

The Santorini Volcanic Complex as a Valuable Source of Enzymes for Bioenergy

Paraskevi N. Polymenakou ^{1,*}, Paraskevi Nomikou ², Haris Zafeiropoulos ^{1,3}, Manolis Mandalakis ¹,
Thekla I. Anastasiou ¹, Stephanos Kiliadis ², Nikos C. Kyrpides ⁴, Georgios Kotoulas ¹ and Antonios Magoulas ¹

- ¹ Hellenic Centre for Marine Research, Institute of Marine Biology, Biotechnology and Aquaculture, Gournes Pediados, P.O. Box 2214, GR 71003 Heraklion, Crete, Greece; haris-zaf@hcmr.gr (H.Z.); mandalakis@hcmr.gr (M.M.); theanast@hcmr.gr (T.I.A.); kotoulas@hcmr.gr (G.K.); magoulas@hcmr.gr (A.M.)
- ² Faculty of Geology and Geoenvironment, National and Kapodistrian University of Athens, Panepistimioupoli Zografou, 15784 Athens, Greece; evinom@geol.uoa.gr (P.N.); kiliadis@geol.uoa.gr (S.K.)
- ³ Department of Biology, University of Crete, GR 71500 Crete, Greece
- ⁴ Department of Energy, Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; nckyrpides@lbl.gov
- * Correspondence: polymen@hcmr.gr

Abstract: Marine microbial communities are an untapped reservoir of genetic and metabolic diversity and a valuable source for the discovery of new natural products of biotechnological interest. The newly discovered hydrothermal vent field of Santorini volcanic complex located in the Aegean Sea is gaining increasing interest for potential biotechnological exploitation. The conditions in these environments, i.e., high temperatures, low pH values and high concentration of heavy metals, often resemble harsh industrial settings. Thus, these environments may serve as pools of enzymes of enhanced catalytic properties that may provide benefits to biotechnology. Here, we screened 11 metagenomic libraries previously constructed from microbial mat samples covering the seafloor and the polymetallic chimneys of Kolumbo volcano as well as mat samples from Santorini caldera, to mine, in silico, genes associated with bioenergy applications. We particularly focused on genes encoding biomass hydrolysis enzymes such as cellulases, hemicellulases and lignin-degrading enzymes. A total of 10,417 genes were found for three specific groups of enzymes—i.e., the endoglucanases, the three different beta-glucosidases BGL, bgIX and bgIB, and the alpha-galactosidases melA, and rafA. Overall, we concluded that the Santorini–Kolumbo volcanic ecosystems constitute a significant resource of novel genes with potential applications in bioenergy that deserve further investigation.

Keywords: bioenergy; metagenomics; extreme environments; microbial mats; submarine volcanoes; Santorini



Citation: Polymenakou, P.N.; Nomikou, P.; Zafeiropoulos, H.; Mandalakis, M.; Anastasiou, T.I.; Kiliadis, S.; Kyrpides, N.C.; Kotoulas, G.; Magoulas, A. The Santorini Volcanic Complex as a Valuable Source of Enzymes for Bioenergy. *Energies* **2021**, *14*, 1414. <https://doi.org/10.3390/en14051414>

Academic Editor: Byong-Hun Jeon

Received: 22 December 2020

Accepted: 28 February 2021

Published: 4 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Among all different extreme habitats that can be found on Earth, marine hydrothermal vents (HVs) represent the very promising targets for the discovery of some of the Earth's oddest and toughest microbes. These microbes have an enormous potential in the so called “blue biotechnology”—i.e., the area of biotechnology which exploits the diversity of marine organisms [1,2]. Hydrothermal systems associated with volcanic activity can support diverse ecosystems providing an oasis for microorganisms with many diverse abilities, yet they are largely unexplored. The geochemical conditions allowing life in such environments could only be met by nonphotosynthetic, lithotrophic microbial communities, having evolved a broad range of new metabolic capabilities [3,4]. However, despite the vital role of microorganisms in HV ecosystems, and their exploitation potential in biotechnology, key questions about their ecology, diversity and metabolic capabilities remain understudied simply because many HV environments remain largely unexplored [5]. Due to the challenges in accessing samples, they are among the least understood ecosystems on Earth [6].

The Hellenic Volcanic Arc (HVA) located in the Aegean Sea contains geologically unique and poorly studied hydrothermal vents. HVA was developed as a response to the northward subduction of the African plate beneath the active margin of the European plate [7,8]. In this arc, volcanism and hydrothermal activity occur through thinned continental crust [9]. HVA consists of five islands—Methana, Milos, Santorini, Nisiros and Kos [10]. In these areas, the vents release substantial amounts of CO₂, H₂, and H₂S into the water column, creating an environment where chemolithoautotrophism plays a key role [10,11]. The Santorini volcano is world famous because of its recent major explosive eruption that happened ~3600 years ago. This volcanic field extends for 20 km and consists of approximately 20 submarine cones. The largest cone is the submarine Kolumbo volcano which is situated 505 m below the sea level. The diameter of this volcano is 3 km and its crater is 1500 m wide. The northernmost part of the crater is covered by a large number of polymetallic chimneys and vents which release fluids of high (up to 220 °C) and low temperatures (up to 70 °C) [12]. The exterior of these chimneys and the seabed of the crater are covered with microbial mats of white/grey and reddish/orange colours. The volcanic discharge of CO₂ results in local pH reduction [13,14]. As a probable consequence of the elevated metal concentrations, high temperatures and low pH values, there is an absence of macrofaunal species which are typically found in HV sites. Indeed, the only thriving biological entities are of bacterial and archaeal origins [9].

Metagenomics, which is the study of the genetic content of all microbial organisms in an environmental sample, has the potential to impact a broad range of biotechnologies. The newly discovered HV field of Kolumbo volcano of the Hellenic Volcanic Arc (HVA), is gaining increasing interest for potential biotechnological exploitation; however, only recently, the first biogeochemical/biotechnological [9,15,16] and metagenomics investigation of its seafloor and water column [5,17,18] has been undertaken. These investigations, that were held in microbial mat samples covering the seafloor of the Santorini–Kolumbo volcanic system and the surfaces of Kolumbo chimneys, revealed that both Kolumbo crater and Santorini caldera harbor highly complex prokaryotic communities composed mainly of chemolithoautotrophs [5,15]. However, our most striking finding at the seafloor of this volcanic complex until now is the enormous number of genes and pathways linked to key metabolic processes with great exploitation potential in different biotechnological fields, including bioremediation and biomedicine [5]. Indeed, in our libraries we identified genes and group of organisms linked to anaerobic degradation of aromatic hydrocarbons, as well as CRISPR elements [5]. According to Barrangou and Horvath [19], these elements are antiviral defense mechanisms with great potential in biomedicine applications.

