acidocaldarius, Thermoproteus texas, Hypertermus butilicus*	Stereoselectibe C-C bond formation	Pharmaceutical industry	Falcicchio et al. (2014)
G6PD, 6PGD, R5PI, others*	*Geobacillus stearothermophilus, Moorella thermoacetica, Thermotoga maritima*	Glucose oxidation	Enzymatic fuel cell	Zhu et al. (2014)
Hydrogenase	*Pyrococcus furiosus*	Final stage of glucose oxidation by oxidative pentose phosphate cycle	Enhanced production of biohydrogen	Woodward et al. (2000)
Pullulanase	*Thermotoga neapolitana*	Hydrolysis of α-1, 6-glucosidic linkages	Biofuel production	Kang et al. (2011)
α-Amylase	Uncultured clone (metagenomic screaning)	Starch hydrolysis at high temperature and high NaCl concentration	Biogas production	Jabbour et al. (2013)
Endoxylanase	*Acidothermus cellulolyticus*	β-1, 4-xylan cleavage	Biofuel production from lignocellulose	Barabote et al. (2010)
Chitinase	*Sulfolobus tokodaii*	Hydrolysis of β-1, 4-glycosidic bond in chitin	Biomedical, pharmaceutical, food and environmental	Staufenberger et al. (2012)
Keratinolytic protease complex	*Meiothermus ruber*	Keratin hydrolysis	Feather degradation (industrial waste)	Kataoka et al. (2014)
Protease	Microbial community from solid state fermentation reactor	Degradation of hair waste from tannery	Leather industry	Abraham et al. (2014)
Lipase	*Geobacillus* sp.	Hydrolysis of divers lipid substrates	Biofuel, cosmetics or perfumes production, leather and pulp industries	Gudiukaitė et al. (2014)
Acidic thermostable lipase	*Bacillus pumilus*	Degradation of palm oil	Treatment of palm oil-containing wastewater	Saranya et al. (2014)
Carboxylesterase	*Geobacillus thermoleovorans*	Carboxyl ester hydrolysis	Agriculture, food, and pharmaceutical industries	Soliman et al. (2014)
Biosurfactants				
Thermostable lipopeptide-type biosurfactant	*Aneurinibacillus thermoaerophilus*	Reduction of the surface tension of water	Agriculture, food, and pharmaceutical, petroleum and paper/pulp industries	
Glicolipid-type biosurfactant	*Alcaligenes faecalis*	Decrease of surface tension at air–water interface, enhance of emulsification	Oil industry	Bharali et al. (2011)
Cyclic lipopeptides	*Bacillus licheniformis, B. subtilis, B. subtilis*	Oil mobilization, enhance in oil recovery	Oil industry	Joshi et al. (2008)

* G6PD (glucose-6-phosphate dehydrogenas), 6PGD (6-phosphogluconate dehydrogenase), R5PI (ribose-5-phosphate isomerase).

salinity; and a greater possibility of obtaining them from renewable sources (Sharafi et al., 2014). As in other fields, thermophiles have demonstrated potential for the production of biosurfactants with higher temperature stabilities and increased resistance to other extreme physicochemical parameters, such as pH and salinity. Species with enhanced performance thus far identified have been related to the *Aneurinibacillus*, *Geobacillus*, *Alcaligenes*, *Bacillus*, and *Brevibacillus* genera (Bharali et al., 2011; Joshi et al., 2008; Mnif et al., 2011; Sharafi et al., 2014).

5. Future prospects on the studies of thermophiles

5.1. Exploring the unculturable thermophiles by culture medium reengineering and single-cell sequencing

Many novel methods have been developed in attempts to cultivate the yet-to-be-cultured prokaryotes (Pham and Kim, 2012). Despite these efforts, a vast number of cells are still unable to be cultured in laboratories. During early work conducted at the hot springs of Yellowstone Park, many unknown 16S rRNA sequences were coded as candidate divisions or also known as *Candidatus*. The occurrence of unculturable bacteria does not signify 'can never be cultured' (Stewart, 2012). Culture medium reengineering is somehow needed. Instead of using blind selection or trial-and-error approaches of medium modification, researchers may need to redesign media recipes based on microbiota populations, as determined via 16S rRNA metagenomics. In addition, functional gene microarrays, such as the GeoChip 4.0 with 83,992 probes (Tu et al., 2014) or the more recent GeoChip version 5.0 (167,044 probes), may provide some insight in surveying the metabolic activity and dynamics of the community. Obtaining such information may be useful for redesigning media.

As its name implies, single-cell manipulation can be used to study microorganism at the one-cell level. This technique can be very useful in characterizing thermophiles that currently cannot be cultured despite massive efforts in growth media experimentation that have been undertaken to culture bacteria from various habitats. Single-cell manipulation and single-cell NGS involve the direct capture of an unculturable bacterium, followed by single-cell NGS. For example, a single cell of candidate division OP11 (ZG1) was isolated using microfluidics approach (Youssef et al., 2011), and almost half of the partial sequenced genome was unique, showing few hits when compared to known genes. In the co-author's (K.G.C.) laboratory, we have used a micromanipulator equipped with a glass micropipettor and a light microscope for guidance to directly capture a bacterium from a water source, followed by whole-genome sequencing. Using this method, we have isolated and whole-genome sequenced several environmental bacteria, all of which are considered unculturable. Other useful cell sorting methods include microfluidic flow, laser tweezers, optoelectronic tweezers, microfluidic arrays, flow cytometry, and microdrop emulsion. The individual features of each method have been reviewed by Blainey (2013) and Lasken (2012). Some techniques and commercial available tools are suitable for large-sized cell lines, but are not applicable for use with smaller microorganisms. In addition, the isolation of single cells using various techniques is inevitably limited by insufficient nucleic acid contents. However, this issue can be well addressed by a technique called multiple strand displacement amplification, which enables whole-genome amplification and hence, has made NGS of single-cell genomes and even transcriptomes possible. Table 6 shows some examples of recent work made in effort to obtain genomic sequences from uncultivated cells; little work has been performed using thermophilic samples, but more is expected in the future.

Table 6
Summary of findings from single-cell and metagenome sequencing for uncultured prokaryotes.

