Bioprospecting of Novel Extremozymes From Prokaryotes—The Advent of Culture-Independent Methods

Maksim Sysoev[1], Stefan W. Grötzinger[1], Dominik Renn[1], Jörg Eppinger[1,2], Magnus Rueping[1,2]* and Ram Karan[1]*

[1] KAUST Catalysis Center (KCC), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia,
[2] Institute for Experimental Molecular Imaging, University Clinic, RWTH Aachen University, Aachen, Germany

Extremophiles are remarkable organisms that thrive in the harshest environments on Earth, such as hydrothermal vents, hypersaline lakes and pools, alkaline soda lakes, deserts, cold oceans, and volcanic areas. These organisms have developed several strategies to overcome environmental stress and nutrient limitations. Thus, they are among the best model organisms to study adaptive mechanisms that lead to stress tolerance. Genetic and structural information derived from extremophiles and extremozymes can be used for bioengineering other nontolerant enzymes. Furthermore, extremophiles can be a valuable resource for novel biotechnological and biomedical products due to their biosynthetic properties. However, understanding life under extreme conditions is challenging due to the difficulties of *in vitro* cultivation and observation since > 99% of organisms cannot be cultivated. Consequently, only a minor percentage of the potential extremophiles on Earth have been discovered and characterized. Herein, we present a review of culture-independent methods, sequence-based metagenomics (SBM), and single amplified genomes (SAGs) for studying enzymes from extremophiles, with a focus on prokaryotic (archaea and bacteria) microorganisms. Additionally, we provide a comprehensive list of extremozymes discovered via metagenomics and SAGs.

Keywords: extremophile, culture-independent methods, metagenomics, SAG, halophiles, thermophiles, psychrophile

INTRODUCTION

A mounting paradigm shift toward using sustainable resources has stimulated exploring new efficient approaches in technological processes (Raddadi et al., 2015; Lamers et al., 2016; Krüger et al., 2018). Enzymes, as natural catalysts, have shown remarkable abilities that have revolutionized the chemical, biotechnological, bioremediation, agricultural, and pharmaceutical industries (Martin and Vandenbol, 2016; Wiltschi et al., 2020). However, the narrow range of stability of most described biocatalysts from mesophilic organisms limits their use for many applications (Raddadi et al., 2015). Enzymes derived from microorganisms thriving under harsh conditions, called extremophiles, can overcome these restrictions, and today, such biocatalysts are in higher demand than ever before (Karan et al., 2012a; Rizk et al., 2012; Johnson, 2014; Raddadi et al., 2015; Sarmiento et al., 2015; Singh et al., 2016; Jorquera et al., 2019). The diversity of

extreme environments promises to reveal biocatalysts capable of withstanding harsh industrial conditions, providing better efficiency with the lower environmental burden (Elleuche et al., 2014; Grötzinger et al., 2017). Extremophiles are present in all three domains of life (bacteria, archaea, and eukarya) (Rothschild and Mancinelli, 2001; Jorquera et al., 2019). According to their natural habitats, extremophiles are classified into thermophiles, alkaliphiles, acidophiles, halophiles, and others (Karan et al., 2012a,b; Raddadi et al., 2015). Therefore, compared to their mesophilic equals, extremozymes are usually able to perform reactions under a broader range of conditions (Karan et al., 2011; Reed et al., 2013; Sarmiento et al., 2015). This extended activity range often allows extremophiles to identify as polyextremophiles, as they tolerate multiple extreme conditions (Karan and Khare, 2010; Rekadwad and Khobragade, 2017; Karan et al., 2019). Therefore, polyextremophiles are perfect candidates as a source of novel enzymes for industrial needs (Raddadi et al., 2015; Sarmiento et al., 2015; Lamers et al., 2016; Krüger et al., 2018).

However, only a small percentage of all microorganisms can be grown in a laboratory environment. This is particularly the case for archaea, representing a large group of prokaryotic extremophiles (Kaeberlein et al., 2002; Zengler et al., 2002; Buerger et al., 2012). Improving computing and bioinformatics technologies have made it possible to study the "dark matter" of the microbial world by looking at the genome data extracted from microbial habitats (Hedlund et al., 2014).

The advent of next-generation sequencing allowed researchers to sequence thousands of microorganisms in parallel, and due to the high sensitivity of these genome-based methods, now we can study microbes with a very low abundance that may be overlooked by other methods (Grötzinger et al., 2014; Laurence et al., 2014; Karan et al., 2020). Accordingly, sequence-based metagenomics (SBM) was born, which studies microorganisms by randomly shearing environmental DNA, sequencing it, and assembling the reads (Hugenholtz and Tyson, 2008).

Further advances allowed researchers to study the genome from single cells (Grötzinger et al., 2014, 2017; Akal et al., 2019; Karan et al., 2019; Vogler et al., 2020). Single amplified genome (SAG) technology separates individual cells before analyzing their DNA, thus giving us information about each cell's genome instead of bulk metagenomes. SAG technology specifically allows whole-genome assemblies from small-sample volumes with low cell yields and low cell abundance compared to those of other cells within a given sample. This technology has proven to be especially useful for studying extremophiles, as their environment makes their cultivation and consequent genomic study particularly complicated (Karan et al., 2019). As the coherent successor of the SBM approach, SAG solved some bottlenecks by introducing a cell-sorting step, making the subsequent sequence analysis of a complex sample easier and more straightforward (Grötzinger et al., 2014). Although SBM and SAG technologies have been available for years, the expected boost in biotechnology has not been realized (Martin and Vandenbol, 2016). Despite the progress in next-generation sequencing technologies, relatively few new extremozymes have been discovered and functionally characterized using culture-independent methods. This shortfall is mainly, in addition to the high sequencing costs, because of the lack of reliable, functional annotation of the genomic data caused by the low amount (0.09%) of experimentally described genes (Grötzinger et al., 2014, 2017; Attrill et al., 2019).

Generally, annotation algorithms depend on existing functional annotations, thus giving reliable results for already well-studied branches of life.

Therefore, culture-based methods are still needed to study a novel enzyme's function and structure, confirmed by several studies (Martin and Vandenbol, 2016; Speda et al., 2017).

Nevertheless, the lack of a specific database limits our choice for the right expression model to study the enzyme of interest. Consequently, the combination of these issues creates a vicious loop resulting in low detection and usage of extremozymes. This review aims to highlight the importance of extremozyme from prokaryotes research for industry by showing how recent developments in sequencing, computation and bioinformatics solve the problems associated with data mining for extremozymes and outline the annotation difficulties, which must be overcome in the future.

ENZYMES: A SUSTAINABLE SOURCE FOR GREEN CHEMISTRY

Enzymes are broadly used in biotechnology and a variety of industries (e.g., agriculture, food, textiles, chemicals, pharmaceuticals, and biofuels) as catalysts, therapeutic agents, analytic reagents, and diagnostic tools (Illanes et al., 2012; Robinson, 2015; Ramesh et al., 2020). Enzymes have broadened the horizon of potential applications by allowing us to perform chemo- and regioselective reactions, which is a big struggle for current chemical techniques (Rasor and Voss, 2001; Schäfer et al., 2007). In general, enzymatic reactions are safer, faster, less hazardous, and generate less waste, thus following the twelve rules of green chemistry (Anastas and Eghbali, 2010). Especially today, in a world with 7.5 billion people and an expected 9.8 billion 30 years from now (United Nations, Department of Economic and Social Affairs, Population Division, 2017), it is crucial to utilize the high potential of enzymes for biotechnological applications and green chemistry to reduce humanity's overconsumption of resources (Anastas and Eghbali, 2010; Sheldon, 2016).

Although enzymes have vast potential in biotechnological applications, they have been used only in very few specific reactions. Most described enzymes can be used only for a limited number of industrial processes (Herbert, 1992; Elleuche et al., 2014). This limitation is caused by the narrow ranges of enzymatic stability, including a majority of essential parameters for chemical reactions, such as temperature, pressure, pH, and the use of organic solvents. For example, although water is considered the solvent of life, it is a poor solvent for synthetic reactions (Aitken and Brown, 1969). Organic solvents are used not only to increase the solubility of hydrophobic substrates but also to shift the thermodynamic equilibrium from hydrolysis

to condensation and suppress water-dependent side reactions (Carrea and Riva, 2000).

A number of genetic (Gatti-Lafranconi et al., 2010) and chemical (Siddiqui et al., 2009) modifications together with immobilization strategies (Mukhopadhyay et al., 2015) have been developed to overcome these restrictions (Stepankova et al., 2013). These strategies aim to increase the enzyme's stability or decrease the denaturing effect of the reaction conditions. To date, there is no ultimate mechanism to increase the stability of enzymes. A point mutation can drastically affect the enzyme's efficiency and stability, although it is difficult to see any apparent trends or patterns to fully control a given enzyme (Vieille and Zeikus, 2001; Sarmiento et al., 2015). It is essential to identify mutagenesis strategies to alter enzymes from mesophilic organisms to achieve stability in extreme physicochemical conditions, which is a challenging task (Reed et al., 2013). Although various stabilizing immobilization strategies exist, several hurdles, such as a limited increase in stability, change in reaction rate, alteration of stabilities under other conditions, and potential influence on the specificity of the enzymes, have not allowed these approaches to be used in a broader range (Coker, 2016).

EXTREMOPHILES

An alternative approach to achieve biocatalysis in extreme physicochemical conditions is to use enzymes derived from organisms that thrive in extreme conditions (Jorquera et al., 2019). These organisms are called extremophiles and live in harsh environments of elevated temperatures (thermophiles), salt concentration (halophiles), pressure (barophiles), osmotic compound content (osmophiles), heavy metal content (metalophiles), and radiation (radiophiles); acidic or basic pH (acid or alkaliphiles); extreme dryness (xerophiles); extreme cold (psychrophiles); or a combination of different extremes (polyextremophiles) (**Figure 1**; Kristjánsson and Hreggvidsson, 1995; Gerday and Glansdorff, 2009; Gabani and Singh, 2013; Coker, 2016; Horikoshi, 2016).
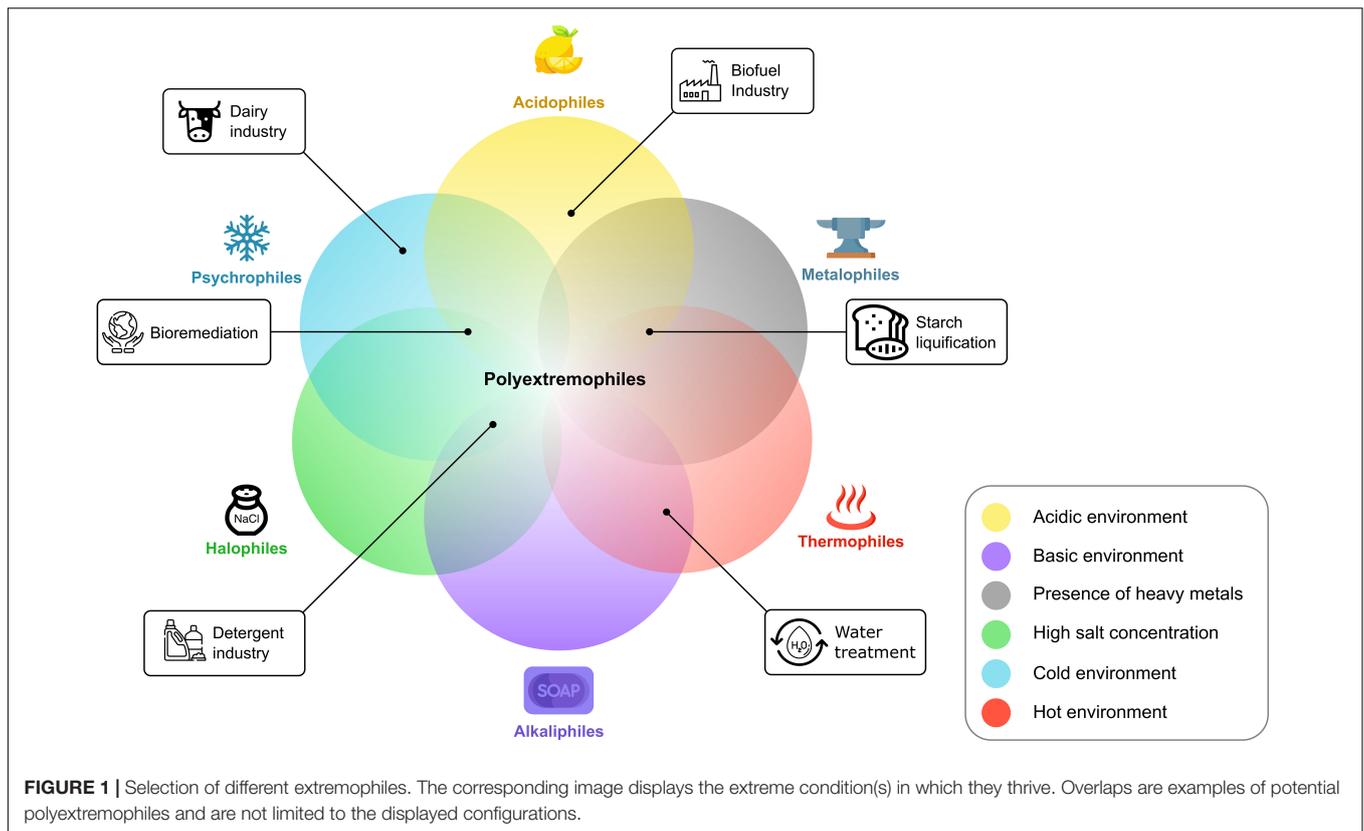
Halophiles and Xerophiles

Halophiles are microorganisms that require elevated salt concentrations to grow. There are four main classifications of halophiles based on their optimal NaCl concentration for growth: slight halophiles (0.2 M), moderate halophiles (0.5–2.5 M), borderline extreme halophiles (2.5–4.0 M), and extreme halophiles (4.0–5.9 M). Halophilic organisms use several adaptation mechanisms to live in these conditions. These include modified electrostatic charge of their proteins, balancing osmotic pressure either by compatible solutes, such as betaine and ectoines, or with chloride and potassium uptake into the cells by transporters (primary or secondary) and the coordinated action of bacteriorhodopsin and ATP synthase (Karan et al., 2012a).