In this study, we investigated genes of interest in the bioenergy sector in order to evaluate the exploitation potential of the Kolumbo volcano in the emerging and challenging field of renewable energy. Current society and the current economy are largely dependent on petroleum as a fuel source; however, petroleum feedstocks are becoming scarcer and more expensive. Thus, there is a growing demand to develop strategies for fuel production from alternative carbon sources such as plant biomass instead of petroleum [20]. Lignocellulose represents an abundant and sustainable raw material for biofuel production accounting for approximately 70% of the carbon fixed by land plant photosynthesis [21,22]. Conversion of lignocellulosic biomass to biofuels has been an attractive alternative to petroleum usage; however, its conversion to biofuels is still complex and expensive. Thus, the development of a strategy capable of finding more effective enzymes with improved catalytic properties in a cost-efficient manner, e.g., enzymes of enhanced thermal and mechanical stability and better pH response, may revolutionize the biotechnology and chemical industries [23].

We explored metagenomic data from 11 samples of microbial mats covering the seafloor and the polymetallic chimneys of Kolumbo volcano, as well as mat samples from Santorini caldera, to mine genes of bioenergy importance. We particularly focused on genes encoding proteins that can effectively hydrolyze biomass such as cellulose, hemicellulose and lignin.

2. Materials and Methods

2.1. Source of the Investigated Metagenomic Data

A total of 11 metagenomic libraries previously constructed by IMBBC-HCMR and the Joint Genome Institute, Department of Energy, USA (JGI, DoE) were investigated [5,17,18].

The metagenomes have been constructed from samples collected within the framework of oceanographic sampling expeditions of the Hellenic Centre for Marine Research and the Rhode Island University, USA (Figures 1 and 2) and they are all available in the IMG/M data management system of JGI, DOE, USA. These included:

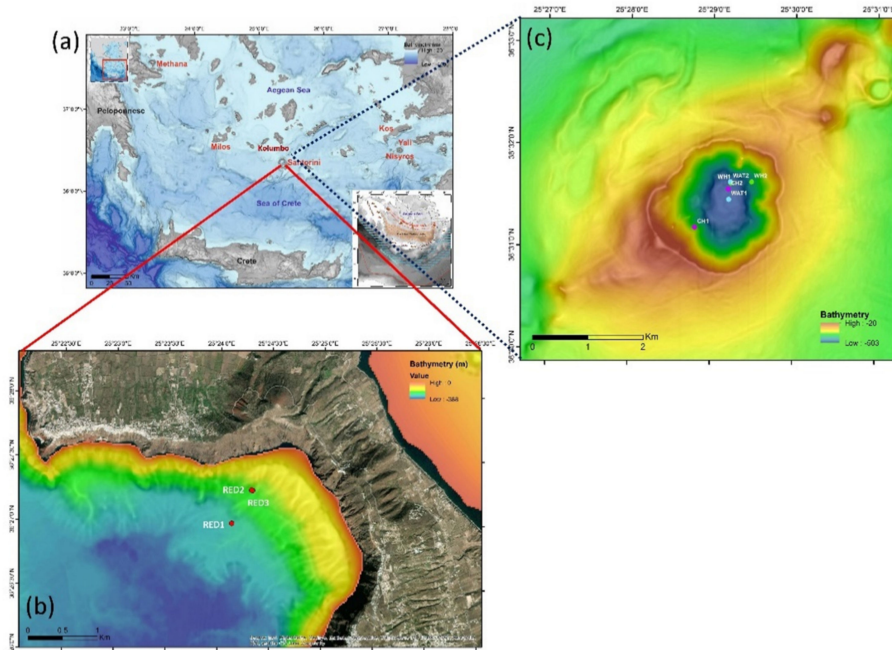


Figure 1. (a) Map showing the positions of the samples that were used as sources of metagenomic data; (b) Details of samples position in Santorini caldera (red dots); (c) Details of samples position in Kolumbo volcano.

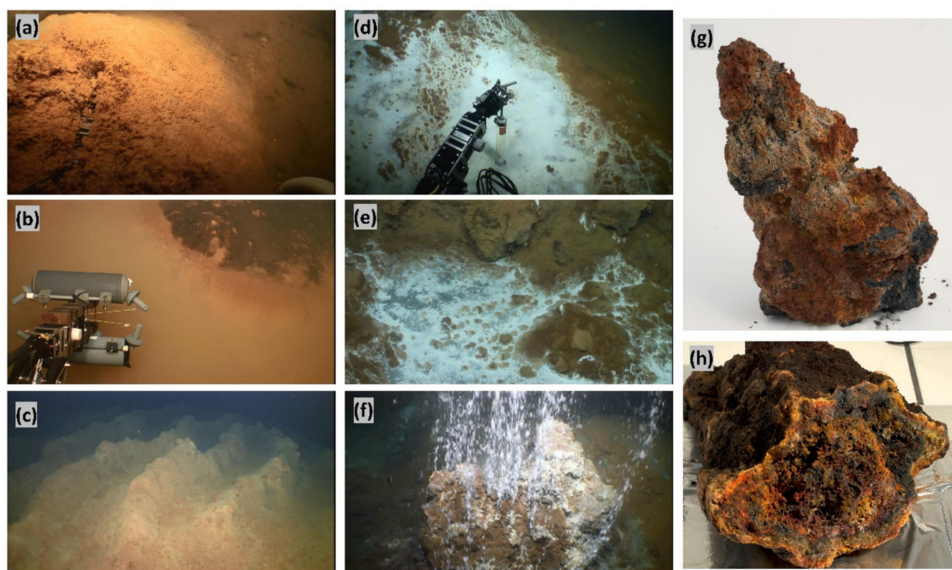


Figure 2. Photos of the samples used in metagenomics: (a) RED1; (b) RED2; (c) RED3; (d) WH1; (e) WH2; (f) WAT1; (g) CH1; (h) CH2. Underwater photos were taken by the cameras attached to the frame of the Remoted Operated Vehicle MaxRover of HCMR that was used for samples collection:

(1) The red mat sample RED1 which was retrieved from the sea-bottom of Santorini caldera (RED1: 36.4499° N, 25.4021° E, depth: 336 m).