Uncultured prokaryotes	Type of sample (source)	Approach	Remarks	References
*Candidate division OP3	Bacteria (flooded paddy soil)	Fosmid metagenome	OP3 was originally detected in Obsidian Pool hot spring. The 16S rRNA phylogeny suggests that OP3 belongs to the Planctomycetes/Verrucomicrobia/Chlamydiae superphylum. OP3 may have metabolic ability similar to that of Deltaproteobacteria.	Glöckner et al. (2010)
*Candidate phylum TM6	Bacteria (biofilm from hospital restroom sink drain)	Single-cell: Fluorescence-activated cell sorting (FACS), Multiple displacement amplification (MDA).	Proposed a new method –the 'mini-metagenome approach', a combination of single-cell assembly tools and metagenome contigs to assemble the genome. Suggested assembly using SPAdes is better than Velvet-SC or CLC.	McLean et al. (2013)
**Candidatus* Sulcia muelleri DMIN	Bacteria (bacteriome of green sharpshooter-insect)	Single-cell: micromanipulation, MDA, combination of Sanger sequencing, pyrosequencing, and mapping with 454-metagenome data	Complete genome with a GC content of 22.5%	Woyke et al. (2010)
**Thiovulum* sp.	*Epsilonproteobacteria* (phototrophic mats)	Single-cell: microfluidic laser-tweezers, MDA, 454 sequencing	Colourless sulphur bacteria with unusually fast motility at 615 μm/s	Marshall et al. (2012)
Candidate division OP11	Bacteria (anoxic, sulphide, and sulphur-rich spring)	Single-cell: microfluidic, MDA, pyrosequencing	Partial genome of ~270 kb and 46% of total ORFs have no predicted function. Cells are heterotrophic and capable to utilizes cellulose, starch, and lignin.	Youssef et al. (2011)
Candidate division OP9	Bacteria (hot spring sediments)	Single-cell: optical trap and microfluidic device, pyrosequencing	Estimated completeness > 96%. OP9 is able to utilize oligosaccharides from cellulose and hemicellulose. Cell is anaerobic and fermentative.	Dodsworth et al. (2013)
Candidatus 'Caldiarchaeum subterraneum'	Archaea (isolated from a microbial mat at a geothermal water stream of a sub-surface gold mine)	Fosmid metagenome library	*C. subterraneum* represents an uncultured crenarchaeotic group. It harbours an ubiquitin-like protein modifier system with structural motifs similar to eukaryotic system proteins.	Nunoura et al. (2010)
Candidatus Chloracido-bacterium thermophilum	Bacteria (alkaline Siliceous hot springs in Yellowstone National Park)	Genome was purified from a co-culture with *Anoxybacillus* and *Meiothermus* spp.	Composed of two chromosomes (2.6 Mbp and 1.0 Mbp), and both encoded essential genes. Cells are aerobic photoheterotrophs.	Costas et al. (2012)

*The sample was taken from a non-thermophilic source.

5.2. Functional metagenomics for uncultured microorganisms and new applications

At present, reports on thermophilic functional metagenomes are limited. Several reviews and commentaries have focused on metagenomics even though the technology is still in its infancy (Chistoserdova, 2014; Davenport and Tümmler, 2013; Lewin et al., 2013; Simon and Daniel, 2011). Functional metagenomes can be investigated using fosmids, (35–45 kb insert), BACs (~200 kb insert), cosmids (30–42 kb insert), or plasmids (<10 kb) (Fig. 2). Briefly, the bulk genomes are sheared partially, cloned into vectors, and transformed into *E. coli* cells maintained in multiwell plates, after which individual clones are sequenced. We suggest that the biggest disadvantages of this approach are that it is expensive and relatively time-consuming. Nevertheless, the reads generated from such libraries are highly accurate and should be better than the consensus reads obtained from shotgun metagenomes if the accuracy of individual genes of interest is critical, i.e. for structure and function studies of proteins or enzymes. The composition of microbial populations in the environment is very complex, and closely related sister-strains or subspecies often cohabitate. In shotgun metagenome studies, sequence assemblers are unable to differentiate reads generated from two or more homologous sequences. Most assemblers use the consensus nucleotide at a particular position in the pool of reads, resulting in contigs that are represented by the consensus of homologous sequences, with the predominant nucleotide often selected as the 'correct' one. If the determined sequence of a given gene is synthesized and the encoded protein is expressed, the protein may not function properly due to the 'mutated' amino acid caused by the consensus issue. In contrast, with library-based metagenomes (for example, in BAC-libraries) genome fragments are cloned and sequenced. Reads generated from library-based metagenomes are definitely more accurate when compared to those generated from shotgun metagenome assemblies.

Interestingly, library-based metagenome can be used to determine the genomic sequences of uncultivated groups. For instance, the unculturable candidate division OP1 was first detected in the heated Obsidian Pool in the late 1990s (Hugenholtz et al., 1998). Several years later, a metagenome analysis of a Japanese epithermal mine fosmid library was performed, and the putative genome of phylotype OP1 *Candidatus Acetothermus autotrophicum* was first reported (Takami et al., 2012). The authors of this study were able to propose its basic metabolism and potential energy conservation mechanisms, despite the fact that this strain is still unculturable. The team had the good fortune of being able to assemble a putative genome of a *Candidatus Caldiarchaerum subterraneum* from the same fosmid library (Nunoura et al., 2010). This *Candidatus* species represents a unique uncultured crenarchaeotic group. In a separate study, data from a BAC-library metagenome was used to fill the gaps of the *Candidatus Chloracidobacterium thermophilum* genome (Costas et al., 2012). These examples illustrate that sequencing cloned libraries is a viable alternative for tapping unculturable thermophiles.

Examples of interesting enzymes obtained through metagenomics include two lipolytic enzymes from a hot spring (Tirawongsaroj et al., 2008) and thermostable β-glucosidases found in a termite gut and hydrothermal spring metagenomes (Schroder et al., 2014; Wang et al., 2012a). There is an ongoing EU FP7 grant 265992 entitled AMYLOMICS (http://amylomics.org/). The acronym AMYLOMICS stands for amylolytic enzymes captured by targeted metagenomics (Hreggvidsson et al., 2011; Zucko et al., 2013). The project aims to harness novel,

thermostable, and efficient enzymes for starch and carbohydrate applications. A similar idea represented with the MetaBioMe program (http://metasystems.riken.jp/metabiome/), which is designed to explore potential commercial enzymes in 44 public metagenome datasets from 10 different samples; however, no data from thermophilic microbiota datasets are included (Sharma et al., 2009). As another example, the NCBI Bioproject PRJNA238853 entitled 'Thermophilic cellulose-degrading bacterium Metagenome' was initiated with the purpose of discovering thermostable cellulases in a Chinese hot spring. The hot spring was fed with sugarcane bagasse as the sole carbon source, and the metagenome was analysed using pyrosequencing. In the near future, interesting findings with thermostable lignozymes should be released. In short, harnessing new proteins by metagenomics approaches will definitely accelerate new discoveries and applications.