Halophilic organisms and their enzymes are used in several different fields, including production of fermented foods, manufacturing of solar salt from seawater, leather industries, environmental bioremediation and textile, pharmaceutical

(Karan and Khare, 2010; Oren, 2010; Karan et al., 2012a; de Lourdes Moreno et al., 2013; Akal et al., 2019). The biochemical properties of the molecules, enzymes, and compatible solutes synthesized by halophiles also present potential implications in fine chemicals, medicines, and bioimplants (Irwin, 2010; Karan et al., 2012a,b; Chen and Jiang, 2018; Jin et al., 2019). One of the main selling points of these enzymes is their stability in such solvents as benzene, toluene or chloroform, which are frequently used in different industrial environments (Oren, 2010). Halophiles are a potential source of novel extremozymes like amylases, proteases, nucleases, cellulases, chitinases, xylanases, esterases, alcohol dehydrogenases and lipases (Karan et al., 2012a; Coker, 2016). For example, nuclease from *Micrococcus varians* has long been used commercially for production of a flavoring agent 5′-guanylic acid (5′-GMP) due to its efficiency in degrading RNA at 60°C and 12% (w/v) salt (Kamekura et al., 1982). Several other enzymes have been heterologously expressed and characterized (**Table 1**). Interestingly, some of the halophilic enzymes display polyextremophilicity i.e., stability toward more than one extreme condition e.g., high salt, elevated or low temperature, alkaline or acidic pH and non-aqueous medium having great potential application for industrial and biotechnological processes (Grötzinger et al., 2017; Akal et al., 2019; Karan et al., 2020).

Along with the enzymes, halophiles possess a number of bioactive molecules that found its application in different areas. The biocompatible solute ectoine has potential use in respiratory medicine (Galinski et al., 1985) and can reduce cell stress effects in nanoparticle-induced lung inflammation by inhibiting the signals (Sydlik et al., 2009). Some halophiles, such as *Halobacterium* and *Haloferax*, were found to accumulate polyhydroxyalkanoates (PHAs), which are a family of biopolyesters with diverse uses in the medical, environmental, and industrial fields (Kirk and Ginzburg, 1972; Quillaguaman et al., 2010; Chen and Patel, 2012). Halophiles, such as *Natronococcus occultus* and *Naloterrigena hispanica,* have been found to generate antimicrobial peptides and diketopiperazines (Charlesworth and Burns, 2015; Coker, 2016). Some haloarchaea produce internal protein gas vesicle nanoparticles (GVNPs), which have been tested as a vaccine scaffold platform and ultrasonic contrast agents (Stuart et al., 2001; Shapiro et al., 2014; DasSarma and DasSarma, 2015). Notably, halophilic nature makes cell lysis, hence farming of the enzymes and other molecules, easier. Since halophiles require high salt to maintain their cell integrity, lowering the salt concentration in the buffer system disrupts the cell wall of most halophiles.

Xerophiles are microorganisms that can survive and grow in arid conditions with water activity $a_w < 0.75$ by forming spores that help them mitigate environmental stress. Adaptive mechanisms are connected to water loss prevention and increased water retention through the accumulation of compatible solutes, production of extracellular polymeric substances (EPSs), adaptations on the cell membrane to retain intracellular water, and synthesis of DNA repair proteins (Lebre et al., 2017). These unique adaptations allow xerophiles to be used in microbial electrochemical systems (Dopson et al., 2015) or in next-generation industrial biotechnology, where they can

FIGURE 1 | Selection of different extremophiles. The corresponding image displays the extreme condition(s) in which they thrive. Overlaps are examples of potential polyextremophiles and are not limited to the displayed configurations.

be used for treating long-chain fatty acids, cellulose, chitin, rubbers, or other compounds (Chen and Jiang, 2018).

Thermophiles and Psychrophiles

Thermophiles are heat-loving microorganisms with an optimum growth temperature of 45°C or above, while hyperthermophiles grow at temperatures above 80°C (Sarmiento et al., 2015). Thermophiles and hyperthermophiles exist in various natural ecosystems, such as geothermal waters, hot springs, volcanoes, deep-sea hydrothermal vents, and other ecosystems with high-temperature parameters. Thermophiles have thermostable proteins and cell membranes that do not become denatured at high temperatures, and some may also resist proteolysis (Sarmiento et al., 2015). Notably, polymerases from thermophiles have laid the foundation for the discovery of polymerase chain reaction (PCR), a technique that has become crucial in medicine and research. Today we may find various enzymes from thermophiles (e.g., from *Thermus aquaticus* and *Pyrococcus furiosus*) finding their use in PCR due to their stability and reasonable cost (Brock, 1997; Bruins et al., 2001; Irwin and Baird, 2004). Along with polymerases, a lot of different thermophilic enzymes, such as lipases, laccases, and xylanases, are also on the market, thus making the industrial process much more environmentally friendly (Damiano et al., 2003; Atalah et al., 2019).

Psychrophiles are cold-loving microorganisms that can grow at temperatures between −20 and 20°C. Psychrophiles possess diverse adaptive molecular mechanisms to survive and thrive at such low temperatures. Psychrophilic bacteria have increased (i) unsaturated fatty acids, cyclopropane-containing fatty acids, and short-chain fatty acids in their membranes, which prevent the loss of membrane fluidity; (ii) cold-shock proteins (CSPs) and chaperones to protect the synthesis of RNA and proteins; (iii) antifreeze proteins (AFPs) that bind to ice crystals and create a state of thermal hysteresis; and (iv) mannitol and other compatible solutes that act as cryoprotectants to prevent cell damage by ultraviolet (UV) radiation and ice formation (Sarmiento et al., 2015; De Santi et al., 2016a).

Psychrophiles have a promising future in pharmaceuticals and medicine since their cell membranes hold surfactants capable of sustaining stability at low temperatures (Cavicchioli et al., 2011). Psychrophiles, such as *Pandalus borealis*, *Euphausia superba*, *Moraxella* species, and *Flavobacterium* species, have been found to produce anticancer and antitumor agents (Margesin and Feller, 2010). Psychrophilic enzymes display high catalytic activity, stability at low temperatures, and pronounced heat lability and may offer useful industrial and biotechnological applications in various domains, such as pharmaceutical science, molecular biology, textiles, paper, food, feed technologies, detergents, and cosmetics (Margesin and Feller, 2010; Cavicchioli et al., 2011; Sarmiento et al., 2015; Dhaulaniya et al., 2019; Al-Ghanayem and Joseph, 2020).

Acidophiles and Alkaliphiles

Acidophiles are microorganisms that grow at an optimum pH < 3 (Baker-Austin and Dopson, 2007). Acid-tolerant

TABLE 1 | List of extremozymes discovered with metagenomics and single amplified genomes.

Enzyme	Source	Heterologous host	Screening approach
	Function-based metagenomics (FBM)		
β-Glucosidase	Hydrothermal spring	*E. coli*	Schröder et al., 2014
Carboxylesterase	Hot vent sediment	*E. coli*	Placido et al., 2015
Esterase	Glacier soil	*E. coli*	De Santi et al., 2015, 2016a,b
Esterase	Hot spring water, Sediment, and compost	*E. coli, Thermus Thermophilus*	Leis et al., 2015
Esterase	Hot spring water	*E. coli*	López-López et al., 2015
Hormone-sensitive lipase	Permafrost	*E. coli*	Petrovskaya et al., 2016
Lipase	Hot spring sediment	*E. coli*	Sahoo et al., 2017
Lipase	Hot spring water	*E. coli*	Yan et al., 2017
Phospholipase	Hydrothermal vent	*E. coli*	Fu et al., 2015
Serine protease	Hot spring sediment	*E. coli*	Singh et al., 2015
	Sequence-based metagenomics (SBM)		
α-Galactosidase	Hot spring water	*E. coli*	Schröder et al., 2014
Endoglucanase	Hot spring sediment	*E. coli*	Zhao et al., 2017
β-Galactosidase	Hot spring water	*E. coli*	Liu et al., 2015
Esterase	Hot spring mud	*E. coli*	Zarafeta et al., 2016
Mercuric reductase	Brine pool	*E. coli*	Sayed et al., 2014
Amine transferase	Hot spring	*E. coli*	Ferrandi et al., 2017
Epoxide hydrolase	Hot spring	*E. coli*	Ferrandi et al., 2015
	Single amplified genomes (SAG)		
Alcohol dehydrogenases	Brine pool	*Hfx. volcanii*	Grötzinger et al., 2017; Akal et al., 2019
Carbonic anhydrase	Brine pool	*Halobacterium* sp. NRC-1	Vogler et al., 2020
Glucose dehydrogenase	Brine pool	*Halobacterium* sp. NRC-1	Grötzinger et al., 2014; Karan et al., 2019
2-hydroxy dehydrogenase	Brine pool	*Halobacterium* sp. NRC-1	Grötzinger et al., 2014; Karan et al., 2019
Protease	Brine pool	*Halobacterium* sp. NRC-1*, E. coli*	Grötzinger et al., 2014; Karan et al., 2019

microbes have optimum growth at pH > 5 but are still active in lower pH environments. Acidophiles maintain their cytoplasmic pH close to neutrality to protect acid-labile cellular constituents by active pumping of protons (proton flux system); by decreased permeability of the cell membrane, which helps to suppress the entry of protons into the cytoplasm; and by improved protein and DNA repair systems compared to those of neutrophiles (Baker-Austin and Dopson, 2007). Acidophilic enzymes can block the activity of matrix metallopeptidases (MMPs), which are essential for tumor metastasis (Irwin, 2010). MMP inhibitors from an acidophilic *Penicillium* species isolated from Berkeley Pit Lake promise a therapeutic approach for cancer (Stierle et al., 2006). Proteolytic enzymes from acidophiles have also been reported as nonallergenic preservatives in medicines (Sharma et al., 2012). Amylolytic enzymes, such as trehalase isolated from acidophilic *Sulfolobus solfataricus,* are used in medicine as preservatives and stabilizers (Schiraldi et al., 2002).

Alkaliphiles are microorganisms that grow in alkaline environments with a pH > 9, usually showing optimal growth at pH ~10. Alkaliphiles may coexist with neutrophiles under mild basic pH conditions and live in specific extreme environments. Alkaliphilic bacteria possess molecular mechanisms that compromise the activation of both symporter and antiporter systems. Electrogenic antiporters produce an electrochemical gradient of Na^+ and H^+, and the symporter system enables the uptake of Na^+ and other solutes into the cells

(Chinnathambi, 2015). Alkaliphilic enzymes have found use in several applications, such as tannery water treatment (Li et al., 2017), food, cosmetics, and pharmaceutical production (Horikoshi, 2016).

Radiophiles

Radiophiles or radiation-resistant extremophiles thrive in high oxidative stress and radiation environments, including UV radiation, gamma radiation and X-ray radiation, and have potential applications in therapeutics pharmacology and biotechnology. Direct intense or prolonged exposure to different forms of radiation, such as UV radiation, can lead to mutagenic and cytotoxic DNA lesions, resulting in various types of human cancers (Gabani and Singh, 2013). Primary and secondary metabolic products from radiophiles can protect the organism's DNA and can be used to manufacture anticancer drugs, antioxidants, and sunscreens (Raddadi et al., 2015). Extremolytes, mycosporin-like amino acids (MAAs) from the red alga *Porphyra rosengurttii*, are commercially utilized to enhance the UV-protective properties of sunscreens (de la Coba et al., 2009) and are therapeutic candidates as preventive agents in UV radiation-induced cancers, such as melanoma. Bacterioruberin isolated from *Halobacterium* and *Rubrobacter* and deinoxanthin isolated from *Deinococcus radiodurans* are other therapeutic candidates for cancer diseases (Singh and Gabani, 2011; Choi et al., 2014).