(2) The sample RED2 which was retrieved from the bottom pool fluid of the Kalisti Limnes located at the slopes of Santorini caldera. This fluid was a mixture of hydrothermal fluid with seawater, plus reddish Fe-mat and Fe-rich suspended particulate material. The suspended particulate material which was further used for metagenomic analysis was collected by pumping the sample fluid through a GF/F filter (RED2: 36.4542° N, 25.4054° E, depth: 229 m).

(3) The red mat sample RED3, a flocculated solid material, which was retrieved from the Kalisti Limnes reddish microbial Fe-mat (RED3: 36.4543° N, 25.4052° E, depth: 229 m).

(4 and 5) The white mat samples WH1 and WH2 which were collected from the north part of the Kolumbo crater (WH1: 36.5273° N, 25.4872° E, depth: 494 m, and WH2: 36.5275° N, 25.4914° E, depth: 482 m).

(6–8) The microbial mat samples covering the outer, middle and inside layers of the polymetallic sulfide chimney V16 (i.e., CH1-V16c.out, CH1-V16c.mid, CH1-V16c.in, respectively; chimney coordinates: 36.5200° N, 25.4800° E, depth: 503 m).

(9) The microbial mat sample CH2 covering the surface of the sulfide chimney V59 (CH2: 36.5263° N, 25.4867° E, depth: 505 m).

(10) The water sample WAT1 which was collected at the water column, at an altitude of 30 m above the active vent chimneys (WAT1: 36.5264° N, 25.4868° E, depth: 470 m).

(11) The water sample WAT2 which was collected from the bottom of the inactive area of Kolumbo at the southern part of its crater (WAT2: 36.5275° N, 25.4871° E, depth: 495 m).

Samples RED1 and WH1 were collected during the sampling expedition EN419 that took place in Santorini volcanic complex in May–June 2006 with the R/V Endeavor of Rhode Island University, in a joint campaign with the Hellenic Centre for Marine Research. High-quality DNA was prepared at IMBBC-HCMR and library construction and bioinformatic analysis were performed at JGI, DoE, USA. During this expedition, the bottom water temperature within the CO₂-poor Santorini caldera was close to 22 °C, whereas the temperatures at the seafloor of the CO₂-rich Kolumbo volcano varied from 60 to 70 °C [5].

Samples RED2, RED3 of Santorini caldera, the white mat sample WH2, the three layers of sulfide chimney CH1-V16c, the sulfide chimney CH2 and the water samples WAT1 and WAT2 were collected within the framework of the sampling campaign 2Biotech of the EU-FP7 SeaBioTech project that took place in September 2013 with the R/V Aegeao. Samples were collected with the Remotely Operated Vehicle (ROV) Max Rover of the HCMR. High-quality DNA extraction, library construction, sequencing and bioinformatic analysis were carried out by IMBBC-HCMR and JGI-DoE. During the Seabiotech expedition, the temperatures that were recorded using the remote operated vehicle-mounted CTD (SBE 39; Seabird Scientific) were as follows: at the slopes of Santorini caldera, the pool fluid temperature of sample RED2 was 29 °C, whereas the ambient seawater above the sample RED3 was 16.5 °C; in Kolumbo crater, the temperature of the ambient seawater above the seabed varied from 15.8 to 16.5 °C (samples WAT1 and WAT2), the temperature in the microbial mats covering the seabed varied from 60 to 70 °C (sample WH2) and the temperatures of the ambient water surrounding the sulfide chimneys V16 (CH1-V16c) and V59 (CH2) were 15.9 and 16.2 °C, respectively [15,17,18].

2.2. List of Genes

We have focused on the biomass hydrolysis process mediated by enzymes with interest in the bioenergy sector. Plant biomass consists of three major biopolymers—i.e., cellulose, hemicellulose and lignin. Cellulose is a highly recalcitrant substrate and it is protected by hemicellulose and lignin. It is a homopolymer of beta-1,4-linked glucose units. It is a major component of primary (up to 15–30%) and secondary cell walls (up to 40%) [22,24]. The second most abundant biomass component is hemicellulose, a group of short chain, branched, substituted polymer of sugars [22,25]. The sugar monomers

are galactose, mannose, xylose, rhamnose, and arabinose [22,26]. Lignin is a class of hydrocarbon polymers consisting of aliphatic and aromatic structures. It is an amorphous polymer that contains phenyl propane units.

In many processes, cellulases must withstand the harsh conditions of the industrial bio-conversion process, such as high temperatures of greater than 50 °C, and even low or high pHs [23,27]. Microorganisms can produce a series of enzymes enabling the degradation of all these biopolymers (Table 1). The enzymes capable of degrading cellulose, named cellulases, contain three major categories of enzymes—namely, endoglucanases (EGs), exoglucanases/cellobiohydrolases (CBH) and the beta-glucosidases (BGL) [22,28]. The enzymes degrading hemicellulose consist of xylanase, which is the major enzyme, xyloglucanase, mannans endo 1,4-beta-mannosidase and beta-1,3- and beta-1,4-glucanase/licheninase. The major group of xylanase enzymes consists of endoxylanases, exoxylanases or glycosyl hydrolases (GHs), glucuronidases, acetylxylan esterases, ferulic acid esterases and alpha-galactosidase. The lignin-degrading enzymes consist of oxidoreductases (i.e., laccases), and heme peroxidases such as lignin (LiP) and manganese peroxidases (MnP) (Table 1) [22]. By using their Enzyme Commission (EC) numbers, we searched for the presence of genes encoding these enzymes in the metagenomic libraries (Table 1). We used the KEGG Ortholog database of the 11 metagenomes which is available in the IMG/M data management system of JGI, DoE, Berkeley, CA, USA.

Table 1. Counts of genes (KEGG Orthologs) encoding the three major categories of biomass degrading enzymes in the 11 metagenomic libraries. Gene counts measure the relative abundance of each gene per sample [29].