5.3. Advanced sequencing and 'omics' techniques

Currently, many investigators are satisfied with incomplete thermophilic genomes containing several gaps. As technology advances and expectations increase, complete genome sequencing will soon become common, considering the continual rapid development of technology. We think that the Pacific Biosciences PacBio RS II (SMRT) platform can potentially serve the scientific community better than other platforms as it generates long sequencing reads that enable scaffolding of thermophilic prokaryotes; the advantages that we experienced using this platform have been summarised recently (Roberts et al., 2013). However, the cost and maintenance of this sequencing instrument is relatively high. As mentioned above, when using incomplete 16S rRNA sequences, it is difficult to determine the presence of new *Candidatus* species. Generating a full-length 16S rRNA metagenome may be possible using PacBio RS II platform, although it is not commonly used at present, as the system is designed for more complex genomes (Bowman et al., 2013; Fichot and Norman, 2013). As an alternative to the long-read PacBio RS II sequencer, Illumina recently launched a new long-read technology in 2014. In addition, it is exciting that Oxford Nanopores has finally introduced the beta-testing sequencer at a price of $1000 USD. This sequencer is the first mobile DNA sequencer and is 4 inches long (approximately the size of common smart phones); it can be powered using USB port, and no sophisticated, bulky, or expensive sequencer hardware is needed (Eisenstein, 2012). Currently, the portable sequencer is being evaluated under the Oxford Nanopores MinION™ Access Program. In the coming years, Oxford Nanopore may drive a new renaissance in genome, RNA, and protein sequencing. These advanced sequencing approaches should drive advances in whole genome, RNA, and metasequencing studies.

However, omics technologies such as proteomics, transcriptomics, and metabolomics are well established for single-strain analyses. It is worth mentioning the growing trend in adapting these techniques in the study of indigenous microbial communities and merging metaproteome, metatranscriptome, and metametabolome data with metagenome data. Examples of related reports or commentaries can be found in Becher et al. (2013), D'haeseleer et al. (2013), Hanreich et al. (2012), Lü et al. (2014), Turnbaugh and Gordon (2008), and Xia et al. (2014).

5.4. Discovering archaeoviruses and bacteriophages

Thermophilic viruses studies are scarce in nature. Less than 5% of prokaryotic viruses listed in the public domain are originated from thermophilic sources (http://www.ncbi.nlm.nih.gov/genomes/GenomesHome.cgi). To date, the majority of thermophilic viruses are archaeoviruses, although a few thermophilic bacteriophages are known (Satyanarayana et al., 2013). The scarcity of studies focused on thermophilic viruses may be overcome in several ways. Improving the growth of fastidious thermophilic hosts is one possibility. The application of culture-independent techniques to study a viral metagenome, or virome, is a potential alternative. Virome analysis has allowed not only estimation of viral abundance but also discovery of novel viruses (Bolduc et al., 2012; Schoenfeld et al., 2008; Schoenfeld et al., 2010). Nevertheless, the biggest challenge in discovering novel viruses lies in the virus identification processes itself. Current identification is highly dependent on sequence homology. Yet, the vast majority of viral reads show a low number of matching records in existing databases or the function of the matching sequences are unknown. Therefore, innovative analyses that break away from the rigidity of sequence identity are highly valuable. The genomes of prokaryotic viruses show unique mosaic patterns, consistent with a high rate of genome mutations and recombination events. This observation is useful for virus identification. For example, genome signatures that are based on oligonucleotide frequencies in the genomic make-up of known viruses or signatures that are based on sequence distances between viruses and hosts have streamlined virus identification by assigning unknown sequences to the genera of known viruses (Deschavanne et al., 2010; Pride and Schoenfeld, 2008). Although the use of signature sequences is appealing for the identification of bacteriophages, considerable complications may arise when studying archaeoviruses due to their diverse morphotypes. In addition, conservation in protein folding was shown to be useful for determining the putative function of unknown sequences (Dellas et al., 2013; Goulet et al., 2009). It is therefore desirable to consider that such conservation may be applicable in virus identification.

6. Conclusions

In this review, we discussed key advances related to thermophilic bacteria, archaea, and viruses, as well as their bioproducts and related bioprocesses. With the emergence of new sequencing technologies, fundamental knowledge related to thermophilic genomes and metabolism will continue to increase. With the huge amount of data generated from whole-genome sequencing and metagenomics, transcriptomics, metabolomics, secretomics, and other possible -omics information sources, extensive amounts of data are available now and will only continue to accumulate. However, as an increasing number of new prokaryotes and eukaryotes are sequenced, the number of unfamiliar sequences (i.e. hypothetical proteins) increases even more rapidly. The authors agree with the remarks made by Lindahl and Kuske (2013) that 'increased sequence data alone will not advance science'. Many laboratories are moving towards high-end NGS technology and bioinformatics; yet, it is important to continue with traditional laboratory approaches to confirm the functions and properties of the numerous novel sequences by protein biochemistry, crystallisation, mutation analysis, and/or gene knockout approaches. In fact, traditional microorganism isolation shall continue and meta- and genome sequencing cannot replace its role.

A paradigm shift from ecology-motivated sequencing to application-driven metagenome sequencing will lead to the discovery of multitudinous novel gene sequences and hopefully enable the development of new biotechnology applications. Thermostable enzymes for industrial applications should remain one of the main focuses, while biomass and biofuels have also become popular areas of research. Thermophiles are excellent candidates for renewable energy development, bioremediation, biomining, production of biosurfactants, and thermozymes. New applications such as developing antimicrobial agents (i.e. bacteriocins or antibiotics) from thermophiles will be interesting fields in the years to come. Currently, the BAGEL database of bacteriocins is comprehensive and can help discover potentially novel antimicrobial peptides from genome sequences (van Heel et al., 2013). In addition, bioprospecting microbial metagenomes for natural products may also become a topic of increasing interest (Novakova and Farkasovsky, 2013). We conclude with the following question: have we reached a new tipping point in thermophile research and the development of related applications with the help

of genomic tools? We shall leave the answer to this question up to the readers.

Acknowledgements

This work was supported by UTM-CINDEFI Flagship Grant Nos. 01G64 and GUP 06H31 and University of Malaya-Ministry of Higher Education High Impact Research Grants (UM.C/625/1/HIR/MOHE/CHAN/01, Grant No. A-000001-50001 and UM-MOHE HIR Grant UM.C/625/1/HIR/MOHE/CHAN/14/1, Grant No. H-50001-A000027 awarded to Kok-Gan Chan). Drs. Urbieta and Donati gratefully acknowledge support from ANPCyT (PICT 2012 0623 and PICT 2014 0630). We also would like to thank Ummirul Mukminin Kahar and Kian Piaw Chai for their helpful discussions.

References

Abdollahi H, Shafaei SZ, Noaparast M, Manafi Z, Niemelä SI, Tuovinen OH. Mesophilic and thermophilic bioleaching of copper from a chalcopyrite-containing molybdenite concentrate. Int J Miner Process 2014;128:25–32.

Abraham J, Gea T, Sánchez A. Substitution of chemical dehairing by proteases from solid-state fermentation of hair wastes. J Clean Prod 2014;74:191–8.

Alakangas E, Valtanen J, Levlin J-E. CEN technical specification for solid biofuels – Fuel specification and classes. Biomass Bioenergy 2006;30:908–14.

Al-Maghrabi IM, Aqil AB, Islam M, Chaalal O. Use of thermophilic bacteria for bioremediation of petroleum contaminants. Energy Sources. Taylor & Francis; 1999. p. 17–29.

Baker GC, Gaffar S, Cowan DA, Suharto AR. Bacterial community analysis of Indonesian hot springs. FEMS Microbiol Lett 2001;200:103–9.