Polyextremophiles

Polyextremophiles i.e., microorganisms growing preferentially under multiple extremes, have developed features that allow them to thrive in harsh environments such as Deep Lake Antarctica, where the temperature reaches as low as $-20°C$ and stays liquid only due to extreme salt concentrations (DasSarma et al., 2013; Karan et al., 2013, 2020). Deep-sea anoxic brine pools at the bottom of the Red Sea are another type of polyextremophilic environment (Antunes et al., 2011). These brine pools are extreme in different physicochemical parameters and vary drastically, with temperatures ranging from 22.6 to $68.2°C$ and NaCl concentrations varying from 2.6 to 5.6 M (Antunes et al., 2011). Additionally, they show a characteristic sharp brine-seawater interface, with steep gradients of dissolved oxygen, density, pH, salinity, and temperature (Emery et al., 1969; Ross, 1972; Anschutz and Blanc, 1995). Because of this variation, brine pools offer a multitude of habitats for different kinds of extremophiles.

Polyextremophiles have a diverse range of uses and applications. Halothermophiles and halopsychrophiles are promising sources of useful enzymes (Elleuche et al., 2014). Enzymes derived from these polyextremophilic microbes possess particularly attractive properties for biotechnology, namely, their function at high salt concentrations and high or low temperatures (Sarmiento et al., 2015). Elevated temperatures are used to shift equilibria, distill products, increase reaction's speed, liquefy compounds, and eliminate microbial contamination (Grötzinger et al., 2017). Low temperatures can save energy (e.g., in washing processes, bioremediation, or food processing), avoid labile or volatile compound production (e.g., in biotransformations or food processing) or prevent bacterial growth. Therefore, cold-active enzymes have great potential for various biotechnological processes. Increasing the salt concentration in solution decreases the water activity, thus mimicking aqueous-organic solvent mixtures. Therefore, halophilic enzymes generally retain high activity and stability in high salt and organic or nonaqueous media (Karan and Khare, 2010, 2011; Karan et al., 2011, 2012a). In turn, organic solvents increase the solubility of hydrophobic substrates and alter the hydrolytic and kinetic equilibria (Sellek and Chaudhuri, 1999). Traditional mesophilic enzymes lose their native structure and thus catalytic activity in organic solvents, limiting their use. However, the high stability of halophilic enzymes toward salt is associated with tolerance to low water activity, such as in mixtures of aqueous and organic or nonaqueous media (Sellek and Chaudhuri, 1999), which emphasizes their high potential for biocatalysis. Cold-active or heat-stable enzymes operating in high-salt or organic solvents are of interest for the sustainable production of value-added chemicals. The world is currently looking for microorganisms with new enzymes, such as hydrolases, amylases, cellulases, peptidases, and lipases (Dumorne et al., 2017). Some examples include alkalithermophilic serine proteases from *Alkalibacillus* sp. NM-Da2, which can potentially be applied in different biotechnological and pharmaceutical industries (Abdel-Hamed et al., 2016), and alkalipsychrophilic esterase from the marine bacterium *Rhodococcus* sp., which can be used in the food industry since the process catalyzed by an esterase can be stopped by increasing the temperature, thus saving food properties (De Santi et al., 2014). The polyextremophilic characteristics of halophilic amylases potentially make them efficient catalysts under alkaline pH and high salinity in processes such as starch hydrolysis and applications such as detergent production, the food industry, and bioremediation (Ali et al., 2014; Kumar et al., 2016; Rekadwad and Khobragade, 2017).

MINING ENZYMES FROM EXTREME ENVIRONMENTS

Enzymes derived from extremophiles, so-called extremozymes, can catalyze chemical reactions in harsh conditions, such as those found in industrial processes (Sarmiento et al., 2015). The current global market for industrial enzymes was 9.9 billion USD in 2019 and is expected to grow to 14.9 billion USD by 2027 (Grand View Research, 2020)[1], in which novel extremozymes could play a significant role and further expand this market.

Extremophiles thrive in extreme habitats, including salt lakes, deep-sea vents, acidic sulfurous lakes, alkaline lakes, hot springs, Arctic and Antarctic waters, and alpine lakes (Kristjánsson and Hreggvidsson, 1995; Gerday and Glansdorff, 2009). The challenges in mining the enzymatic potential of extremophilic locations are the hurdle of mimicking these harsh conditions in the laboratory (Sarmiento et al., 2015), the remote areas, and the unusually low cell density, causing a minimal amount of biomass yields (Ferrer et al., 2007). As a result, despite the high scientific potential and industrial value of extremozymes, very little is known about their structure and function. Therefore, most enzymes currently used in industry originate from either fungi or mesophilic bacteria (Elleuche et al., 2014).

Microbial Dark Matter—The Big Unknown

According to recent estimations, Earth is home to an upward range of 10^{12} (1 trillion) microbial species, with microorganisms being the most abundant, widespread, and taxonomically, metabolically, and functionally diverse organisms (Locey and Lennon, 2016). During the past decade, high-throughput sequencing in combination with advanced bioinformatics algorithms has allowed for enhanced insight into microbial taxa and expanded the estimation of the global microbial load by orders of magnitude via projects such as the "Earth Microbiome project" (EMP), which analyzed more than 200,000 environmental samples (Gilbert et al., 2014). Out of the 10^{12} microbial species, only approximately 10^5 have been sequenced (0.00001%), 10^4 have been cultured, and even the EMP, which also uses rRNA sequences for identification, has cataloged fewer than 10^7 species (0.001%), of which 29% were detected only twice (Locey and Lennon, 2016). It is estimated that approximately 99.999% of microbial taxa remain undiscovered (Locey and Lennon, 2016). Focusing on the ~10^4 cultivated species and looking at the highest order, among the 60 significant lines of

---

[1] https://www.grandviewresearch.com/industry-analysis/enzymes-industry, last access date: 2/1/2021 4:43 p.m.

descent (phyla or divisions) that are known within the archaeal and bacterial domain (Hugenholtz and Kyrpides, 2009), 50% are uncultured and make up the "microbial dark matter" (Marcy et al., 2007). Furthermore, 88% of all isolated microbes are members of only four bacterial phyla, *Proteobacteria, Firmicutes, Actinobacteria*, and *Bacteroidetes* (Rinke et al., 2013).

The Challenge of Cultivation

Traditional cultural methods employ the ability of microorganisms to grow in a specific laboratory environment. Historically, these methods have been used in microbiology, including selective or differential media, microscopy, Gram-staining, and biochemical tests. These methods are sensitive, reliable, inexpensive, and provide qualitative and quantitative results on the bacterial populations (Nowrotek et al., 2019). However, only a small fraction of all microbes can be grown in a laboratory setting. Thus, although cultural methods are a well-established, simple, and inexpensive way of isolating, detecting, and quantifying microorganisms, the methods lack speed and are labor-intensive with the risk of contamination. Based on this legacy, microbial culturomics has emerged as a tool that complements metagenomics data and gives another way to determine composition of microbial populations. Several studies have shown that culture-dependent and culture-independent methods often deliver different results (Steven et al., 2007; Carraro et al., 2011; Stefani et al., 2015; Nowrotek et al., 2019; Rego et al., 2019). For example, using 212 different culture conditions, such as temperature, various oxygen levels, and selective media, many researchers have isolated 340 different bacterial species (Lagier et al., 2012). Interestingly, the metagenomics approach could identify only 51 species out of 340 cultured species in this study. Therefore, it is crucial to combine both culture-dependent and culture-independent ways to study a given microbiome (Steven et al., 2007; Lagier et al., 2016; Nowrotek et al., 2019; Sarhan et al., 2019).

The enormous lack of isolates is based on the fact that very few microbes are cultivable, e.g., only as few as 0.001–0.1% of the microbes found in seawater can be cultivated under laboratory conditions (Amann et al., 1995). Therefore, archaea enzymes are typically studied after expression in a heterologous host, mostly *Escherichia coli* (*E. coli*). However, this approach is impossible for some archaeal proteins due to low expression rates, inaccurate protein folding, or the lack of functionally necessary posttranslational protein modifications in eubacterial hosts. Besides, it is challenging to identify archaeal enzymes directly from genome samples in bacterial host-based functional screening assays, as archaeal promoter structures differ from bacterial ones (Zweerink et al., 2017). A direct *in vivo* investigation of archaeal enzymes in an archaeal strain or an archaeal host with the right expression machinery of the respective proteins represents a valuable alternative. For this reason, some archaeal expression models have been used, such as *Halobacterium salinarum*/NRC-1 or *Haloferax volcanii*. Using these expression systems, researchers have studied archaeal TATA-binding proteins, discovered

16S gene transfer between bacteria and archaea (Baliga et al., 2000; Fuchsman et al., 2017), and several different archaeal enzymes have been studied using these archaeal expression systems (**Table 1**). However, this approach requires technically sensitive and straightforward technologies to detect active enzymes and a methodology to express proteins in an archaeal host heterologously. Ultimately, as extremophiles preferentially live in harsh conditions, it is difficult to satisfy all of these microbes' extreme requirements. However, with the progress in computational methods and improving bioinformatics algorithms, it is now possible to study the mysterious "dark matter" of the microbial world using culture-independent approaches, including bulk metagenomics and SAGs.

Function-Based Metagenomics (FBM), Sequence-Based Metagenomics (SBM), and Single Amplified Genomes (SAGs) as a Gateway to Novel Species

Culture-independent methods developed and utilized over the last few years, such as metagenomics, have recently gained momentum. Metagenomics analyses rely on the direct isolation of genomic DNA from the environment. These analyses can be either sequence-based (i.e., putative enzymes are discovered based on their conserved sequences) or function-based (i.e., functional enzymes are found based on the expressed features, such as a specific enzyme activity).

Function-Based Metagenomics (FBM)

The function-based metagenomics (FBM) approach is based on cloning random environmental DNA into expression hosts, such as *E. coli,* forming a library. This library is then screened for function by different assay methods (Rashid and Stingl, 2015). The main challenge of this method is the need for hosts to create functional expression libraries and sufficient screening methods (Rashid and Stingl, 2015). The alternative is the sequence-based approach, which historically was based on sequence homology and used a colony hybridization technique to screen metagenomic clones using an oligonucleotide primer or probes for the target gene. In a subsequent step, the desired gene was amplified by PCR and cloned into suitable expression vectors. This technique led to discovering novel sequences similar to existing known sequences and may help find enzymes more efficiently (Raghunathan et al., 2005; Marcy et al., 2007).

However, metagenomics approaches give little information about expression and regulation of genes in the environment. Recently, metatranscriptomics approaches have been used to reveal gene expression profiles and ecophysiology of natural microbial communities (Hua et al., 2015; Tripathy et al., 2016). Combined genomic and transcriptomic analyses can discover and characterize the relative transcriptional levels of multiple genes and disclose the functional diversity in microbial communities. Such approach already has been utilized for studying extremophiles in acidophilic conditions and provided important data on *Ferrovum* population and its metabolic potentials and gene expression profile (Tripathy et al., 2016).

In the end, metatranscriptomics approach allowed researchers to peak into the ecological role of these unexplored, but potentially important microorganisms.