	EC Number	WH1	WH2	RED1	RED2	RED3	CH1-V16c.out	CH1-V16c.mid	CH1-V16c.in	CH2	WAT1	WAT2
1. Cellulases												
endoglucanases (EG)	EC 3.2.1.4	637	222	1646	60	72	29	32	17	372	119	178
beta-glucosidase (BGL) bgIX	EC 3.2.1.21	586	226	1164	33	98	21	19	22	282	262	296
beta-glucosidase (BGL) E3.2.1.21	EC 3.2.1.21	163	4	271	0	1	0	0	0	5	1	5
beta-glucosidase (BGL) bgIB	EC 3.2.1.21	124	70	450	17	27	4	8	7	164	82	125
cellulose 1,4-beta-cellobiosidase CBHA_B & CBHC	EC 3.2.1.91	18	21	36	3	3	0	2	0	9	9	12
exoglucanases/cellobiohydrolases (CBH)	EC 3.2.1.176	0	0	0	0	0	0	0	0	0	0	0
exoglucanases/cellobiohydrolases (CBH)	EC 3.2.1.74	0	0	0	0	0	0	0	0	0	0	0
2. Hemicellulases												
alpha-galactosidases E3.2.1.22B, galA, rafA	EC 3.2.1.22	153	216	530	9	23	14	13	10	180	86	179
endo-1,4-beta-xylanase xynA	EC 3.2.1.8	129	38	285	10	11	4	7	5	114	38	55
alpha-galactosidases melA	EC 3.2.1.22	111	50	445	25	48	3	4	1	88	102	206
alpha-N-arabinofuranosidases abfA	EC 3.2.1.55	92	28	487	20	18	0	0	1	57	50	63
beta-glucuronidase uidA, GUSB	EC 3.2.1.31	78	17	388	6	21	0	3	0	80	12	16
xylan 1,4-beta-xylosidase xynB	EC 3.2.1.37	60	8	163	0	1	0	0	0	21	14	6
mannans endo 1,4-beta-mannosidase gmuG	EC 3.2.1.78	38	5	159	2	6	0	1	0	28	6	5
feruloyl esterase	EC 3.1.1.73	21	4	83	0	0	5	12	6	28	53	71
acetylxylan esterases AXE1	EC 3.1.1.72	7	0	5	0	0	0	0	0	0	0	1
licheninase	EC 3.2.1.73	7	1	23	1	1	0	0	1	2	1	7
arabinoxylan arabinofuranohydrolase xynD	EC 3.2.1.55	2	2	13	0	0	0	0	0	1	0	0
alpha-galactosidases GLA	EC 3.2.1.22	1	0	0	0	0	0	0	0	1	0	0
beta-D-xylosidase 4 XYL4	EC 3.2.1.37	0	0	0	0	0	0	0	0	0	10	3
xyloglucan-specific endo-beta-1,4-glucanase XEG	EC 3.2.1.151	0	0	0	0	0	0	0	0	0	0	0
mannans endo 1,4-beta-mannosidase MAN	EC 3.2.1.78	0	4	0	0	3	0	0	0	6	1	4
3. Lignin-degrading enzymes												
laccases	EC 1.10.3.2	0	0	0	0	0	0	0	0	0	3	0
lignin peroxidase (LiP)	EC 1.11.1.14	0	0	0	0	0	0	0	0	0	0	0
manganese peroxidase (MnP)	EC 1.11.1.13	0	0	0	0	0	0	0	0	0	0	0

2.3. Phylogenetic Tree of Endoglucanases, Beta-Glucosidases and Alpha-Galactosidases

A total of 222 endoglucanases, 226 beta-glucosidases bgIX and 216 alpha-galactosidases, melA and rafA, from the metagenome WH2 were selected for the construction of the phylogenetic trees. Putative amino acid sequences of these enzymes were extracted from the IMG/M database. A series of reference enzymes from the public database of Genbank—i.e., 14 endoglucanases, 17 beta-glucosidases bgIX and 14 alpha-galactosidases melA and rafA from different microbial strains and metagenomes from hydrothermal vent sites were obtained and used for the tree constructions. Signal peptides were then identified and removed by using the software SignalIP (<http://www.cbs.dtu.dk/services/SignalIP/>; accessed on 24 February 2021) [30]. Aligned and trimmed sequences were used to construct the phylogeny trees using both RAXML-NG (<https://github.com/amkozlov/raxml-ng>; accessed on 24 February 2021) [31] and IQ-TREE (<http://www.iqtree.org/>; accessed on 24 February 2021) [32]. The phylogenetic trees were built with RAXML-NG using the model DEN and with IQ-TREE using the models LG+R5 for endoglucanases and WAG+R4 for alpha-galactosidases and beta-glucosidases. In the case of RAXML-NG, 1000 bootstrapped

replicates were required; however, the algorithm converged after 900 bootstrap trees. Trees were visualized and obtained by using iTOL (<https://itol.embl.de/upload.cgi>; accessed on 24 February 2021) [33]. Both programs produced similar results and only the trees obtained by RAxML-NG are presented.

3. Results and Discussion

Through the enormous amount of metagenomic data being produced from a wide range of different environments, we can develop a better understanding of how to effectively harness various renewable energy sources [34]. Here, we explored the genetic content of 11 large metagenomic libraries previously developed from the extreme environments of the Santorini–Kolumbo volcanic complex in the context of bioenergy. We were particularly interested in genes involved in lignocellulosic biomass conversion to biofuels. Different groups of gene-encoded enzymes were mined and the results of gene counts are presented in Table 1.

One of the biggest challenges in biofuel research is the discovery of new cellulases of improved performance in order to reduce the amount and the cost of their production. The investigated environment of the Santorini volcanic complex benefits from this point of view because of its harsh conditions. A high number of genes encoding cellulases and hemicellulases were encountered in most of the metagenomes of the present study, whereas lignin-degrading enzymes were almost absent. Interestingly, high gene counts were found for three specific groups of enzymes—i.e., the endoglucanases (EG: 3384 genes), the three different beta-glucosidases (BGL: 450 genes, bglX: 3009 genes, bglB: 1078 genes) and the alpha-galactosidases (melA: 1083, gala, rafA: 1413). Endoglucanases (EGs) are very important components in biomass conversion to biofuels since they are considered as the initiators of cellulose hydrolysis. They are involved in the conversion of plant components into sugar. These enzymes act by cleaving internal beta-glycosidic bonds in the cellulose chain. Then, the ends of cellulose chain become accessible to cellobiohydrolase that produce cellobiose. Beta-glucosidase further cleaves cellobiose into glucose molecules [35]. Endoglucanases have an EC number of 3.2.1.4 and consist of one the most important groups of cellulases which is widely distributed in archaeal, bacterial, fungal and other eukaryotic species [36]. For cellulases and endoglucanases, a lot of researches have been carried out on fungus *Trichoderma reesei* [37] as well as on the Gram-positive, endospore forming and cellulolytic bacterium *Clostridium* that can produce a variety of chemicals including n-butanol [38]. However, the expression of endoglucanases from their host microbial communities adds to the final cost of the biofuel product. Therefore, there is a growing interest in exploiting the potential of thermostable bioprocessing enzymes. In the present study, we were able to mine more than 3000 gene-encoding endoglucanases. Interestingly, these genes were characterized by high diversity. Figure 3a presents the phylogenetic analysis of 222 partial endoglucanase amino acid metagenome sequences that were mined from the high-temperature WH2 sample. This analysis revealed three distinct clusters of diverse set of endoglucanase enzymes which are related to members of bacterial (green color; Figure 3a) and archaeal communities (red color; Figure 3a), as well as to community samples (blue color; Figure 3a).