Barabote RD, Parales JV, Guo Y-Y, Labavitch JM, Parales RE, Berry AM. Xyn10A, a thermostable endoxylanase from *Acidothermus cellulolyticus* 11B. Appl Environ Microbiol 2010;76:7363–6.

Barns SM, Fundyga RE, Jeffries MW, Pace NR. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. Proc Natl Acad Sci 1994;91:1609–13.

Basen M, Rhaesa AM, Kataeva I, Prybol CJ, Scott IM, Poole FL, et al. Degradation of high loads of crystalline cellulose and of unpretreated plant biomass by the thermophilic bacterium *Caldicellulosiruptor bescii*. Bioresour Technol 2014;152:384–92.

Becher D, Bernhardt J, Fuchs S, Riedel K. Metaproteomics to unravel major microbial players in leaf litter and soil environments: challenges and perspectives. Proteomics. Wiley Online Library; 2013. p. 2895–909.

Bergquist LP, W. Morgan H, Saul D. Selected enzymes from extreme thermophiles with applications in biotechnology. Curr Biotechnol 2014;3:45–59.

Bhalla A, Bansal N, Kumar S, Bischoff KM, Sani RK. Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. Bioresour Technol 2013;128:751–9.

Bhandiwad A, Guseva A, Lynd L. Metabolic engineering of *Thermoanaerobacterium thermosaccharolyticum* for increased n-Butanol production. Adv Microbiol 2013;3:46–51.

Bharali P, Das S, Konwar B, Thakur A. Crude biosurfactant from thermophilic *Alcaligenes faecalis*: feasibility in petro-spill bioremediation. Int Biodeter Biodegr 2011;65:682–90.

Blainey PC. The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 2013;37:407–27.

Bolduc B, Shaughnessy DP, Wolf YI, Koonin EV, Roberto FF, Young M. Identification of novel positive-strand RNA viruses by metagenomic analysis of archaea-dominated Yellowstone hot springs. J Virol 2012;86:5562–73.

Boonyaratanakornkit BB, Simpson AJ, Whitehead TA, Fraser CM, El-Sayed N, Clark DS. Transcriptional profiling of the hyperthermophilic methanarchaeon *Methanococcus jannaschii* in response to lethal heat and non-lethal cold shock. Environ Microbiol 2005;7:789–97.

Borges N, Matsumi R, Imanaka T, Atomi H, Santos H. *Thermococcus kodakarensis* mutants deficient in di-*myo*-inositol phosphate use aspartate to cope with heat stress. J Bacteriol 2010;192:191–7.

Bowen De León K, Gerlach R, Peyton BM, Fields MW. Archaeal and bacterial communities in three alkaline hot springs in Heart Lake Geyser Basin, Yellowstone National Park. Front Microbiol 2013;4:330.

Bowman B, Shin MY, Lee J-E, Cho Y-J, Chun J. Analysis of full-length metagenomic 16S genes by SMRT sequencing. Chem 2013;4:C2.

Cai L, Ye L, Tong AHY, Lok S, Zhang T. Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. PLoS One 2013;8:e53649.

Canganella F, Wiegel J. Anaerobic thermophiles. Life 2014;4:77–104.

Chai YY, Rahman RNZRA, Illias RM, Goh KM. Cloning and characterization of two new thermostable and alkalitolerant α-amylases from the *Anoxybacillus* species that produce high levels of maltose. J Ind Microbiol Biotechnol 2012;39:731–41.

Chatterjee S, Bhattacharjee I, Chandra G. Biosorption of heavy metals from industrial waste water by *Geobacillus thermodenitrificans*. J Hazard Mater 2010;175:117–25.

Chernyh NA, Gavrilov SN, Sorokin VV, German KE, Sergeant C, Simonoff M, et al. Characterization of technetium (vII) reduction by cell suspensions of thermophilic bacteria and archaea. Appl Microbiol Biotechnol 2007;76:467–72.

Chistoserdova L. Is metagenomics resolving identification of functions in microbial communities? Microb Biotechnol 2014;7:1–4.

Costas G, Amaya M, Liu Z, Tomsho LP, Schuster SC, Ward DM, et al. Complete genome of *Candidatus* Chloracidobacterium thermophilum, a chlorophyll-based photoheterotroph belonging to the phylum acidobacteria. Environ Microbiol 2012;14:177–90.

Davenport CF, Tümmler B. Advances in computational analysis of metagenome sequences. Environ Microbiol 2013;15:1–5.

Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham DE, et al. The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature 1998;392:353–8.

Deive F, Dominguez A, Barrio T, Moscoso F, Moran P, Longo M, et al. Decolorization of dye Reactive Black 5 by newly isolated thermophilic microorganisms from geothermal sites in Galicia (Spain). J Hazard Mater 2010;182:735–42.

Dellas N, Lawrence CM, Young MJ. A survey of protein structures from archaeal viruses. Life 2013;3:118–30.

Deschavanne P, DuBow MS, Regeard C. The use of genomic signature distance between bacteriophages and their hosts displays evolutionary relationships and phage growth cycle determination. Virol J 2010;7:163.

D'haeseleer P, Gladden JM, Allgaier M, Chain PS, Tringe SG, Malfatti SA, et al. Proteogenomic analysis of a thermophilic bacterial consortium adapted to deconstruct switchgrass. PLoS One 2013;8:e68465.

d'Hugues P, Foucher S, Galle-Cavalloni P, Morin D. Continuous bioleaching of chalcopyrite using a novel extremely thermophilic mixed culture. Int J Miner Process 2002;66:107–19.

Dodsworth JA, Blainey PC, Murugapiran SK, Swingley WD, Ross CA, Tringe SG, et al. Single-cell and metagenomic analyses indicate a fermentative and saccharolytic lifestyle for members of the OP9 lineage. Nat Commun 2013;4:1854.

Donati ER, Sand W. Microbial processing of metal sulfides. Springer; 2007.

Eisenstein M. Oxford Nanopore announcement sets sequencing sector abuzz. Nat Biotechnol 2012;30:295–6.

Elleuche S, Schröder C, Sahm K, Antranikian G. Extremozymes-biocatalysts with unique properties from extremophilic microorganisms. Curr Opin Biotechnol 2014;29:116–23.

Empadinhas N, da Costa MS. Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. Environ Microbiol 2011;13:2056–77.

Falcicchio P, Wolterink-Van Loo S, Franssen MC, van der Oost J. DHAP-dependent aldolases from (hyper) thermophiles: biochemistry and applications. Extremophiles 2014;18:1–13.

Faria TQ, Mingote A, Siopa F, Ventura R, Maycock C, Santos H. Design of new enzyme stabilizers inspired by glycosides of hyperthermophilic microorganisms. Carbohydr Res 2008;343:3025–33.

Fichot EB, Norman RS. Microbial phylogenetic profiling with the Pacific Biosciences sequencing platform. Microbiome 2013;1:10–4.