Sequence-Based Metagenomics (SBM)

SBM became more attractive approach with the advent of high-throughput sequencing, when the main biases and bottlenecks of environmental sampling along with the requirement of DNA cloning before sequencing were removed (Bohmann et al., 2014). Today, the DNA from metagenomic samples can be sequenced without cloning. Furthermore, sometimes it is possible to extract desired gene sequences directly from metagenomics data and synthesize the genes *de novo* after bioinformatic annotation and codon optimization if required. These developments shifted the bottleneck of culture-independent sequencing, mainly to bioinformatics-based sequence alignment, data management, and, in particular, annotation algorithms. Annotation algorithms initially relied on sequence homology-based methods and, therefore, depended on existing functional annotations, giving excellent results only for already well-studied life branches. Therefore, it is advisable to use additional methods to account for this point. Different annotation approaches arose over time, still based on existing functional annotations but offering more degrees of freedom regarding the DNA sequences. Big data analysis requires tremendous computing powers, and for a long time, computers were not ready to handle that kind of load (Fan et al., 2014). Since introduction of the first supercomputer in 1964, the performance has dramatically increased, and now the fastest supercomputer is more than 100 times more powerful than those from a decade ago (**Figure 2A**). Along with the development of powerful computers and advanced computing algorithms, the cost of sequencing has been continuously decreasing, and in the last two decades, it has dropped by a factor of over one million (Wetterstrand, 2013; **Figure 2B**). Increased availability opened up multiple possibilities to study genomes, and several technologies appeared, including first-, second-, and third-generation sequencing. Due to the rapid development of next-generation sequencing, shotgun and high-throughput versions of SBM have become more popular. The SBM method allows for the sequencing of environmental DNA (Thomas et al., 2012) by randomly shearing all metagenomic DNA from one sample, a method depending on the cloning step into vectors for amplification, and sequencing the vector DNA, followed by assembly via algorithms (Rashid and Stingl, 2015).
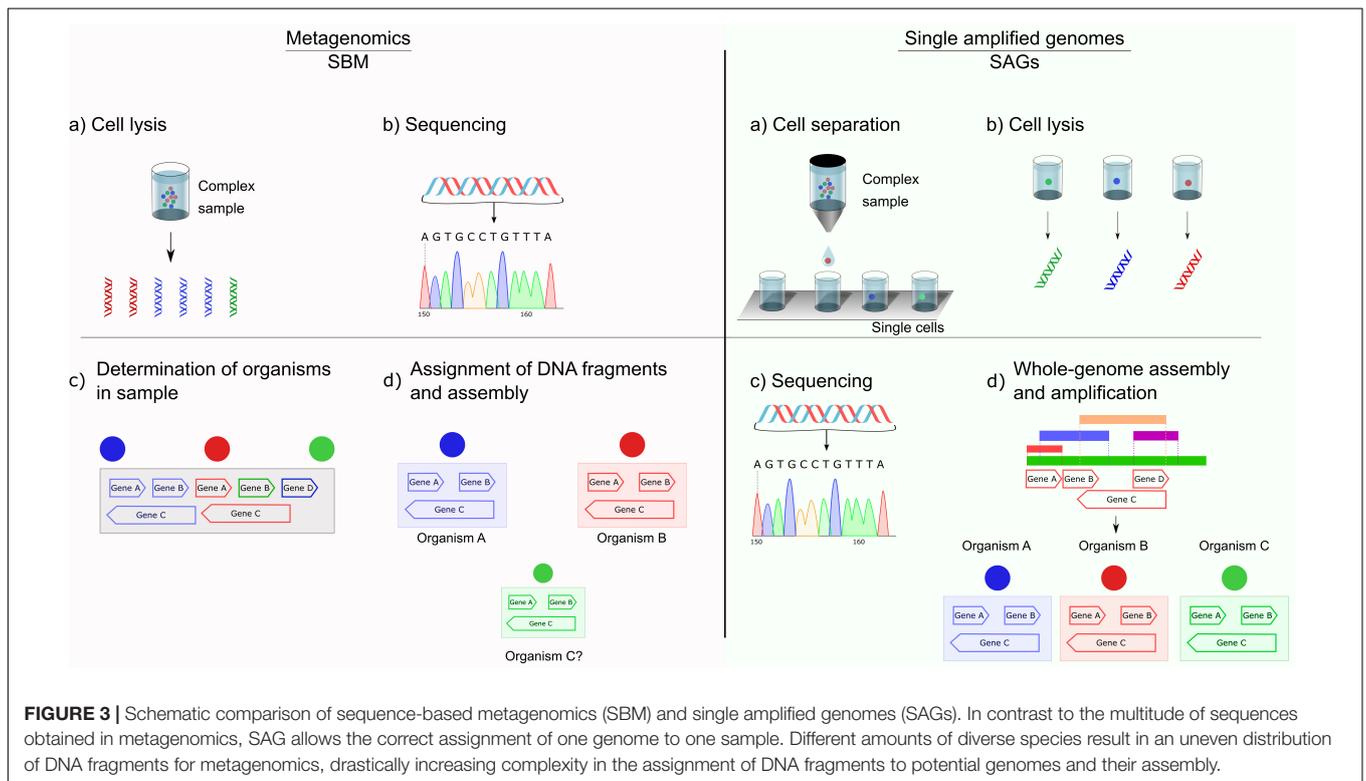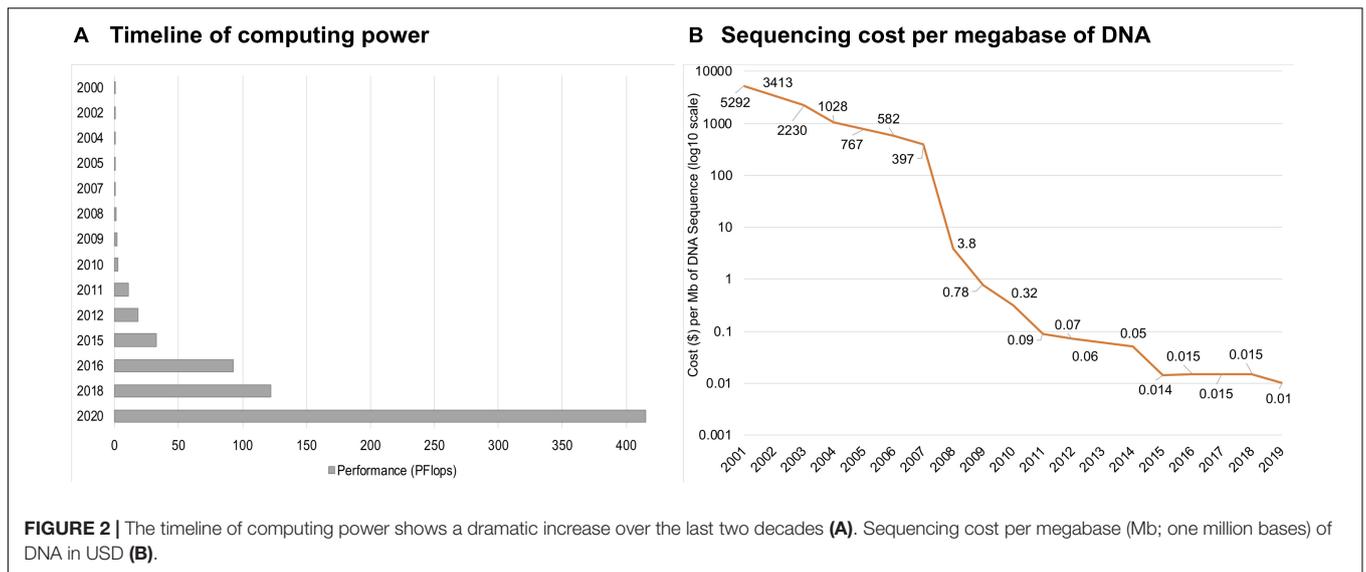
Despite many benefits, SBM also suffers from limitations and challenges. Not only does the short read length of the DNA sequencing and the size and complexity of the sequence data pose analytic and informatics challenges (Rashid and Stingl, 2015), but also the ability to assemble independent genomes is drastically reduced at both very low abundances and increasing genomic heterogeneity (Albertsen et al., 2013). These challenges result in a low probability of identifying rare populations, occurring in less than 1% of the total metagenomic DNA (Kunin et al., 2008). This effect comes mainly from cross-strain assemblies (Rinke et al., 2014) that originate from different organisms but can also come from unwanted host DNA or environmental

contamination (Schmieder and Edwards, 2011) that cannot be avoided. Hence, despite high expectations, metagenomics data have not yet led to the expected boost in biotechnology (Chistoserdova, 2010).

Single Amplified Genomes (SAGs)

SAGs gain the momentum due to the breakthrough of single-cell genome sequencing that has alleviated many bottlenecks in metagenomics by physically separating the genomic material of uncultured cells (Raghunathan et al., 2005; Kvist et al., 2007; Marcy et al., 2007). The **Figure 3** shows a schematic comparison of the critical steps in SBM and SAGs screening using SAGs. In recent years, single-cell genome sequencing has become a highly accessible tool (Chi, 2014). In part, this utility is due to the improvements that have taken place in the protocols for DNA or RNA isolation, leading to more simplified procedures. Additionally, several industrial kits have become available, leading to the amplification of genetic material from single cells (Chi, 2014). To prepare samples for further operations, they need to be separated from the bulk cell mass. There are several different methods for sample preparation of single cells for single-cell genomics (SCG) (Blainey, 2013). These methods include serial dilution (Zhang et al., 2006), micromanipulation (Woyke et al., 2010), optofluidics (optical tweezing in conjunction with microfluidics) (Landry et al., 2013), laser-capture microdissection of tissue samples (Frumkin et al., 2008) and fluorescence-activated cell sorting (FACS) (Swan et al., 2011; Dupont et al., 2012; McLean et al., 2013; Wilson et al., 2014), ultimately resulting in individually sorted samples enriched with a particular microbe into 96-well plates, with a minimal amount of either or both host and environmental contamination. FACS has become the most popular method due to its high performance and the ability to separate individual environmental cells based on various cellular properties (e.g., size, fluorescence, and granularity). Moreover, FACS can also be used to study populations (Thompson et al., 2013). The enriched pool of symbiont cells can undergo whole-genome amplification (WGA) followed by sequencing, yielding a population genome assembly or a homogeneous draft assembly in a clonal population. However, FACS has some technical problems for the sorting of microbial cells. Because it is impossible to confirm cell identity visually, sometimes non-cellular fluorescent particles present in environmental samples can be sorted along with targeted microbial cells (Davey and Kell, 1996; Müller and Nebe-Von-Caron, 2010). For the same reason, FACS also retains a low efficiency in recovering rare cells (Yamamura et al., 2005; Yoshimoto et al., 2013; Nakamura et al., 2016).

After sorting, cells need to be lysed (e.g., with an alkaline solution, detergent, or heat). Furthermore, to obtain DNA contigs, the sample undergoes multiple displacement DNA amplification to sequence genomic DNA from a single microorganism. Subsequently, microorganisms, or unamplified single DNA fragments can be screened in each well using PCR with specific gene primers. WGA can also be performed using bacteriophage Φ29 polymerase to identify specific genes in a genome. Finally, it is vital to monitor DNA contamination to

FIGURE 2 | The timeline of computing power shows a dramatic increase over the last two decades **(A)**. Sequencing cost per megabase (Mb; one million bases) of DNA in USD **(B)**.

FIGURE 3 | Schematic comparison of sequence-based metagenomics (SBM) and single amplified genomes (SAGs). In contrast to the multitude of sequences obtained in metagenomics, SAG allows the correct assignment of one genome to one sample. Different amounts of diverse species result in an uneven distribution of DNA fragments for metagenomics, drastically increasing complexity in the assignment of DNA fragments to potential genomes and their assembly.

ensure that all reagents and equipment used are decontaminated (Giddings and Newman, 2015).

The Challenge of Gene Function Annotation

Although the accessibility of uncultured species' genomes has increased, computational challenges hinder discovering novel extremozymes using culture-independent methods. These hindrances are not u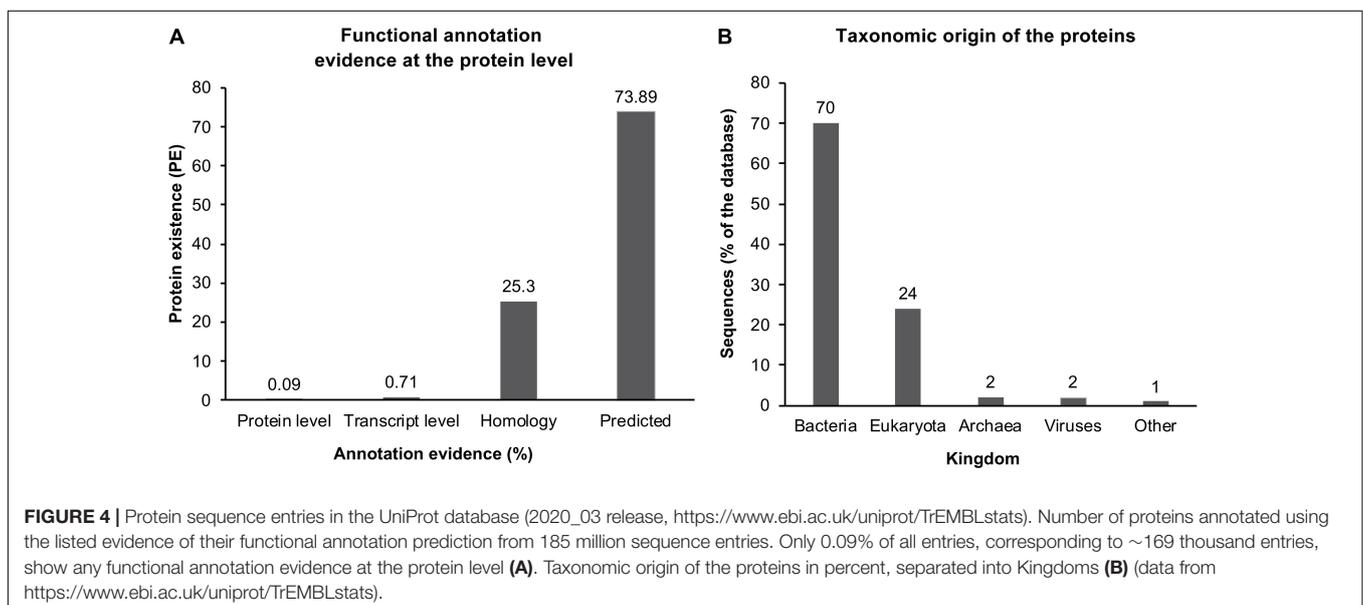nique to extremophiles and affects all organisms, albeit every group has its own unique hurdles contributing to the major cause. The major challenge in mining uncultured organisms' genomic data is the reliable large-scale annotation of DNA sequences from non-mesophilic organisms with low sequence homology to experimentally described organisms. Annotation by homology faces a significant dilemma: annotation reliability is reciprocally coupled to protein diversity. Therefore, the difference between gene sequences negatively correlates with the overall homology to any related gene. Moreover, studying extremozymes amplifies not only in

culture-based approaches but also culture-independent methods. First reason is linked to the fact that extremophiles tend to alter the overall amino acid composition of enzymes to adapt to their environment, which can be observed in halophiles and thermophiles (Madern et al., 2000; Sterner and Liebl, 2001; Berezovsky and Shakhnovich, 2005; Radestock and Gohlke, 2011; Siglioccolo et al., 2011; Reed et al., 2013; Karan et al., 2020). The second reason is mainly ought to the very limited amount of characterized extremozymes. For example, alcohol dehydrogenases and γ-carbonic anhydrase (CA_D) discovered from uncharacterized archaea collected from brine pool at the bottom of the Red Sea showed sequence homology of about 30–37% to the nearest mesophilic homologs (Grötzinger et al., 2017; Akal et al., 2019; Vogler et al., 2020).