These include enzymes from thermophilic communities such as the archaeal Eurarchaeota (e.g., *Thermococcus* sp., *Thermococcus chitinophagus*, *Thermococcus radiotolerans*), and the diverse thermophilic bacterial members of Gammaproteobacteria (e.g., *Alteromonas* sp., *Pseudoalteromonas telluritireducens*) and Thermotogae (e.g., *Thermotoga* sp. *Ku-13t*). Although several studies have been carried out to mine novel endoglucanases from microbial strains [39,40] or metagenomic samples [41,42], few investigations exist from high-temperature environments such as hydrothermal vents. Indeed, despite the high demand for new extremozymes, the number of known thermoactive cellulases such as endoglucanases is low and mostly originates from archaeal communities [43,44]. In the present study, we presented a large set of diverse enzymes closely related to enzymes from both

thermophilic bacterial and archaeal communities, providing promising candidates for high-temperature biotechnological processes.

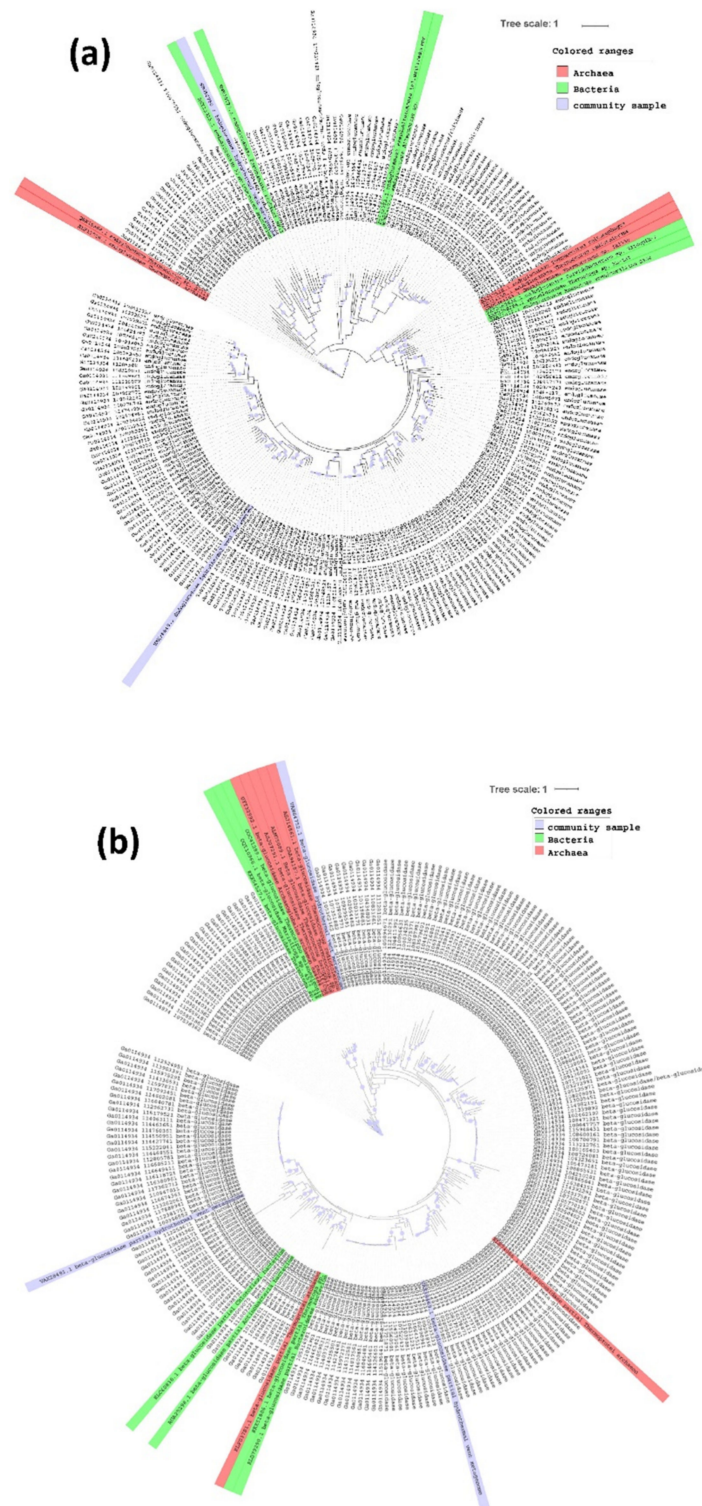


Figure 3. Cont.