Fierer N, Leff JW, Adams BJ, Nielsen UN, Bates ST, Lauber CL, et al. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. Proc Natl Acad Sci 2012;109:21390–5.

Fukui T, Atomi H, Kanai T, Matsumi R, Fujiwara S, Imanaka T. Complete genome sequence of the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1 and comparison with *Pyrococcus* genomes. Genome Res 2005;15:352–63.

Futterer O, Angelov A, Liesegang H, Gottschalk G, Schleper C, Schepers B, et al. Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. Proc Natl Acad Sci U S A 2004;101:9091–6.

Giaveno MA, MSa Urbieta, Ulloa JR, EGl Toril, Donati ER. Physiologic versatility and growth flexibility as the main characteristics of a novel thermoacidophilic *Acidianus* strain isolated from Copahue geothermal area in Argentina. Microb Ecol 2013;65:336–46.

Glass EM, Wilkening J, Wilke A, Antonopoulos D, Meyer F. Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. Cold Spring Harb Protoc 2010:5. [pdb.prot5368].

Glöckner J, Kube M, Shrestha PM, Weber M, Glöckner FO, Reinhardt R, et al. Phylogenetic diversity and metagenomics of candidate division OP3. Environ Microbiol 2010;12:1218–29.

Goh KM, Chua YS, Abdul RNZR. A comparison of conventional and miniprimer PCR to elucidate bacteria diversity in Malaysia Ulu Slim hot spring using 16S rDNA clone library. Rom Biotechnol Lett 2011;16:6247–55.

Goh KM, Kahar UM, Chai YY, Chong CS, Chai KP, Ranjani V, et al. Recent discoveries and applications of *Anoxybacillus*. Appl Microbiol Biotechnol 2013;97:1475–88.

Goh KM, Gan HM, Chan K-G, Chan GF, Shahar S, Chong CS, et al. Analysis of *Anoxybacillus* genomes from the aspects of lifestyle adaptations, prophage diversity, and carbohydrate metabolism. PLoS One 2014;9:e90549.

Gorke B, Stulke J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol 2008;6:613–24.

Goulet A, Spinelli S, Blangy S, van Tilbeurgh H, Leulliot N, Basta T, et al. The thermo- and acido-stable ORF-99 from the archaeal virus AFV1. Protein Sci 2009;18:1316–20.

Gudiukaitė R, Gegeckas A, Kazlauskas D, Citavicius D. Influence of N-and/or C-terminal regions on activity, expression, characteristics and structure of lipase from *Geobacillus* sp. 95. Extremophiles 2014;18:131–45.

Guerrero-Barajas C, Ordaz A, Garibay-Orijel C, García-Solares SM, Bastida-González F, Zárate-Segura PB. Enhanced sulfate reduction and trichloroethylene (TCE) biodegradation in a UASB reactor operated with a sludge developed from hydrothermal vents sediments: process and microbial ecology. Int Biodeterior Biodegradation 2014;94:182–91.

Guo F, Ju F, Cai L, Zhang T. Taxonomic precision of different hypervariable regions of 16S rRNA gene and annotation methods for functional bacterial groups in biological wastewater treatment. PLoS One 2013;8:e76185.

Han Y, Agarwal V, Dodd D, Kim J, Bae B, Mackie RI, et al. Biochemical and structural insights into xylan utilization by the thermophilic bacterium *Caldanaerobius polysaccharolyticus*. J Biol Chem 2012;287:34946–60.

Hanreich A, Heyer R, Benndorf D, Rapp E, Pioch M, Reichl U, et al. Metaproteome analysis to determine the metabolically active part of a thermophilic microbial community producing biogas from agricultural biomass. Can J Microbiol 2012;58:917–22.

Hawkes FR, Hussy I, Kyazze G, Dinsdale R, Hawkes DL. Continuous dark fermentative hydrogen production by mesophilic microflora: principles and progress. Int J Hydrog Energy 2007;32:172–84.

Heine M, Chandra SB. The linkage between reverse gyrase and hyperthermophiles: a review of their invariable association. J Microbiol 2009;47:229–34.

Hickey DA, Singer G. Genomic and proteomic adaptations to growth at high temperature. Genome Biol 2004;5:117.

Hongoh Y, Yuzawa H, Ohkuma M, Kudo T. Evaluation of primers and PCR conditions for the analysis of 16S rRNA genes from a natural environment. FEMS Microbiol Lett 2003;221:299–304.

Hou W, Wang S, Dong H, Jiang H, Briggs BR, Peacock JP, et al. A comprehensive census of microbial diversity in hot springs of Tengchong, Yunnan Province China using 16S rRNA gene pyrosequencing. PLoS One 2013;8:e53350.

Hreggvidsson OG, Fridjonsson HO, Petursdóttir KS, Aevarsson A, Altenbuchner J, Fotheringham I, et al. Amyloenzymes captured by targetedmetagenomics. Carbohydr Bioeng 2011.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR. Novel division level bacterial diversity in a Yellowstone hot spring. J Bacteriol 1998;180:366–76.

Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. Genome Res 2007;17:377–86.

Ilyas S. Lee J-c, Kim B-s. Bioremoval of heavy metals from recycling industry electronic waste by a consortium of moderate thermophiles: process development and optimization. J Clean Prod 2014;70:194–202.

Inskeep WP, Rusch DB, Jay ZJ, Herrgard MJ, Kozubal MA, Richardson TH, et al. Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. PLoS One 2010;5:e9773.

Inskeep WP, Jay ZJ, Tringe SG, Herrgård MJ, Rusch DB, Committee YMPS, et al. The YNP metagenome project: environmental parameters responsible for microbial distribution in the Yellowstone geothermal ecosystem. Front Microbiol 2013;4:67–81.

Isenbarger TA, Finney M, Ríos-Velázquez C, Handelsman J, Ruvkun G. Miniprimer PCR, a new lens for viewing the microbial world. Appl Environ Microbiol 2008;74:840–9.

Jabbour D, Sorger A, Sahm K, Antranikian G. A highly thermoactive and salt-tolerant α-amylase isolated from a pilot-plant biogas reactor. Appl Microbiol Biotechnol 2013;97:2971–8.

Jiang B, Liang X, Chen Y, Ma T, Liu L, Li J, et al. Integrating next-generation sequencing and traditional tongue diagnosis to determine tongue coating microbiome. Sci Rep 2012;2:936.

Jiménez DJ, Andreote FD, Chaves D, Montaña JS, Osorio-Forero C, Junca H, et al. Structural and functional insights from the metagenome of an acidic hot spring microbial planktonic community in the Colombian Andes. PLoS One 2012;7:e52069.

Johnson DB, Okibe N, Wakeman K, Yajie L. Effect of temperature on the bioleaching of chalcopyrite concentrates containing different concentrations of silver. Hydrometallurgy 2008;94:42–7.

Jones EJ, Voytek MA, Warwick PD, Corum MD, Cohn A, Bunnell JE, et al. Bioassay for estimating the biogenic methane-generating potential of coal samples. Int J Coal Geol 2008;76:138–50.