Error propagation further complicates the situation, especially in enzymes with few or highly diverged homologs. The function of the encoded protein can be proven experimentally only for a small and continuously decreasing fraction of gene sequences available from databases, which is currently 0.09% of the UniProt database (**Figure 4A**). Furthermore, more than 94% of all sequences originate from eukaryotes or bacteria (**Figure 4B**). Therefore, the challenge of a reliable annotation is even more pronounced for genes that originate from other organisms, such as archaea, including most extremophiles (Rampelotto, 2013). Hence, the more distinct a protein is, smaller the probability of being related to an experimentally described enzyme. At the same time, the more distinct an enzyme is, the higher is its impact on future annotations. For example, a gene from the newly discovered mesophilic microorganism can be compared to many homologs. The function of the majority of these homologs will be used as annotation of the new gene. In contrast, a gene from a newly discovered extremophilic microorganism might have far fewer homologs and therefore, the annotation of these homologs plays a significant role for future studies. Recent developments in annotation algorithms allow a higher

degree of freedom between DNA sequences that will be annotated compared to the closest homologs. These methods include algorithms that focus on specific and short conserved sequences only, such as essential amino acids of the active center along with conserved regions such as catalytic sites and cofactor binding sites (Grötzinger et al., 2014). More elegant but substantially more complicated and computational resource-demanding are structure-based approaches. Here, the expression product of potential genes is modeled, and its function is assessed by its structure, not by the sequence. The biennial Critical Assessment of protein Structure Prediction (CASP) is a competition where their efficiency compares novel modeling algorithms to predict unusual and novel tertiary structures of proteins whose crystal and/or NMR structure was measured but not yet published. The overall *in silico* modeling success was limited until the rise of AI-based protein folding prediction. Last year, in 2020, a new version of a deep machine learning-based algorithm called AlphaFold 2 was released (Senior et al., 2020). This algorithm has substantially increased computational protein folding prediction potential and showed impressive results in the CASP assessment. AlphaFold 2 reached 92.4 out of 100 Global Distance Test (GDT), whereas the average crystal structure only gets 90. Thus, at least for the test samples, it was as good as a crystal structure. Such an impressive result gives hope for the future computational study of unknown proteins. In theory, it could mean that any given protein's structure and function could be predicted solely based on its amino acid sequence in the future. Still, today only a limited amount of algorithms have experimentally been shown to be efficient in the annotation of far distant related genes (Grötzinger et al., 2014; Akal et al., 2019; Vogler et al., 2020). Reliable annotation of the entire genome of an organism that is very distantly related to described organisms is not yet available.

SAGs reduce data complexity significantly compared to that of the metagenomics approach, but the analysis and management of next-generation whole-genome sequencing (NGS) data include

FIGURE 4 | Protein sequence entries in the UniProt database (2020_03 release, https://www.ebi.ac.uk/uniprot/TrEMBLstats). Number of proteins annotated using the listed evidence of their functional annotation prediction from 185 million sequence entries. Only 0.09% of all entries, corresponding to ~169 thousand entries, show any functional annotation evidence at the protein level **(A)**. Taxonomic origin of the proteins in percent, separated into Kingdoms **(B)** (data from https://www.ebi.ac.uk/uniprot/TrEMBLstats).

an impressive number of various software applications. These applications are used for sequence read assembly, mapping to the reference genome, variant/SNP calling and annotation, transcript assembly and quantification, and small RNA identification (Horner et al., 2010; Garber et al., 2011; Pabinger et al., 2013). Compared to classical genomic sequencing, experimental characterization of SAG gene products requires gene synthesis, expression, purification, and functional characterization and, therefore, is several orders of magnitude more time-consuming and cost-intensive. Hence, false-positive results from flawed annotation are much more problematic than false-negative results (due to incomplete annotation) when genomic data are searched for the desired function, which is particularly true for genes from extremophilic organisms that require slow-growing expression systems (Grötzinger et al., 2014).

Therefore, the initial challenge of cultivating organisms can be addressed to a considerable degree by using novel culture-independent methods. However, these advance shifts the problem toward bioinformatics-based handling and interpretation of the data, where developments are rapidly progressing but still need time and data, to allow reliable, functional annotation of whole genomes. In the **Table 1**, a comprehensive list of extremozymes discovered with metagenomics (Berini et al., 2017) and SAGs are shown.

CONCLUSION/OUTLOOK

In the past decades, DNA sequence analysis has made tremendous progress with advances in sequencing technologies and the rapid development of data analysis algorithms. Through these developments, metagenomic studies of complex microbial communities have become more straightforward and more applicable. Given that most microbes cannot be cultivated in a laboratory, obtaining genome information directly from the environment is a massive step forward for genome profiling of

unknown microbial communities. At the time, when read lengths are long enough to allow a confident assembly of genomes, SBM will become the method of choice for bioprospecting non-cultivable microorganisms. For now, the implementation of single-cell genomics has facilitated recovering proteins from uncultured microorganism, helped us understand the species diversity, and how enzymes adapt to harsh environments. Recent advances in developing new sequencing technologies with longer reads, higher throughput, and better cell sorting methods have certainly enhanced SCG as a tool for studying uncultivable microbes. Consequently, this relatively cost-efficient technology is essential to broaden the knowledge of extremophiles and extremozymes, which are very distantly related to the majority of described organisms. We believe that single-cell technologies will not only shed further light on microbial "dark matter" but also facilitate the development of *in silico*-designed and fine-tailored biocatalysts for specific reactions, thus bringing us closer to a sustainable future.

AUTHOR CONTRIBUTIONS

All authors contributed to and commented on this manuscript.

REFERENCES

Abdel-Hamed, A. R., Abo-Elmatty, D. M., Wiegel, J., and Mesbah, N. M. (2016). Biochemical characterization of a halophilic, alkalithermophilic protease from *Alkalibacillus* sp. NM-Da2. *Extremophiles* 20, 885–894. doi: 10.1007/s00792-016-0879-x

Aitken, D. M., and Brown, A. D. (1969). Citrate and glyoxylate cycles in the halophil, Halobacterium salinarium. *Biochim. Biophys. Acta* 177, 351–354. doi: 10.1016/0304-4165(69)90148-2

Akal, A. L., Karan, R., Hohl, A., Alam, I., Vogler, M., Grötzinger, S. W., et al. (2019). A polyextremophilic alcohol dehydrogenase from the Atlantis II deep red sea brine pool. *FEBS Open Bio* 9, 194–205. doi: 10.1002/2211-5463.12557

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K. L., Tyson, G. W., and Nielsen, P. H. (2013). Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* 31, 533–538. doi: 10.1038/nbt.2579

Al-Ghanayem, A. A., and Joseph, B. (2020). Current prospective in using cold-active enzymes as eco-friendly detergent additive. *Appl. Microbiol. Biotechnol.* 104, 2871–2882. doi: 10.1007/s00253-020-10429-x

Ali, I., Akbar, A., Yanwisetpakdee, B., Prasongsuk, S., Lotrakul, P., and Punnapayak, H. (2014). Purification, characterization, and potential of saline waste water

remediation of a polyextremophilic α-amylase from an obligate halophilic *Aspergillus gracilis*. *BioMed Res. Int.* 2014, 106937.

Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in-situ detection of individual microbial-cells without cultivation. *Microbiol. Rev.* 59, 143–169. doi: 10.1128/mmbr.59.1.143-169.1995

Anastas, P., and Eghbali, N. (2010). Green chemistry: principles and practice. *Chem. Soc. Rev.* 39, 301–312. doi: 10.1039/b918763b

Anschutz, P., and Blanc, G. (1995). Geochemical dynamics of the Atlantis II Deep (Red Sea): silica behavior. *Mar. Geol.* 128, 25–36. doi: 10.1016/0025-3227(95)00085-d

Antunes, A., Ngugi, D. K., and Stingl, U. (2011). Microbiology of the Red Sea (and other) deep-sea anoxic brine lakes. *Environ. Microbiol. Rep.* 3, 416–433. doi: 10.1111/j.1758-2229.2011.00264.x

Atalah, J., Cáceres-Moreno, P., Espina, G., and Blamey, J. M. (2019). Thermophiles and the applications of their enzymes as new biocatalysts. *Bioresour. Technol.* 280, 478–488. doi: 10.1016/j.biortech.2019.02.008

Attrill, H., Gaudet, P., Huntley, R. P., Lovering, R. C., Engel, S. R., Poux, S., et al. (2019). Annotation of gene product function from high-throughput studies using the gene ontology. *Database* 2019, baz007.

Baker-Austin, C., and Dopson, M. (2007). Life in acid: pH homeostasis in acidophiles. *Trends Microbiol.* 15, 165–171. doi: 10.1016/j.tim.2007.02.005

Baliga, N. S., Goo, Y. A., Ng, W. V., Hood, L., Daniels, C. J., and DasSarma, S. (2000). Is gene expression in *Halobacterium* NRC-1 regulated by multiple TBP and TFB transcription factors? *Mol. Microbiol.* 36, 1184–1185. doi: 10.1046/j.1365-2958.2000.01916.x

Berezovsky, I. N., and Shakhnovich, E. I. (2005). Physics and evolution of thermophilic adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12742–12747. doi: 10.1073/pnas.0503890102

Berini, F., Casciello, C., Marcone, G. L., and Marinelli, F. (2017). Metagenomics: novel enzymes from non-culturable microbes. *FEMS Microbiol. Lett.* 364, fnx211. doi: 10.1093/femsle/fnx211

Blainey, P. C. (2013). The future is now: Single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* 37, 407–427. doi: 10.1111/1574-6976.12015

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., et al. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends Ecol. Evol.* 29, 358–367. doi: 10.1016/j.tree.2014.04.003

Brock, T. D. (1997). The value of basic research: discovery of *Thermus aquaticus* and other extreme thermophiles. *Genetics* 146, 1207–1210. doi: 10.1093/genetics/146.4.1207

Bruins, M. E., Janssen, A. E., and Boom, R. M. (2001). Thermozymes and their applications: a review of recent literature and patents. *Appl. Biochem. Biotechnol.* 90, 155–186. doi: 10.1385/abab:90:2:155

Buerger, S., Spoering, A., Gavrish, E., Leslin, C., Ling, L., and Epstein, S. S. (2012). Microbial scout hypothesis and microbial discovery. *Appl. Environ. Microbiol.* 78, 3229–3233. doi: 10.1128/aem.07308-11

Carraro, L., Maifreni, M., Bartolomeoli, I., Martino, M. E., Novelli, E., Frigo, F., et al. (2011). Comparison of culture-dependent and -independent methods for bacterial community monitoring during Montasio cheese manufacturing. *Res. Microbiol.* 162, 231–239. doi: 10.1016/j.resmic.2011.01.002

Carrea, G., and Riva, S. (2000). Properties and synthetic applications of enzymes in organic solvents. *Angew. Chem. Int. Ed. Engl.* 39, 2226–2254. doi: 10.1002/1521-3773(20000703)39:13<2226::aid-anie2226>3.0.co;2-l

Cavicchioli, R., Charlton, T., Ertan, H., Omar, S. M., Siddiqui, K., and Williams, T. (2011). Biotechnological uses of enzymes from psychrophiles. *Microb. Biotechnol.* 4, 449–460. doi: 10.1111/j.1751-7915.2011.00258.x

Charlesworth, J. C., and Burns, B. P. (2015). Untapped resources: biotechnological potential of peptides and secondary metabolites in archaea. *Archaea* 2015, 282035. doi: 10.1155/2015/282035

Chen, G.-Q., and Jiang, X.-R. (2018). Next generation industrial biotechnology based on extremophilic bacteria. *Curr. Opin. Biotechnol.* 50, 94–100. doi: 10.1016/j.copbio.2017.11.016

Chen, G. Q., and Patel, M. K. (2012). Plastics derived from biological sources: present and future: a technical and environmental review. *Chem. Rev.* 112, 2082–2099. doi: 10.1021/cr200162d

Chi, K. R. (2014). Singled out for sequencing. *Nat. Methods* 11, 13–17. doi: 10.1038/nmeth.2768

Chinnathambi, A. (2015). Industrial important enzymes from alkaliphiles–an overview. *Biosci. Biotechnol. Res. Asia* 12, 2007–2016. doi: 10.13005/bbra/1868

Chistoserdova, L. (2010). Recent progress and new challenges in metagenomics for biotechnology. *Biotechnol. Lett.* 32, 1351–1359. doi: 10.1007/s10529-010-0306-9

Choi, Y. J., Hur, J. M., Lim, S., Jo, M., Kim, D. H., and Choi, J. I. (2014). Induction of apoptosis by deinoxanthin in human cancer cells. *Anticancer Res.* 34, 1829–1835.