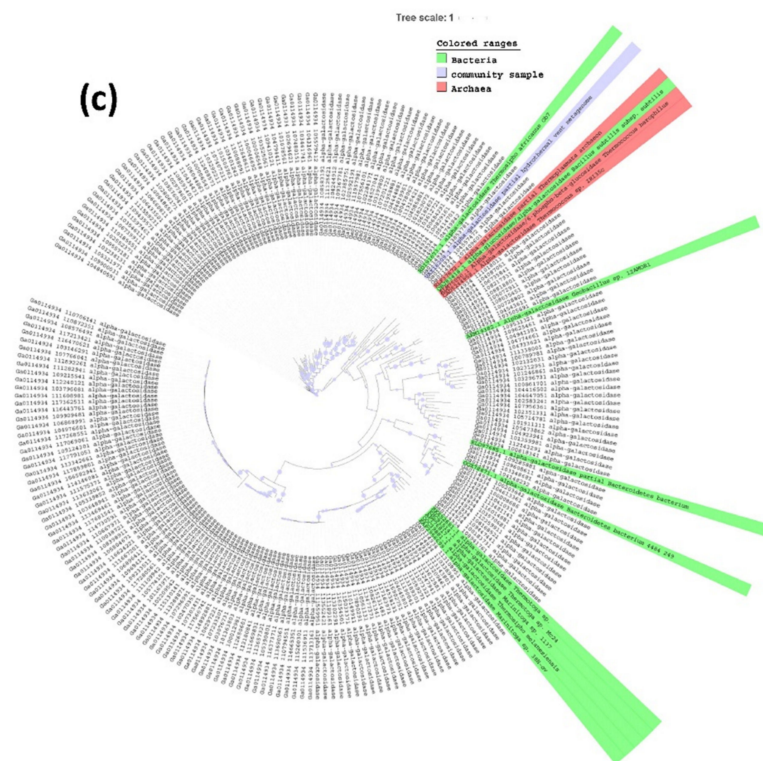


Figure 3. Phylogenetic maximum-likelihood tree based on partial sequences of (a) endoglucanase, (b) beta-glucosidase and (c) alpha-galactosidase amino acid sequences of the WH2 metagenome. Reference sequences from various microbial species and other metagenomes that were retrieved from Genbank are also included. The branches denoting sequences derived from GenBank were colored on the taxonomy domain they are coming from—i.e., green for bacteria, red for archaea and blue for community samples. Scale bar represents 1 amino acid substitution per 100. Dots on the branches represent bootstrap support values. The larger the dots, the higher the bootstrap values (close to 100%).

In addition to endoglucanases, beta-glucosidase genes with an EC number of 3.2.1.21 were also very abundant in almost all metagenomic libraries. The most abundant beta-glucosidase *bglX*, also known as periplasmic beta-glucosidase, can perform hydrolysis of terminal, nonreducing beta-D-glycosyl residues with release of beta-D-glucose. Beta-glucosidases are of great importance in industrial processes and they are mainly used for the enzymatic saccharification of cellulosic matrixes [45]. These enzymes are ubiquitous, occurring in organisms representing all domains of life from bacteria to highly evolved mammals. The *Thermotoga* species is an important source of hyperthermophilic glycosidases [46], whereas different *Aspergillus* species have also been studied as a source of glucosidases that can improve the efficiency of the saccharification step [47]. The phylogenetic analysis of beta-glucosidase amino acid sequences revealed the presence of diverse enzymes. A large number of different clusters were produced containing reference enzymes from both thermophilic bacterial and archaeal communities, as well as from metagenomic libraries from high-temperature environments (Figure 3b). These enzymes were most closely related to beta-glucosidases mined from metagenomes of the hydrothermal vent sediments in the hydrocarbon-rich Guayamas Basin [48,49] and to beta-glucosidase of the thermophilic *Thermotoga* sp. [48,49].

High gene counts were also recorded for alpha-galactosidases, *mela* and *rafA*, with an EC number of 3.2.1.22 (Table 1). These enzymes catalyze the hydrolysis of alpha-(1,6) bonds of galactose residues in galacto-oligosaccharides. A number of industrial applications of alpha-galactosidases are known, mainly in the food and pharmaceutical industry, but also in bioenergy sector [50]. In the present study, alpha-galactosidases phylogeny also

showed the presence of a large number of clusters. These enzyme clusters are closely related to diverse enzymes including those of the thermophilic *Geobacillus* sp. 12AMOR1 which have been isolated from an Arctic deep-sea hydrothermal vent site [51] and of the hyperthermophilic archaeon *Thermococcus barophilus*, which has been isolated from a deep-sea hydrothermal vent of the Mid-Atlantic Ridge [52]. Interestingly, many clusters were not closely related to any known alpha-galactosidases, implying the novelty characteristics of these enzymes.

Advances in genetics and biotechnology are introducing a new view of converting complex biomass into value-added products such as biofuels. In this context, emphasis has been placed upon the genetic conversion of biomass into fuels. Indeed, the production of cost-effective cellulases and hemicellulases for biomass hydrolysis is the major challenge in successful development of biofuels. In these categories, the enzymes that have gained more importance in recent years include the endoglucanases, beta-glucosidases and alpha-galactosidases due to their high thermostability [36] and their high biomass hydrolysis efficiency [53–55].

Techno-economic studies have indicated that a cost-efficient biomass to ethanol conversion strategy requires a series of steps in biorefinery operation to be comprehensively addressed, including the use of more stable enzymes [22]. According to Gusakov et al. [56] the production cost of an enzyme is connected to the productivity of the enzyme-producing strains while the hydrolytic efficiency of the enzymes depends on both the properties of the individual enzymes and the synergies between them [22]. In order to improve the efficiency of these enzymes, several strategies have been implemented to obtain biocatalysts with improved performances. Although these strategies ranged from rational and computational designs to de novo enzyme designs, the highly demanding industrial conditions, in terms of extreme pH values and temperature, still limits the expression or/and activity of enzymes. Thus, the finding of more suitable blueprints to improve the catalytic properties of enzymes in a cost-efficient way can revolutionize the field of bioenergy [23,57].

The presence of a large number of enzymes from a high-temperature submarine hydrothermal vent site holds great promise for the identification of novel enzymes of high thermostability and improved catalytic properties. Further research on these data will provide valuable information to aid the discovery of endoglucanases, beta-glucosidases and alpha-galactosidases with potential in biofuel production.

4. Conclusions

Enzymes from extreme environments such as submarine volcanoes and hydrothermal vents, which are characterized by elevated temperatures and low pH conditions, are expected to display enhanced thermal or mechanical stability and better pH responses. Here, we explored the metagenomes of the extreme environments of the Santorini–Kolumbo volcanic complex. We mined a large number of gene-encoded enzymes involved in lignocellulosic biomass conversion to biofuels. Since this pool of enzymes originates from environments where temperature and pH resemble those of industrial settings, we anticipate that they can offer a cost-efficient solution in bioenergy production.

Author Contributions: P.N.P., P.N., H.Z., M.M., T.I.A., S.K., N.C.K., G.K., and A.M. contributed to the study design, data interpretation and article preparation. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the project “Centre for the study and sustainable exploitation of Marine Biological Resources (CMBR)” (MIS 5002670), which is implemented under the Action “Reinforcement of the Research and Innovation Infrastructure” and funded by the Operational Programme “Competitiveness, Entrepreneurship and Innovation” (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund), the project “Managing and Analysing Life Sciences Data (ELIXIR-GR)” (MIS 5002780), which is also cofinanced by Greece and the European Union-European Regional Development Fund and the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in the IMG/Mer data management system of JGI, DOE, Ca, USA under IMG submission ID numbers: RED1: 28983; RED2: 97772; RED3: 97773; WH1: 28984; WH2: 99695; CH1-V16c.out: 98329; CH1-V16c.mid: 98334; CH1-V16c.in: 98330; CH2: 89579; WAT1: 182564; WAT2: 89691. Their access and use should be conducted in compliance with the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization.