Joshi S, Bharucha C, Jha S, Yadav S, Nerurkar A, Desai AJ. Biosurfactant production using molasses and whey under thermophilic conditions. Bioresour Technol 2008;99:195–9.

Kahar UM, Chan K-G, Salleh MM, Hii SM, Goh KM. A high molecular-mass *Anoxybacillus* sp. SK3-4 amylopullulanase: characterization and its relationship in carbohydrate utilization. Int J Mol Sci 2013;14:11302–18.

Kang J, Park K-M, Choi K-H, Park C-S, Kim G-E, Kim D, et al. Molecular cloning and biochemical characterization of a heat-stable type I pullulanase from *Thermotoga neapolitana*. Enzyme Microb Technol 2011;48:260–6.

Kataoka M, Yamaoka A, Kawasaki K, Shigeri Y, Watanabe K. Extraordinary denaturant tolerance of keratinolytic protease complex assemblies produced by *Meiothermus ruber* H328. Appl Microbiol Biotechnol 2014;98:2973–80.

Kembel SW, Wu M, Eisen JA, Green JL. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol 2012;8:e1002743.

Klindworth A, Pruesse E, Schweer T. Peplies Jr, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 2012;41:e1.

Kongjan P, Kotay M, Min B, Angelidaki I. Biohydrogen production from wheat straw hydrolysate by dark fermentation using extreme thermophilic mixed culture. Biotechnol Bioeng 2010;105:899–908.

Kotelnikova S. Microbial production and oxidation of methane in deep subsurface. Earth Sci Rev 2002;58:367–95.

Lasken RS. Genomic sequencing of uncultured microorganisms from single cells. Nat Rev Microbiol 2012;10:631–40.

Lavania M, Cheema S, Sarma PM, Ganapathi R, Lal B. Methanogenic potential of a thermophilic consortium enriched from coal mine. Int Biodeter Biodegr 2014;93:177–85.

Lee J, Lee C, Hugunin K, Maute C, Dysko R. Bacteria from drinking water supply and their fate in gastrointestinal tracts of germ-free mice: a phylogenetic comparison study. Water Res 2010;44:5050–8.

Lewin A, Wentzel A, Valla S. Metagenomics of microbial life in extreme temperature environments. Curr Opin Biotechnol 2013;24:516–25.

Li C, Fang HH. Fermentative hydrogen production from wastewater and solid wastes by mixed cultures. Crit Rev Environ Sci Technol 2007;37:1–39.

Li S, Zhong H, Hu Y, Zhao J, He Z, Gu G. Bioleaching of a low-grade nickel-copper sulfide by mixture of four thermophiles. Bioresour Technol 2014;153:300–6.

Lin L, Xu J. Dissecting and engineering metabolic and regulatory networks of thermophilic bacteria for biofuel production. Biotechnol Adv 2013;31:827–37.

Lindahl BD, Kuske CR. Metagenomics for study of fungal ecology. The Ecological Genomics of Fungi. Wiley Online Library; 2013. p. 279–303.

Liu Z, Klatt CG, Wood JM, Rusch DB, Ludwig M, Wittekindt N, et al. Metatranscriptomic analyses of chlorophototrophs of a hot-spring microbial mat. ISME J 2011;5:1279–90.

Lü F, Bize A, Guillot A, Monnet V, Madigou C, Chapleur O, et al. Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. ISME J 2014;8:88–102.

Margesin R, Schinner F. Biodegradation and bioremediation of hydrocarbons in extreme environments. Appl Microbiol Biotechnol 2001;56:650–63.

Marshall IP, Blainey PC, Spormann AM, Quake SR. A single-cell genome for *Thiovulum* sp. Appl Environ Microbiol 2012;78:8555–63.

McClendon SD, Batth T, Petzold CJ, Adams PD, Simmons BA, Singer SW. *Thermoascus aurantiacus* is a promising source of enzymes for biomass deconstruction under thermophilic conditions. Biotechnol Biofuels 2012;5:1–10.

McLean JS, Lombardo M-J, Badger JH, Edlund A, Novotny M, Yee-Greenbaum J, et al. *Candidate phylum* TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. Proc Natl Acad Sci 2013;110:E2390–9.

Millard AD, Hands-Portman I, Zwirglmaier K. Morphotypes of virus-like particles in two hydrothermal vent fields on the East Scotia Ridge. Antarctia. Bacteriophage 2014;4:e28732.

Meyer-Dombard D, Shock E, Amend J. Archaeal and bacterial communities in geochemically diverse hot springs of Yellowstone National Park, USA. Geobiology 2005;3:211–27.

Mnif S, Chamkha M, Labat M, Sayadi S. Simultaneous hydrocarbon biodegradation and biosurfactant production by oilfield-selected bacteria. J Appl Microbiol 2011;111:525–36.

Mnif S, Sayadi S, Chamkha M. Biodegradative potential and characterization of a novel aromatic-degrading bacterium isolated from a geothermal oil field under saline and thermophilic conditions. Int Biodeter Biodegr 2014;86:258–64.

Mori H, Maruyama F, Kurokawa K. VITCOMIC: visualization tool for taxonomic compositions of microbial communities based on 16S rRNA gene sequences. BMC Bioinformatics 2010;11:332–40.

Murakami S, Fujishima K, Tomita M, Kanai A. Metatranscriptomic analysis of microbes in an oceanfront deep-subsurface hot spring reveals novel small RNAs and type-specific tRNA degradation. Appl Environ Microbiol 2012;78:1015–22.

Namiki T, Hachiya T, Tanaka H, Sakakibara Y. MetaVelvet: an extension of Velvet assembler to *de novo* metagenome assembly from short sequence reads. Nucleic Acids Res 2012;40:e155.

Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, et al. Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 1999;399:323–9.

Novakova J, Farkasovsky M. Bioprospecting microbial metagenome for natural products. Biologia 2013;68:1079–86.

Nunoura T, Takaki Y, Kakuta J, Nishi S, Sugahara J, Kazama H, et al. Insights into the evolution of archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. Nucleic Acids Res 2010;39:3204–23.

Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a web browser. BMC Bioinformatics 2011;12:385–93.

Parisutham V, Kim TH, Lee SK. Feasibilities of consolidated bioprocessing microbes: from pretreatment to biofuel production. Bioresour Technol 2014;161:431–40.

Parmar A, Singh NK, Pandey A, Gnansounou E, Madamwar D. Cyanobacteria and microalgae: a positive prospect for biofuels. Bioresour Technol 2011;102:10163–72.

Peng Y, Leung HC, Yiu S-M, Chin FY. Meta-IDBA: a *de novo* assembler for metagenomic data. Bioinformatics 2011;27:i94–i101.

Peng Y, Leung HC, Yiu S-M, Chin FY. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics 2012;28:1420–8.

Pham VH, Kim J. Cultivation of unculturable soil bacteria. Trends Biotechnol 2012;30:475–84.