Coker, J. A. (2016). Extremophiles and biotechnology: current uses and prospects. *F1000Res.* 5:F1000 Faculty Rev-396. doi: 10.12688/f1000research.7432.1

Damiano, V., Bocchini, D., Gomes, E., and Da Silva, R. (2003). Application of crude xylanase from *Bacillus licheniformis* 77-2 to the bleaching of eucalyptus Kraft pulp. *World J. Microbiol. Biotechnol.* 19, 139–144.

DasSarma, S., Capes, M. D., Karan, R., and DasSarma, P. (2013). Amino acid substitutions in cold-adapted proteins from *Halorubrum lacusprofundi*, an extremely halophilic microbe from Antarctica. *PLoS One* 8:e58587. doi: 10.1371/journal.pone.0058587

DasSarma, S., and DasSarma, P. (2015). Gas vesicle nanoparticles for antigen display. *Vaccines* 3, 686–702. doi: 10.3390/vaccines3030686

Davey, H. M., and Kell, D. B. (1996). Flow cytometry and cell sorting of heterogeneous microbial populations: the importance of single-cell analyses. *Microbiol. Rev.* 60, 641–696. doi: 10.1128/mr.60.4.641-696.1996

de la Coba, F., Aguilera, J., de Galvez, M. V., Alvarez, M., Gallego, E., Figueroa, F. L., et al. (2009). Prevention of the ultraviolet effects on clinical and histopathological changes, as well as the heat shock protein-70 expression in mouse skin by topical application of algal UV-absorbing compounds. *J. Dermatol. Sci.* 55, 161–169. doi: 10.1016/j.jdermsci.2009.06.004

de Lourdes Moreno, M., Pérez, D., García, M. T., and Mellado, E. (2013). Halophilic bacteria as a source of novel hydrolytic enzymes. *Life* 3, 38–51. doi: 10.3390/life3010038

De Santi, C., Altermark, B., Pierechod, M. M., Ambrosino, L., De Pascale, D., and Willassen, N. P. (2016a). Characterization of a cold-active and salt tolerant esterase identified by functional screening of Arctic metagenomic libraries. *BMC Biochem.* 17:1. doi: 10.1186/s12858-016-0057-x

De Santi, C., Ambrosino, L., Tedesco, P., de Pascale, D., Zhai, L., Zhou, C., et al. (2015). Identification and characterization of a novel salt-tolerant esterase from a *Tibetan glacier* metagenomic library. *Biotechnol. Prog.* 31, 890–899. doi: 10.1002/btpr.2096

De Santi, C., Tedesco, P., Ambrosino, L., Altermark, B., Willassen, N.-P., and de Pascale, D. (2014). A new alkaliphilic cold-active esterase from the psychrophilic marine bacterium *Rhodococcus* sp.: functional and structural studies and biotechnological potential. *Appl. Biochem. Biotechnol.* 172, 3054–3068. doi: 10.1007/s12010-013-0713-1

De Santi, C., Willassen, N. P., and Williamson, A. (2016b). Biochemical characterization of a family 15 carbohydrate esterase from a bacterial marine Arctic metagenome. *PLoS One* 11:e0159345. doi: 10.1371/journal.pone.0159345

Dhaulaniya, A. S., Balan, B., Agrawal, P. K., and Singh, D. K. (2019). Cold survival strategies for bacteria, recent advancement and potential industrial applications. *Arch. Microbiol.* 201, 1–16. doi: 10.1007/s00203-018-1602-3

Dopson, M., Ni, G., and Sleutels, T. H. (2015). Possibilities for extremophilic microorganisms in microbial electrochemical systems. *FEMS Microbiol. Rev.* 40, 164–181. doi: 10.1093/femsre/fuv044

Dumorne, K., Córdova, D. C., Astorga-Elo, M., and Renganathan, P. (2017). Extremozymes: a potential source for industrial applications. *J. Microbiol. Biotechnol.* 27, 649–659. doi: 10.4014/jmb.1611.11006

Dupont, C. L., Rusch, D. B., Yooseph, S., Lombardo, M. J., Alexander Richter, R., Valas, R., et al. (2012). Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* 6, 1186–1199. doi: 10.1038/ismej.2011.189

Elleuche, S., Schroder, C., Sahm, K., and Antranikian, G. (2014). Extremozymes–biocatalysts with unique properties from extremophilic microorganisms. *Curr. Opin. Biotechnol.* 29, 116–123. doi: 10.1016/j.copbio.2014.04.003

Emery, K. O., Hunt, J. M., and Hays, E. E. (1969). "Summary of hot brines and heavy metal deposits in the red sea," in *Hot Brines and Recent Heavy Metal Deposits in the Red Sea*, eds E. Degens and D. Ross (Berlin: Springer), 557–571. doi: 10.1007/978-3-662-28603-6_50

Fan, J., Han, F., and Liu, H. (2014). Challenges of big data analysis. *Natl. Sci. Rev.* 1, 293–314. doi: 10.1093/nsr/nwt032

Ferrandi, E. E., Previdi, A., Bassanini, I., Riva, S., Peng, X., and Monti, D. (2017). Novel thermostable amine transferases from hot spring metagenomes. *Appl. Microbiol. Biotechnol.* 101, 4963–4979. doi: 10.1007/s00253-017-8228-2

Ferrandi, E. E., Sayer, C., Isupov, M. N., Annovazzi, C., Marchesi, C., Iacobone, G., et al. (2015). Discovery and characterization of thermophilic limonene-1,2-epoxide hydrolases from hot spring metagenomic libraries. *FEBS J.* 282, 2879–2894. doi: 10.1111/febs.13328

Ferrer, M., Golyshina, O., Beloqui, A., and Golyshin, P. N. (2007). Mining enzymes from extreme environments. *Curr. Opin. Microbiol.* 10, 207–214. doi: 10.1016/j.mib.2007.05.004

Frumkin, D., Wasserstrom, A., Itzkovitz, S., Harmelin, A., Rechavi, G., and Shapiro, E. (2008). Amplification of multiple genomic loci from single cells isolated by laser micro-dissection of tissues. *BMC Biotechnol.* 8:17. doi: 10.1186/1472-6750-8-17

Fu, L., He, Y., Xu, F., Ma, Q., Wang, F., and Xu, J. (2015). Characterization of a novel thermostable patatin-like protein from a Guaymas basin metagenomic library. *Extremophiles* 19, 829–840. doi: 10.1007/s00792-015-0758-x

Fuchsman, C. A., Collins, R. E., Rocap, G., and Brazelton, W. J. (2017). Effect of the environment on horizontal gene transfer between bacteria and archaea. *PeerJ* 5, e3865–e3865. doi: 10.7717/peerj.3865

Gabani, P., and Singh, O. V. (2013). Radiation-resistant extremophiles and their potential in biotechnology and therapeutics. *Appl. Microbiol. Biotechnol.* 97, 993–1004. doi: 10.1007/s00253-012-4642-7

Galinski, E. A., Pfeiffer, H. P., and Truper, H. G. (1985). 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus Ectothiorhodospira. *Eur. J. Biochem.* 149, 135–139. doi: 10.1111/j.1432-1033.1985.tb08903.x

Garber, M., Grabherr, M. G., Guttman, M., and Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nat. Methods* 8, 469–477. doi: 10.1038/nmeth.1613

Gatti-Lafranconi, P., Natalello, A., Rehm, S., Doglia, S. M., Pleiss, J., and Lotti, M. (2010). Evolution of stability in a cold-active enzyme elicits specificity relaxation and highlights substrate-related effects on temperature adaptation. *J. Mol. Biol.* 395, 155–166. doi: 10.1016/j.jmb.2009.10.026

Gerday, C., and Glansdorff, N. (2009). *EXTREMOPHILES*, Vol. I. Paris: Eolss Publishers.

Giddings, L.-A., and Newman, D. J. (2015). "Bioactive compounds from extremophiles," in *Bioactive Compounds from Extremophiles: Genomic Studies, Biosynthetic Gene Clusters, and New Dereplication Methods*, eds L.-A. Giddings and D. J. Newman (Cham: Springer International Publishing), 1–47. doi: 10.1007/978-3-319-14836-6_1

Gilbert, J. A., Jansson, J. K., and Knight, R. (2014). The Earth Microbiome project: successes and aspirations. *BMC Biol.* 12:69. doi: 10.1186/s12915-014-0069-1

Grand View Research (2020). *Enzymes Market Size, Share, Global Industry Report, 2020-2027*. Available: https://www.grandviewresearch.com/industry-analysis/enzymes-industry (accessed January 2, 2021).

Grötzinger, S., Alam, I., Alawi, W., Bajic, V., Stingl, U., and Eppinger, J. (2014). Mining a database of single amplified genomes from Red Sea brine pool extremophiles—improving reliability of gene function prediction using a profile and pattern matching algorithm (PPMA). *Front. Microbiol.* 5:134. doi: 10.3389/fmicb.2014.00134

Grötzinger, S. W., Karan, R., Strillinger, E., Bader, S., Frank, A., Al Rowaihi, I. S., et al. (2017). Identification and experimental characterization of an extremophilic brine pool alcohol dehydrogenase from single amplified genomes. *ACS Chem. Biol.* 13, 161–170. doi: 10.1021/acschembio.7b00792

Hedlund, B. P., Dodsworth, J. A., Murugapiran, S. K., Rinke, C., and Woyke, T. (2014). Impact of single-cell genomics and metagenomics on the emerging view of extremophile "microbial dark matter". *Extremophiles* 18, 865–875. doi: 10.1007/s00792-014-0664-7

Herbert, R. A. (1992). A perspective on the biotechnological potential of extremophiles. *Trends Biotechnol.* 10, 395–402. doi: 10.1016/0167-7799(92)90282-z

Horikoshi, K. (2016). "Alkaliphiles," in *Extremophiles: Where It All Began* (Tokyo: Springer Japan), 53–78.

Horner, D. S., Pavesi, G., Castrignanò, T., De Meo, P. D. O., Liuni, S., Sammeth, M., et al. (2010). Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. *Brief. Bioinform.* 11, 181–197. doi: 10.1093/bib/bbp046

Hua, Z.-S., Han, Y.-J., Chen, L.-X., Liu, J., Hu, M., Li, S.-J., et al. (2015). Ecological roles of dominant and rare prokaryotes in acid mine drainage revealed by metagenomics and metatranscriptomics. *ISME J.* 9, 1280–1294. doi: 10.1038/ismej.2014.212

Hugenholtz, P., and Kyrpides, N. C. (2009). A changing of the guard. *Environ. Microbiol.* 11, 551–553. doi: 10.1111/j.1462-2920.2009.01888.x

Hugenholtz, P., and Tyson, G. W. (2008). Metagenomics. *Nature* 455, 481–483.