Acknowledgments: We acknowledge support of this work by project “Centre for the study and sustainable exploitation of Marine Biological Resources (CMBR)” (MIS 5002670) and project “Managing and Analysing Life Sciences Data (ELIXIR-GR)” (MIS 5002780). This research was supported in part through computational resources provided by IMBBC (Institute of Marine Biology, Biotechnology and Aquaculture) of the HCMR. Funding for establishing the IMBBC HPC has been received by the MARBIGEN (EU Regpot) project, LifeWatchGreece RI and the CMBR (Centre for the study and sustainable exploitation of Marine Biological Resources) RI. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dalmazo, G.Z.L.; Ferreira, D.; Vermelho, A.B. Marine Extremophiles: A source of hydrolases for biotechnological applications. *Mar. Drugs* **2015**, *13*, 1925–1965. [[CrossRef](#)]
2. Coker, J.A. Extremophiles and biotechnology: Current uses and prospects. *F100Research* **2016**, *5*, 396. [[CrossRef](#)] [[PubMed](#)]
3. Orcutt, B.; Sylvan, J.B.; Knab, N.J.; Edwards, K.J. Microbial ecology of the dark ocean above, at, and below the seafloor. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 361–422. [[CrossRef](#)] [[PubMed](#)]
4. Ivarsson, M.; Broman, C.; Sturkell, E.; Ormö, J.; Siljeström, S.; van Zuilen, M.; Bengtson, S. Fungal colonization of an Ordovician impact-induced hydrothermal system. *Nat. Sci. Rep.* **2013**, *3*, 3487. [[CrossRef](#)]
5. Oulas, A.; Polymenakou, P.N.; Seshadri, R.; Tripp, H.J.; Mandalakis, M.; Paez-Espino, D.A.; Pati, A.; Chain, P.; Nomikou, P.; Carey, S.; et al. Metagenomic investigation of the geologically unique Hellenic Volcanic Arc reveals a distinctive ecosystem with unexpected physiology. *Env. Microbiol.* **2016**, *18*, 1122–1136. [[CrossRef](#)] [[PubMed](#)]
6. Drake, H.; Ivarsson, M.; Bengtson, S.; Heim, C.; Siljeström, S.; Whitehouse, M.J.; Broman, C.; Belivanova, V.; Åström, M.E. Anaerobic consortia of fungi and sulfate reducing bacteria in deep granite fractures. *Nat. Commun.* **2017**, *8*, 55. [[CrossRef](#)]
7. Le Pichon, X.; Angelier, J. The Hellenic Arc and Trench system: A key to the neotectonic evolution of the Eastern Mediterranean area. *Tectonophysics* **1979**, *60*, 1–42. [[CrossRef](#)]
8. Nomikou, P.; Carey, S.; Papanikolaou, D.; Croff Bell, K.; Sakellariou, D.; Alexandri, M.; Bejelou, K. Submarine Volcanoes of the Kolumbo volcanic zone NE of Santorini Caldera, Greece. *Glob. Planet. Chang.* **2012**, *90–91*, 135–151. [[CrossRef](#)]
9. Kiliyas, S.P.; Nomikou, P.; Papanikolaou, D.; Polymenakou, P.N.; Godelitsas, A.; Argyraki, A.; Carey, S.; Gamaletsos, P.; Mertimekis, T.J.; Stathopoulou, E.; et al. New insights into hydrothermal vent processes in the unique shallow-submarine arc-volcano Kolumbo (Santorini), Greece. *Nat. Sci. Rep.* **2013**, *3*, 2421. [[CrossRef](#)]
10. Dando, P.R.; Aliani, S.; Arab, H.; Bianchi, C.N.; Brehmer, M.; Cocito, S.; Fowlers, S.W.; Gundersen, J.; Hooper, L.E.; Kölbh, R.; et al. Hydrothermal studies in the Aegean Sea. *Phys. Chem. Earth B* **2000**, *25*, 1–8. [[CrossRef](#)]
11. Dando, P.R.; Hughes, J.A.; Leahy, Y.; Niven, S.J.; Smith, C. Gas venting rates from submarine hydrothermal areas around the island of Milos, Hellenic Volcanic Arc. *Cont. Shelf Res.* **1995**, *15*, 913–929. [[CrossRef](#)]
12. Sigurdsson, H.; Carey, S.; Alexandri, M.; Vougioukalakis, G.; Croff, K.; Roman, C.; Sakellariou, D.; Anagnostou, C.; Rousakis, G.; Ioakim, C.; et al. Marine investigations of Greece’s Santorini volcanic field. *Eos* **2006**, *87*, 337–342. [[CrossRef](#)]
13. Hall-Spencer, J.M.; Rodolfo-Metalpa, R.; Martin, S.; Ransome, E.; Fine, M.; Turner, S.M.; Rowley, S.J.; Tedesco, D.; Buia, M.-C. Volcanic carbon dioxide vents show ecosystem effects of ocean acidification. *Nature* **2008**, *454*, 96–99. [[CrossRef](#)]
14. Carey, S.; Nomikou, P.; Croff-Bell, K.; Lilley, M.; Lupton, J.; Roman, C.; Stathopoulou, E.; Bejelou, K.; Ballard, R. CO₂ degassing from hydrothermal vents at Kolumbo submarine volcano, Greece, and the accumulation of acidic crater water. *Geology* **2013**, *41*, 1035–1038. [[CrossRef](#)]
15. Christakis, C.A.; Polymenakou, P.N.; Mandalakis, M.; Nomikou, P.; Kristoffersen, J.B.; Lampridou, D.; Kotoulas, G.; Mandalakis, M. Microbial community differentiation between active and inactive sulfide chimneys of the Kolumbo submarine volcano, Hellenic Volcanic Arc. *Extremophiles* **2018**, *22*, 13–27. [[CrossRef](#)] [[PubMed](#)]
16. Mandalakis, M.; Gabriilidou, A.; Polymenakou, P.N.; Christakis, C.A.; Nomikou, P.; Medvecký, M.; Kiliyas, S.P.; Kentouri, M.; Kotoulas, G.; Magoulas, A. Microbial strains isolated from CO₂-venting Kolumbo submarine volcano show enhanced co-tolerance to acidity and antibiotics. *Mar. Env. Res.* **2019**, *144*, 102–110. [[CrossRef](#)] [[PubMed](#)]