Plumb JJ, Haddad CM, Gibson JA, Franzmann PD. *Acidianus sulfidivorans* sp. nov., an extremely acidophilic, thermophilic archaeon isolated from a solfatara on Lihir Island, Papua New Guinea, and emendation of the genus description. Int J Syst Evol Microbiol 2007;57:1418–23.

Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA, et al. A genomic analysis of the archaeal system *Ignicoccus hospitalis-Nanoarchaeum equitans*. Genome Biol 2008;9:R158.

Pride DT, Schoenfeld T. Genome signature analysis of thermal virus metagenomes reveals archaea and thermophilic signatures. BMC Genomics 2008;9:420.

Purcell D, Sompong U, Yim LC, Barraclough TG, Peerapornpisal Y, Pointing SB. The effects of temperature, pH and sulphide on the community structure of hyperthermophilic streamers in hot springs of northern Thailand. FEMS Microbiol Ecol 2007;60:456–66.

Qin W, Yang C, Lai S, Wang J, Liu K, Zhang B. Bioleaching of chalcopyrite by moderately thermophilic microorganisms. Bioresour Technol 2013;129:200–8.

Rastogi G, Muppidi GL, Gurram RN, Adhikari A, Bischoff KM, Hughes SR, et al. Isolation and characterization of cellulose-degrading bacteria from the deep subsurface of the Homestake gold mine, Lead, South Dakota, USA. J Ind Microbiol Biotechnol 2009;36:585–98.

Rastogi G, Bhalla A, Adhikari A, Bischoff KM, Hughes SR, Christopher LP, et al. Characterization of thermostable cellulases produced by *Bacillus* and *Geobacillus* strains. Bioresour Technol 2010;101:8798–806.

Raven N, Ladwa N, Cossar D, Sharp R. Continuous culture of the hyperthermophilic archaeum *Pyrococcus furiosus*. Appl Microbiol Biotechnol 1992;38:263–7.

Roberts RJ, Carneiro MO, Schatz MC. The advantages of SMRT sequencing. Genome Biol 2013;14:405.

Ruepp A, Graml W, Santos-Martinez M-L, Koretke KK, Volker C, Mewes HW, et al. The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. Nature 2000;407:508–13.

Ryu H-W, Yoo S-K, Choi JM, Cho K-S, Cha DK. Thermophilic biofiltration of H2S and isolation of a thermophilic and heterotrophic H2S-degrading bacterium, *Bacillus* sp. TSO3. J Hazard Mater 2009;168:501–6.

Sabath N, Ferrada E, Barve A, Wagner A. Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic streamlining during thermal adaptation. Genome Biol Evol 2013;5:966–77.

Santos H, Da Costa MS. Compatible solutes of organisms that live in hot saline environments. Environ Microbiol 2002;4:501–9.

Sar P, Kazy SK, Paul D, Sarkar A. Metal bioremediation by thermophilic microorganisms. Thermophilic Microbes in Environmental and Industrial Biotechnology. Springer; 2013. p. 171–201.

Saranya P, Kumari HS, Rao BP, Sekaran G. Lipase production from a novel thermo-tolerant and extreme acidophile *Bacillus pumilus* using palm oil as the substrate and treatment of palm oil-containing wastewater. Environ Sci Pollut Res Int 2014;21:3907–19.

Satoh T, Watanabe K, Yamamoto H, Yamamoto S, Kurosawa N. Archaeal community structures in the Solfataric acidic hot springs with different temperatures and elemental compositions. Archaea 2013 2013.

Satyanarayana T, Littlechild J, Kawarabayasi Y. Thermophilic microbes in environmental and industrial biotechnology. biotechnology of thermophiles.. Springer; 2013.

Schoenfeld T, Patterson M, Richardson PM, Wommack KE, Young M, Mead D. Assembly of viral metagenomes from yellowstone hot springs. Appl Environ Microbiol 2008;74:4164–74.

Schoenfeld T, Liles M, Wommack KE, Polson SW, Godiska R, Mead D. Functional viral metagenomics and the next generation of molecular tools. Trends Microbiol 2010;18:20–9.

Schroder C, Elleuche S, Blank S, Antranikian G. Characterization of a heat-active archaeal β-glucosidase from a hydrothermal spring metagenome. Enzyme Microb Technol 2014;57:48–54.

Sen SK, Raut S, Dora TK, Mohapatra PKD. Contribution of hot spring bacterial consortium in cadmium and lead bioremediation through quadratic programming model. J Hazard Mater 2014;265:47–60.

Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M. CAMERA: a community resource for metagenomics. PLoS Biol 2007;5:e75.

Sharafi H, Abdoli M, Hajfarajollah H, Samie N, Alidoust L, Abbasi H, et al. First report of a lipopeptide biosurfactant from thermophilic bacterium *Aneurinibacillus thermoaerophilus* MK01 newly isolated from municipal landfill site. Appl Biochem Biotechnol 2014;173:1236–49.

Sharma VK, Kumar N, Prakash T, Taylor TD. MetaBioME: a database to explore commercially useful enzymes in metagenomic datasets. Nucleic Acids Res 2009;38:D468–72.

Simon C, Daniel R. Metagenomic analyses: past and future trends. Appl Environ Microbiol 2011;77:1153–61.

Singer GA, Hickey DA. Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. Gene 2003;317:39–47.

Sizova M, Izquierdo J, Panikov N, Lynd L. Cellulose- and xylan-degrading thermophilic anaerobic bacteria from biocompost. Appl Environ Microbiol 2011;77:2282–91.

Soliman N, Abdel-Fattah Y, Mostafa HE, Gaballa A. Heterologous expression of thermostable esterase gene from *Geobacillus thermoleovorans* YN under different expression promoters. Int J Environ Sci Technol 2014;11:119–26.

Somboonna N, Assawamakin A, Wilantho A, Tangphatsornruang S, Tongsima S. Metagenomic profiles of free-living archaea, bacteria and small eukaryotes in coastal areas of Sichang island, Thailand. BMC Genomics 2012;13:S29.

Staufenberger T, Imhoff JF, Labes A. First crenarchaeal chitinase found in *Sulfolobus tokodaii*. Microbiol Res 2012;167:262–9.

Stewart EJ. Growing unculturable bacteria. J Bacteriol 2012;194:4151–60.

Su X, Han Y, Dodd D, Moon YH, Yoshida S, Mackie RI, et al. Reconstitution of a thermostable xylan-degrading enzyme mixture from the bacterium *Caldicellulosiruptor bescii*. Appl Environ Microbiol 2013;79:1481–90.

Swingley WD, Meyer-Dombard DR, Shock EL, Alsop EB, Falenski HD, Havig JR, et al. Coordinating environmental genomics and geochemistry reveals metabolic transitions in a hot spring ecosystem. PLoS One 2012;7:e38108.

Tabak HH, Lens P, van Hullebusch ED, Dejonghe W. Developments in bioremediation of soils and sediments polluted with metals and radionuclides-1. Microbial processes and mechanisms affecting bioremediation of metal contamination and influencing metal toxicity and transport. Rev Environ Sci Biotechnol 2005;4:115–56.