Illanes, A., Cauerhff, A., Wilson, L., and Castro, G. R. (2012). Recent trends in biocatalysis engineering. *Bioresour. Technol.* 115, 48–57. doi: 10.1016/j.biortech.2011.12.050

Irwin, J. A. (2010). Extremophiles and their application to veterinary medicine. *Environ. Technol.* 31, 857–869. doi: 10.1080/09593330.2010.484073

Irwin, J. A., and Baird, A. W. (2004). Extremophiles and their application to veterinary medicine. *Ir. Vet. J.* 57, 348–354. doi: 10.1186/2046-0481-57-6-348

Jin, M., Gai, Y., Guo, X., Hou, Y., and Zeng, R. (2019). Properties and applications of extremozymes from deep-sea extremophilic microorganisms: a mini review. *Mar. Drugs* 17, 656. doi: 10.3390/md17120656

Johnson, D. B. (2014). Biomining-biotechnologies for extracting and recovering metals from ores and waste materials. *Curr. Opin. Biotechnol.* 30, 24–31. doi: 10.1016/j.copbio.2014.04.008

Jorquera, M. A., Graether, S. P., and Maruyama, F. (2019). Bioprospecting and biotechnology of extremophiles. *Front. Bioeng. Biotechnol.* 7:204. doi: 10.3389/fbioe.2019.00204

Kaeberlein, T., Lewis, K., and Epstein, S. S. (2002). Isolating "Uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* 296, 1127–1129. doi: 10.1126/science.1070633

Kamekura, M., Hamakawa, T., and Onishi, H. (1982). Application of halophilic nuclease H of *Micrococcus* varians subsp. halophilus to commercial production of flavoring agent 5'-GMP. *Appl. Environ. Microbiol.* 44, 994–995. doi: 10.1128/aem.44.4.994-995.1982

Karan, R., Capes, M. D., DasSarma, P., and DasSarma, S. (2013). Cloning, overexpression, purification, and characterization of a polyextremophilic β-galactosidase from the Antarctic haloarchaeon *Halorubrum lacusprofundi. BMC Biotechnol.* 13:3. doi: 10.1186/1472-6750-13-3

Karan, R., Capes, M. D., and DasSarma, S. (2012a). Function and biotechnology of extremophilic enzymes in low water activity. *Aquat. Biosyst.* 8, 4. doi: 10.1186/2046-9063-8-4

Karan, R., Grötzinger, S., Vogler, M., Akal, A., Renn, D., Vancea, A., et al. (2019). Single amplified genomes from Red Sea brine pool extremophiles-a source of novel enzymes. *Stud. Univ. Babes Bolyai Biol.* 64, 76.

Karan, R., and Khare, S. (2010). Purification and characterization of a solvent−stable protease from *Geomicrobium* sp. EMB2. *Environ. Technol.* 31, 1061–1072. doi: 10.1080/09593331003674556

Karan, R., and Khare, S. (2011). Stability of haloalkaliphilic *Geomicrobium* sp. protease modulated by salt. *Biochemistry (Moscow)* 76, 686. doi: 10.1134/s0006297911060095

Karan, R., Kumar, S., Sinha, R., and Khare, S. (2012b). "Halophilic microorganisms as sources of novel enzymes," in *Microorganisms in Sustainable Agriculture and Biotechnology*, eds T. Satyanarayana, B. Johri, and A. Prakash (Dordrecht: Springer), 555–579. doi: 10.1007/978-94-007-2214-9_25

Karan, R., Mathew, S., Muhammad, R., Bautista, D. B., Vogler, M., Eppinger, J., et al. (2020). Understanding high-salt and cold adaptation of a polyextremophilic enzyme. *Microorganisms* 8, 1594. doi: 10.3390/microorganisms8101594

Karan, R., Singh, S. P., Kapoor, S., and Khare, S. K. (2011). A novel organic solvent tolerant protease from a newly isolated *Geomicrobium* sp. EMB2 (MTCC 10310): production optimization by response surface methodology. *N. Biotechnol.* 28, 136–145. doi: 10.1016/j.nbt.2010.10.007

Kirk, R. G., and Ginzburg, M. (1972). Ultrastructure of two species of *Halobacterium. J. Ultrastruct. Res.* 41, 80–94. doi: 10.1016/s0022-5320(72)90040-8

Kristjánsson, J. K., and Hreggvidsson, G. O. (1995). Ecology and habitats of extremophiles. *World J. Microbiol. Biotechnol.* 11, 17–25. doi: 10.1007/BF00339134

Krüger, A., Schäfers, C., Schröder, C., and Antranikian, G. (2018). Towards a sustainable biobased industry – highlighting the impact of extremophiles. *N. Biotechnol.* 40, 144–153. doi: 10.1016/j.nbt.2017.05.002

Kumar, S., Grewal, J., Sadaf, A., Hemamalini, R., and Khare, S. K. (2016). Halophiles as a source of polyextremophilic α-amylase for industrial applications. *AIMS Microbiol.* 2, 1–26. doi: 10.3934/microbiol.2016.1.1

Kunin, V., Copeland, A., Lapidus, A., Mavromatis, K., and Hugenholtz, P. (2008). A bioinformatician's guide to metagenomics. *Microbiol. Mol. Biol. Rev.* 72, 557–578.

Kvist, T., Ahring, B. K., Lasken, R. S., and Westermann, P. (2007). Specific single-cell isolation and genomic amplification of uncultured microorganisms. *Appl. Microbiol. Biotechnol.* 74, 926–935. doi: 10.1007/s00253-006-0725-7

Lagier, J.-C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., et al. (2012). Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* 18, 1185–1193. doi: 10.1111/1469-0691.12023

Lagier, J.-C., Khelaifia, S., Alou, M. T., Ndongo, S., Dione, N., Hugon, P., et al. (2016). Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* 1, 1–8. doi: 10.1007/978-1-4614-6418-1_540-3

Lamers, P., Searcy, E., Hess, J. R., and Stichnothe, H. (2016). *Developing the Global Bioeconomy: Technical, Market, and Environmental Lessons from Bioenergy*. Cambridge, MA: Academic Press.

Landry, Z. C., Giovanonni, S. J., Quake, S. R., and Blainey, P. C. (2013). Optofluidic cell selection from complex microbial communities for single-genome analysis. *Methods Enzymol.* 531, 61–90. doi: 10.1016/b978-0-12-407863-5.00004-6

Laurence, M., Hatzis, C., and Brash, D. E. (2014). Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes. *PLoS One* 9:e97876. doi: 10.1371/journal.pone.0097876

Lebre, P. H., De Maayer, P., and Cowan, D. A. (2017). Xerotolerant bacteria: surviving through a dry spell. *Nat. Rev. Microbiol.* 15, 285–296. doi: 10.1038/nrmicro.2017.16

Leis, B., Angelov, A., Mientus, M., Li, H., Pham, V. T. T., Lauinger, B., et al. (2015).). Identification of novel esterase-active enzymes from hot environments by use of the host bacterium *Thermus thermophilus*. *Front. Microbiol.* 6:275. doi: 10.3389/fmicb.2015.00275

Li, J., Yun, Y.-q, Xing, L., and Song, L. (2017). Novel bioflocculant produced by salt-tolerant, alkaliphilic strain *Oceanobacillus polygoni* HG6 and its application in tannery wastewater treatment. *Biosci. Biotechnol. Biochem.* 81, 1018–1025. doi: 10.1080/09168451.2016.1274635

Liu, Z., Zhao, C., Deng, Y., Huang, Y., and Liu, B. (2015). Characterization of a thermostable recombinant β-galactosidase from a thermophilic anaerobic bacterial consortium YTY-70. *Biotechnol. Biotechnol. Equip.* 29, 547–554. doi: 10.1080/13102818.2015.1015244

Locey, K. J., and Lennon, J. T. (2016). Scaling laws predict global microbial diversity. *Proc. Natl. Acad. Sci. U.S.A.* 113, 5970–5975. doi: 10.1073/pnas.1521291113

López-López, O., Knapik, K., Cerdán, M. E., and González-Siso, M. I. (2015). Metagenomics of an alkaline hot spring in Galicia (Spain): microbial diversity analysis and screening for novel lipolytic enzymes. *Front. Microbiol.* 6:1291. doi: 10.3389/fmicb.2015.01291

Madern, D., Ebel, C., and Zaccai, G. (2000). Halophilic adaptation of enzymes. *Extremophiles* 4, 91–98. doi: 10.1007/s007920050142

Marcy, Y., Ouverney, C., Bik, E. M., Losekann, T., Ivanova, N., Martin, H. G., et al. (2007). Dissecting biological "dark matter" with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11889–11894. doi: 10.1073/pnas.0704662104

Margesin, R., and Feller, G. (2010). Biotechnological applications of psychrophiles. *Environ. Technol.* 31, 835–844. doi: 10.1080/09593331003663328

Martin, M., and Vandenbol, M. (2016). The hunt for original microbial enzymes: an initiatory review on the construction and functional screening of (meta) genomic libraries. *Biotechnol. Agron. Soc. Environ.* 20, 523–532.

McLean, J. S., Lombardo, M. J., Badger, J. H., Edlund, A., Novotny, M., Yee-Greenbaum, J., et al. (2013). Candidate phylum TM6 genome recovered from a hospital sink biofilm provides genomic insights into this uncultivated phylum. *Proc. Natl. Acad. Sci. U.S.A.* 110, E2390–E2399. doi: 10.1073/pnas.1219809110

Mukhopadhyay, A., Dasgupta, A. K., and Chakrabarti, K. (2015). Enhanced functionality and stabilization of a cold active laccase using nanotechnology based activation-immobilization. *Bioresour. Technol.* 179, 573–584. doi: 10.1016/j.biortech.2014.12.070

Müller, S., and Nebe-Von-Caron, G. (2010). Functional single-cell analyses: flow cytometry and cell sorting of microbial populations and communities. *FEMS Microbiol. Rev.* 34, 554–587. doi: 10.1111/j.1574-6976.2010.00214.x

Nakamura, K., Iizuka, R., Nishi, S., Yoshida, T., Hatada, Y., Takaki, Y., et al. (2016). Culture-independent method for identification of microbial enzyme-encoding genes by activity-based single-cell sequencing using a water-in-oil microdroplet platform. *Sci. Rep.* 6, 22259. doi: 10.1038/srep22259

Nowrotek, M., Jałowiecki, Ł., Harnisz, M., and Płaza, G. A. (2019). Culturomics and metagenomics: in understanding of environmental resistome. *Front. Environ. Sci. Eng.* 13:40. doi: 10.1007/s11783-019-1121-8

Oren, A. (2010). Industrial and environmental applications of halophilic microorganisms. *Environ. Technol.* 31, 825–834. doi: 10.1080/09593330903370026

Pabinger, S., Dander, A., Fischer, M., Snajder, R., Sperk, M., Efremova, M., et al. (2013). A survey of tools for variant analysis of next-generation genome sequencing data. *Brief. Bioinform.* 15, 256–278.

Petrovskaya, L. E., Novototskaya-Vlasova, K. A., Spirina, E. V., Durdenko, E. V., Lomakina, G. Y., Zavialova, M. G., et al. (2016). Expression and characterization of a new esterase with GCSAG motif from a permafrost metagenomic library. *FEMS Microbiol. Ecol.* 92, fiw046. doi: 10.1093/femsec/fiw046

Placido, A., Hai, T., Ferrer, M., Chernikova, T. N., Distaso, M., Armstrong, D., et al. (2015). Diversity of hydrolases from hydrothermal vent sediments of the Levante Bay, Vulcano Island (*Aeolian archipelago*) identified by activity-based metagenomics and biochemical characterization of new esterases and

an arabinopyranosidase. *Appl. Microbiol. Biotechnol.* 99, 10031–10046. doi: 10.1007/s00253-015-6873-x

Quillaguaman, J., Guzman, H., Van-Thuoc, D., and Hatti-Kaul, R. (2010). Synthesis and production of polyhydroxyalkanoates by halophiles: current potential and future prospects. *Appl. Microbiol. Biotechnol.* 85, 1687–1696. doi: 10.1007/s00253-009-2397-6

Raddadi, N., Cherif, A., Daffonchio, D., Neifar, M., and Fava, F. (2015). Biotechnological applications of extremophiles, extremozymes and extremolytes. *Appl. Microbiol. Biotechnol.* 99, 7907–7913. doi: 10.1007/s00253-015-6874-9

Radestock, S., and Gohlke, H. (2011). Protein rigidity and thermophilic adaptation. *Proteins* 79, 1089–1108. doi: 10.1002/prot.22946

Raghunathan, A., Ferguson, H. R. Jr., Bornarth, C. J., Song, W., Driscoll, M., and Lasken, R. S. (2005). Genomic DNA amplification from a single bacterium. *Appl. Environ. Microbiol.* 71, 3342–3347. doi: 10.1128/AEM.71.6.3342-3347.2005

Ramesh, A., Harani Devi, P., Chattopadhyay, S., and Kavitha, M. (2020). "Commercial applications of microbial enzymes," in *Microbial Enzymes: Roles and Applications in Industries*, eds N. K. Arora, J. Mishra, and V. Mishra (Singapore: Springer Singapore), 137–184. doi: 10.1007/978-981-15-1710-5_6

Rampelotto, P. H. (2013). Extremophiles and extreme environments. *Life (Basel)* 3, 482–485. doi: 10.3390/life3030482

Rashid, M., and Stingl, U. (2015). Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. *Biotechnol. Adv.* 33, 1755–1773. doi: 10.1016/j.biotechadv.2015.09.005

Rasor, J. P., and Voss, E. (2001). Enzyme-catalyzed processes in pharmaceutical industry. *Appl. Catal. Gen.* 221, 145–158. doi: 10.1016/S0926-860x(01)00804-3

Reed, C. J., Lewis, H., Trejo, E., Winston, V., and Evilia, C. (2013). Protein adaptations in archaeal extremophiles. *Archaea* 2013, 373275. doi: 10.1155/2013/373275

Rego, A., Raio, F., Martins, T. P., Ribeiro, H., Sousa, A. G., Séneca, J., et al. (2019). Actinobacteria and cyanobacteria diversity in terrestrial antarctic microenvironments evaluated by culture-dependent and independent methods. *Front. Microbiol.* 10:1018. doi: 10.3389/fmicb.2019.01018

Rekadwad, B., and Khobragade, C. (2017). "Marine polyextremophiles and their biotechnological applications," in *Microbial Applications Vol.1: Bioremediation and Bioenergy*, eds V. C. Kalia and P. Kumar (Cham: Springer International Publishing), 319–331. doi: 10.1007/978-3-319-52666-9_15

Rinke, C., Lee, J., Nath, N., Goudeau, D., Thompson, B., Poulton, N., et al. (2014). Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. *Nat. Protoc.* 9, 1038–1048. doi: 10.1038/nprot.2014.067

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., et al. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437. doi: 10.1038/nature12352

Rizk, M., Antranikian, G., and Elleuche, S. (2012). End-to-end gene fusions and their impact on the production of multifunctional biomass degrading enzymes. *Biochem. Biophys. Res. Commun.* 428, 1–5. doi: 10.1016/j.bbrc.2012.09.142

Robinson, P. K. (2015). Enzymes: principles and biotechnological applications. *Essays Biochem.* 59, 1–41. doi: 10.1042/bse0590001

Ross, D. A. (1972). Red Sea hot brine area: revisited. *Science* 175, 1455–1457. doi: 10.1126/science.175.4029.1455

Rothschild, L. J., and Mancinelli, R. L. (2001). Life in extreme environments. *Nature* 409, 1092–1101.

Sahoo, R. K., Kumar, M., Sukla, L. B., and Subudhi, E. (2017). Bioprospecting hot spring metagenome: lipase for the production of biodiesel. *Environ. Sci. Pollut. Res.* 24, 3802–3809. doi: 10.1007/s11356-016-8118-7

Sarhan, M. S., Hamza, M. A., Youssef, H. H., Patz, S., Becker, M., ElSawey, H., et al. (2019). Culturomics of the plant prokaryotic microbiome and the dawn of plant-based culture media–A review. *J. Adv. Res.* 19, 15–27. doi: 10.1016/j.jare.2019.04.002

Sarmiento, F., Peralta, R., and Blamey, J. M. (2015). Cold and hot extremozymes: industrial relevance and current trends. *Front. Bioeng. Biotechnol.* 3:148. doi: 10.3389/fbioe.2015.00148

Sayed, A., Ghazy, M. A., Ferreira, A. J. S., Setubal, J. C., Chambergo, F. S., Ouf, A., et al. (2014). A novel mercuric reductase from the unique deep brine environment of atlantis II in the red sea. *J. Biol. Chem.* 289, 1675–1687. doi: 10.1074/jbc.M113.493429

Schäfer, T., Borchert, T. W., Nielsen, V. S., Skagerlind, P., Gibson, K., Wenger, K., et al. (2007). "Industrial enzymes," in *White Biotechnology*, eds R. Ulber and D. Sell (Berlin: Springer), 59–131.

Schiraldi, C., Giuliano, M., and De Rosa, M. (2002). Perspectives on biotechnological applications of archaea. *Archaea* 1, 75–86. doi: 10.1155/2002/436561

Schmieder, R., and Edwards, R. (2011). Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *PLoS One* 6:e17288. doi: 10.1371/journal.pone.0017288

Schröder, C., Elleuche, S., Blank, S., and Antranikian, G. (2014). Characterization of a heat-active archaeal β-glucosidase from a hydrothermal spring metagenome. *Enzyme Microb. Technol.* 57, 48–54. doi: 10.1016/j.enzmictec.2014.01.010

Sellek, G. A., and Chaudhuri, J. B. (1999). Biocatalysis in organic media using enzymes from extremophiles. *Enzyme Microb. Technol.* 25, 471–482. doi: 10.1016/S0141-0229(99)00075-7

Senior, A. W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., et al. (2020). Improved protein structure prediction using potentials from deep learning. *Nature* 577, 706–710. doi: 10.1038/s41586-019-1923-7

Shapiro, M. G., Ramirez, R. M., Sperling, L. J., Sun, G., Sun, J., Pines, A., et al. (2014). Genetically encoded reporters for hyperpolarized xenon magnetic resonance imaging. *Nat. Chem.* 6, 629. doi: 10.1038/nchem.1934

Sharma, A., Kawarabayasi, Y., and Satyanarayana, T. (2012). Acidophilic bacteria and archaea: acid stable biocatalysts and their potential applications. *Extremophiles* 16, 1–19. doi: 10.1007/s00792-011-0402-3

Sheldon, R. A. (2016). Green chemistry and resource efficiency: towards a green economy. *Green Chem.* 18, 3180–3183. doi: 10.1039/c6gc90040b

Siddiqui, K. S., Parkin, D. M., Curmi, P. M., De Francisci, D., Poljak, A., Barrow, K., et al. (2009). A novel approach for enhancing the catalytic efficiency of a protease at low temperature: reduction in substrate inhibition by chemical modification. *Biotechnol. Bioeng.* 103, 676–686. doi: 10.1002/bit.22300

Siglioccolo, A., Paiardini, A., Piscitelli, M., and Pascarella, S. (2011). Structural adaptation of extreme halophilic proteins through decrease of conserved hydrophobic contact surface. *BMC Struct. Biol.* 11:50. doi: 10.1186/1472-6807-11-50

Singh, O. V., and Gabani, P. (2011). Extremophiles: radiation resistance microbial reserves and therapeutic implications. *J. Appl. Microbiol.* 110, 851–861. doi: 10.1111/j.1365-2672.2011.04971.x

Singh, R., Chopra, C., Gupta, V. K., Akhlaq, B., Verma, V., and Rasool, S. (2015). Purification and characterization of CHpro1, a thermotolerant, alkali-stable and oxidation-resisting protease of Chumathang hotspring. *Sci. Bull.* 60, 1252–1260. doi: 10.1007/s11434-015-0834-8

Singh, R., Kumar, M., Mittal, A., and Mehta, P. K. (2016). Microbial enzymes: industrial progress in 21st century. *3 Biotech* 6, 174.

Speda, J., Jonsson, B.-H., Carlsson, U., and Karlsson, M. (2017). Metaproteomics-guided selection of targeted enzymes for bioprospecting of mixed microbial communities. *Biotechnol. Biofuels* 10, 128.

Stefani, F. O., Bell, T. H., Marchand, C., de la Providencia, I. E., El Yassimi, A., St-Arnaud, M., et al. (2015). Culture-dependent and -independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. *PLoS One* 10:e0128272. doi: 10.1371/journal.pone.0128272

Stepankova, V., Bidmanova, S., Koudelakova, T., Prokop, Z., Chaloupkova, R., and Damborsky, J. (2013). Strategies for stabilization of enzymes in organic solvents. *ACS Catal.* 3, 2823–2836. doi: 10.1021/cs400684x

Sterner, R.h, and Liebl, W. (2001). Thermophilic adaptation of proteins. *Crit. Rev. Biochem. Mol. Biol.* 36, 39–106. doi: 10.1080/20014091074174

Steven, B., Briggs, G., McKay, C. P., Pollard, W. H., Greer, C. W., and Whyte, L. G. (2007). Characterization of the microbial diversity in a permafrost sample from the Canadian high Arctic using culture-dependent and culture-independent methods. *FEMS Microbiol. Ecol.* 59, 513–523. doi: 10.1111/j.1574-6941.2006.00247.x

Stierle, A. A., Stierle, D. B., and Kelly, K. (2006). Berkelic acid, a novel spiroketal with selective anticancer activity from an acid mine waste fungal extremophile. *J. Org. Chem.* 71, 5357–5360. doi: 10.1021/jo060018d

Stuart, E. S., Morshed, F., Sremac, M., and DasSarma, S. (2001). Antigen presentation using novel particulate organelles from halophilic archaea. *J. Biotechnol.* 88, 119–128. doi: 10.1016/s0168-1656(01)00267-x

Swan, B. K., Martinez-Garcia, M., Preston, C. M., Sczyrba, A., Woyke, T., Lamy, D., et al. (2011). Potential for chemolithoautotrophy among ubiquitous bacteria lineages in the dark ocean. *Science* 333, 1296–1300. doi: 10.1126/science.1203690

Sydlik, U., Gallitz, I., Albrecht, C., Abel, J., Krutmann, J., and Unfried, K. (2009). The compatible solute ectoine protects against nanoparticle-induced neutrophilic lung inflammation. *Am. J. Respir. Crit. Care Med.* 180, 29–35. doi: 10.1164/rccm.200812-1911OC

Thomas, T., Gilbert, J., and Meyer, F. (2012). Metagenomics - a guide from sampling to data analysis. *Microb. Inform. Exp.* 2, 3. doi: 10.1186/2042-5783-2-3

Thompson, A., Bench, S., Carter, B., and Zehr, J. (2013). Coupling FACS and genomic methods for the characterization of uncultivated symbionts. *Methods Enzymol.* 531, 45–60. doi: 10.1016/b978-0-12-407863-5.00003-4

Tripathy, S., Padhi, S. K., Mohanty, S., Samanta, M., and Maiti, N. K. (2016). Analysis of the metatranscriptome of microbial communities of an alkaline hot sulfur spring revealed different gene encoding pathway enzymes associated with energy metabolism. *Extremophiles* 20, 525–536. doi: 10.1007/s00792-016-0846-6

United Nations, Department of Economic and Social Affairs, Population Division (2017). *World Population Prospects: The 2017 Revision, Key Findings and Advance Tables*. New York, NY: United Nations.

Vieille, C., and Zeikus, G. J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1–43. doi: 10.1128/MMBR.65.1.1-43.2001

Vogler, M., Karan, R., Renn, D., Vancea, A., Vielberg, M.-T., Grötzinger, S. W., et al. (2020). Crystal Structure and Active Site Engineering of a Halophilic γ-Carbonic Anhydrase. *Front. Microbiol.* 11:742. doi: 10.3389/fmicb.2020.00742

Wetterstrand, K. A. (2013). *DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)*. Burlington, MA: ScienceOpen, Inc.

Wilson, M. C., Mori, T., Rückert, C., Uria, A. R., Helf, M. J., Takada, K., et al. (2014). An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* 506, 58–62. doi: 10.1038/nature12959

Wiltschi, B., Cernava, T., Dennig, A., Casas, M. G., Geier, M., Gruber, S., et al. (2020). Enzymes revolutionize the bioproduction of value-added compounds: from enzyme discovery to special applications. *Biotechnol. Adv.* 40, 107520. doi: 10.1016/j.biotechadv.2020.107520

Woyke, T., Tighe, D., Mavromatis, K., Clum, A., Copeland, A., Schackwitz, W., et al. (2010). One bacterial cell, one complete genome. *PLoS One* 5:e10314. doi: 10.1371/journal.pone.0010314

Yamamura, S., Kishi, H., Tokimitsu, Y., Kondo, S., Honda, R., Ramachandra Rao, S., et al. (2005). Single-cell microarray for analyzing cellular response. *Anal. Chem.* 77, 8050–8056. doi: 10.1021/ac0515632

Yan, W., Li, F., Wang, L., Zhu, Y., Dong, Z., and Bai, L. (2017). Discovery and characterizaton of a novel lipase with transesterification activity from hot spring metagenomic library. *Biotechnol. Rep.* 14, 27–33. doi: 10.1016/j.btre.2016.12.007

Yoshimoto, N., Kida, A., Jie, X., Kurokawa, M., Iijima, M., Niimi, T., et al. (2013). An automated system for high-throughput single cell-based breeding. *Sci. Rep.* 3, 1191. doi: 10.1038/srep01191

Zarafeta, D., Moschidi, D., Ladoukakis, E., Gavrilov, S., Chrysina, E. D., Chatziioannou, A., et al. (2016). Metagenomic mining for thermostable esterolytic enzymes uncovers a new family of bacterial esterases. *Sci. Rep.* 6:38886. doi: 10.1038/srep38886

Zengler, K., Toledo, G., Rappé, M., Elkins, J., Mathur, E. J., Short, J. M., et al. (2002). Cultivating the uncultured. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15681–15686. doi: 10.1073/pnas.252630999

Zhang, K., Martiny, A. C., Reppas, N. B., Barry, K. W., Malek, J., Chisholm, S. W., et al. (2006). Sequencing genomes from single cells by polymerase cloning. *Nat. Biotechnol.* 24, 680–686. doi: 10.1038/nbt1214

Zhao, C., Chu, Y., Li, Y., Yang, C., Chen, Y., Wang, X., et al. (2017). High-throughput pyrosequencing used for the discovery of a novel cellulase from a thermophilic cellulose-degrading microbial consortium. *Biotechnol. Lett.* 39, 123–131. doi: 10.1007/s10529-016-2224-y

Zweerink, S., Kallnik, V., Ninck, S., Nickel, S., Verheyen, J., Blum, M., et al. (2017). Activity-based protein profiling as a robust method for enzyme identification and screening in extremophilic Archaea. *Nat. Commun.* 8, 15352. doi: 10.1038/ncomms15352