Takami H, Noguchi H, Takaki Y, Uchiyama I, Toyoda A, Nishi S, et al. A deeply branching thermophilic bacterium with an ancient acetyl-CoA pathway dominates a subsurface ecosystem. PLoS One 2012;7:e30559.

Talluri S, Raj SM, Christopher LP. Consolidated bioprocessing of untreated switchgrass to hydrogen by the extreme thermophile *Caldicellulosiruptor saccharolyticus* DSM 8903. Bioresour Technol 2013;139:272–9.

Techtmann SM, Colman AS, Robb FT. 'That which does not kill us only makes us stronger': the role of carbon monoxide in thermophilic microbial consortia. Environ Microbiol 2009;11:1027–37.

Thielemann T, Cramer B, Schippers A. Coalbed methane in the Ruhr Basin, Germany: a renewable energy resource? Org Geochem 2004;35:1537–49.

Thomas T, Gilbert J, Meyer F. Metagenomics-a guide from sampling to data analysis. Microb Inform Exp 2012;2:3.

Tirawongsaroj P, Sriprang R, Harnpicharnchai P, Thongaram T, Champreda V, Tanapongpipat S, et al. Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. J Biotechnol 2008;133:42–9.

Treangen TJ, Koren S, Sommer DD, Liu B, Astrovskaya I, Ondov B, et al. MetAMOS: a modular and open source metagenomic assembly and analysis pipeline. Genome Biol 2013;14:R2.

Trivedi S, Rao SR, Gehlot HS. Nucleic acid stability in thermophilic prokaryotes: a review. J Mol Cell Biol 2005;4:61–9.

Tsudome M, Deguchi S, Tsujii K, Ito S, Horikoshi K. Versatile solidified nanofibrous cellulose-containing media for growth of extremophiles. Appl Environ Microbiol 2009;75:4616–9.

Tu Q, Yu H, He Z, Deng Y, Wu L, Nostrand JD, et al. GeoChip 4: a functional gene-array-based high-throughput environmental technology for microbial community analysis. Mol Ecol Resour 2014;14:914–28.

Turnbaugh PJ, Gordon JI. An invitation to the marriage of metagenomics and metabolomics. Cell 2008;134:708–13.

Urbieta MS, Rascovan N, Castro C, Revale S, Giaveno MA, Vazquez M, et al. Draft genome sequence of the novel thermoacidophilic archaeon *Acidianus copahuensis* strain ALE1, isolated from the Copahue volcanic area in Neuquén, Argentina. Genome Announc 2014a;2:e00259-4.

Urbieta MS, Toril EG, Giaveno MA, Bazan AA, Donati ER. Archaeal and bacterial diversity in five different hydrothermal ponds in the Copahue region in Argentina. Syst Appl Microbiol 2014b;37:429–41.

van Heel AJ, de Jong A, Montalbán-López M, Kok J, Kuipers OP. BAGEL3: automated identification of genes encoding bacteriocins and (non-) bactericidal posttranslationally modified peptides. Nucleic Acids Res 2013;41:W448–53.

Vartoukian SR, Palmer RM, Wade WG. Strategies for culture of 'unculturable' bacteria. FEMS Microbiol Lett 2010;309:1–7.

Vick T, Dodsworth JA, Costa K, Shock E, Hedlund BP. Microbiology and geochemistry of Little Hot Creek, a hot spring environment in the Long Valley Caldera. Geobiology 2010;8:140–54.

Walther J, Sierocinski P, van der Oost J. Hot transcriptomics. Archaea 2011;2010.

Wang S. Copper leaching from chalcopyrite concentrates. JOM 2005;57:48–51.

Wang Q, Qian C, Zhang X-Z, Liu N, Yan X, Zhou Z. Characterization of a novel thermostable β-glucosidase from a metagenomic library of termite gut. Enzyme Microb Technol 2012a;51:319–24.

Wang Y, Leung HC, Yiu S-M, Chin FY. MetaCluster 5.0: a two-round binning approach for metagenomic data for low-abundance species in a noisy sample. Bioinformatics 2012b;28:i356–62.

Wang Z, Tong W, Wang Q, Bai X, Chen Z, Zhao J, et al. The temperature dependent proteomic analysis of *Thermotoga maritima*. PLoS One 2012c;7:e46463.

Wemheuer B, Taube R, Akyol P, Wemheuer F, Daniel R. Microbial diversity and biochemical potential encoded by thermal spring metagenomes derived from the Kamchatka Peninsula. Archaea 2013 2013:136714.

Wilson MS, Siering PL, White CL, Hauser ME, Bartles AN. Novel archaea and bacteria dominate stable microbial communities in North America's largest hot spring. Microb Ecol 2008;56:292–305.

Woodward J, Orr M, Cordray K, Greenbaum E. Biotechnology: enzymatic production of biohydrogen. Nature 2000;405:1014–5.

Woyke T, Tighe D, Mavromatis K, Clum A, Copeland A, Schackwitz W, et al. One bacterial cell, one complete genome. PLoS One 2010;5:e10314.

Wu D, Raymond J, Wu M, Chatterji S, Ren Q, Graham JE, et al. Complete genome sequence of the aerobic CO-oxidizing thermophile *Thermomicrobium roseum*. PLoS One 2009;4:e4207.

Xia Y, Wang Y, Fang HH, Jin T, Zhong H, Zhang T. Thermophilic microbial cellulose decomposition and methanogenesis pathways recharacterized by metatranscriptomic and metagenomic analysis. Scientific reports, 4. Nature Publishing Group; 2014.

Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, et al. SOAPdenovo-Trans: *de novo* transcriptome assembly with short RNA-Seq reads. Bioinformatics 2014;30:1660–6.

Youssef NH, Blainey PC, Quake SR, Elshahed MS. Partial genome assembly for a candidate division OP11 single cell from an anoxic spring (Zodletone Spring, Oklahoma). Appl Environ Microbiol 2011;77:7804–14.

Zeldovich KB, Berezovsky IN, Shakhnovich EI. Protein and DNA sequence determinants of thermophilic adaptation. PLoS Comput Biol 2007;3:e5.

Zhaxybayeva O, Swithers KS, Lapierre P, Fournier GP, Bickhart DM, DeBoy RT, et al. On the chimeric nature, thermophilic origin, and phylogenetic placement of the *Thermotogales*. Proc Natl Acad Sci 2009;106:5865–70.

Zhu Z, Tam TK, Sun F, You C, Zhang Y-HP. A high-energy-density sugar biobattery based on a synthetic enzymatic pathway. Nat Commun 2014;5:3026.

Zucko J, Fridjonsson HO, Petursdottir KS, Gacesa R, Diminic J, Long FP, et al. Browsing metagenomes for novel starch and carbohydrate industry enzymes-AMYLOMICS case study. Curr Opin Biotechnol 2013;245:521–7.