17. Christakis, C.A.; Polymenakou, P.N.; Kiliyas, S.P.; Nomikou, P.; Nielsen, T.; Kyrpides, N.; Kristoffersen, J.B.; Kotoulas, G.; Magoulas, A. Metagenomic evidence of microbial iron oxidation in the iron-mat ecosystem of the Kallisti Limnes low-temperature CO₂-rich seafloor hydrothermal pools, Santorini. Unpublished work. 2021.
18. Christakis, C.A.; Polymenakou, P.N.; Kiliyas, S.P.; Nomikou, P.; Nielsen, T.; Kyrpides, N.; Kristoffersen, J.B.; Kotoulas, G.; Magoulas, A. Metagenomic exploration of three different microbial mat layers of a polymetallic gas chimney, Kolumbo volcano, Hellenic Volcanic Arc. Unpublished work. 2021.
19. Barrangou, R.; Horvarth, P. A decade of discovery: CRISPR functions and applications. *Nat. Microbiol.* **2017**, *2*, 17092. [[CrossRef](#)]
20. Lauria, M.; Molinari, F.; Motto, M. *Genetic Strategies to Enhance Plant Biomass Yield and Quality—Related Traits for Bio-Renewable Fuel and Chemical Productions*; Intechopen: London, UK, 2014. [[CrossRef](#)]
21. Wang, H.C.; Chen, Y.C.; Hseu, R.S. Purification and characterization of a cellulolytic multienzyme complex produced by *Neocallimastix patriciarum* J11. *Biochem. Biophys. Res. Commun.* **2014**, *451*, 190–195. [[CrossRef](#)]
22. Sukumaran, K.; Abraham, A.; Mathew, A.K. Enzymes for Bioenergy. In *Bioresources and Bioprocess in Biotechnology*; Sugathan, S., Pradeep, N.S., Abdulhameed, S., Eds.; Springer: Singapore, 2017; Volume 2, pp. 3–44.
23. Barruetabeña, N.; Alonso-Lerma, B.; Galera-Prat, A.; Joudeh, N.; Barandiaran, L.; Aldazabal, L.; Arbulu, M.; Alcalde, M.; De Sancho, D.; Gavira, J.A.; et al. Resurrection of efficient Precambrian endoglucanases for lignocellulosic biomass hydrolysis. *Comm. Chem.* **2019**, *2*, 76. [[CrossRef](#)]
24. Sticklen, M.B. Plant genetic engineering for biofuel production: Towards affordable cellulosic ethanol. *Nat. Rev. Genet.* **2008**, *9*, 433–443. [[CrossRef](#)]
25. Zhao, X.; Zhang, L.; Liu, D. Biomass recalcitrance Part I: The chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels Bioprod. Biorefin.* **2012**, *6*, 465–482. [[CrossRef](#)]
26. Gírio, F.M.; Fonseca, C.; Carvalheiro, F.; Duarte, L.C.; Marques, S.; Bogel-Lukasik, R. Hemicelluloses for fuel ethanol: A review. *Bioresour. Technol.* **2010**, *101*, 4775–4800. [[CrossRef](#)]
27. Kumar, P.; Barrett, D.M.; Delwiche, M.J.; Stroeve, P. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. *Ind. Eng. Chem. Res.* **2009**, *48*, 3713–3729. [[CrossRef](#)]
28. Schulein, M. Cellulases of *Trichoderma reesei*. In *Methods in Enzymology*; Wodd, W.A., Abelson, J.N., Eds.; Elsevier: New York, NY, USA, 1998; Volume 160, pp. 234–242.
29. Jonsson, V. Statistical Analysis and Modelling of Gene Count Data in Metagenomics. Ph.D. Thesis, Clamers University of Technology and Göteborg University, Gothenburg, Sweden, 2017.
30. Nielsen, H.; Tsigirigos, K.D.; Brunak, S.; von Heijne, G. A brief history of protein sorting prediction. *Protein J.* **2019**, *38*, 200–216. [[CrossRef](#)]
31. Kozlov, A.M.; Darriba, D.; Flouri, T.; Morel, B.; Stamatakis, A. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **2019**, *35*, 4453–4455. [[CrossRef](#)]
32. Minh, B.Q.; Schmidt, H.A.; Chernomor, O.; Schrempf, D.; Woodhams, M.D.; von Haeseler, A.; Lanfear, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **2020**, *37*, 1530–1534. [[CrossRef](#)]
33. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v4: Recent updates and new developments. *Nucleic Acids Res.* **2019**, *47*, W256–W259. [[CrossRef](#)] [[PubMed](#)]
34. Davidson, S. Sustainable bioenergy: Genomics and biofuels development. *Nat. Educ.* **2008**, *1*, 175.
35. Sukumaran, R.K.; Singhanian, R.R.; Pandey, A. Microbial cellulases—Production, applications and challenges. *J. Sci. Ind. Res.* **2005**, *64*, 832–844.
36. Yennamalli, R.M.; Rader, A.J.; Kenny, A.J.; Wolt, J.D.; Sen, T.Z. Endoglucanases: Insights into thermostability for biofuel applications. *Biotechnol. Biofuels* **2013**, *6*, 136. [[CrossRef](#)]
37. Chen, H.; Wang, L. Chapter 4—Enzymatic hydrolysis of pretreated biomass. In *Technologies for Biochemical Conversion of Biomass*; Academic Press: Cambridge, CA, USA, 2017; pp. 65–99.
38. Xue, C.; Cheng, C. Chapter Two—Butanol production by *Clostridium*. *Adv. Bioenergy* **2019**, *4*, 35–77.
39. Kim, D.-S.; Chi, W.-J.; Hong, S.-K. Molecular characterization of an endo- β -1,4-glucanase, CelAJ93, from the recently isolated marine bacterium, *Cellulophaga* sp. J9-3. *Appl. Sci.* **2019**, *9*, 4061. [[CrossRef](#)]
40. Thomas, L.; Ram, H.; Singh, V.P. Inducible cellulase production from an organic solvent tolerant *Bacillus* sp. SV1 and evolutionary divergence of endoglucanase in different species of the genus *Bacillus*. *Braz. J. Microbiol.* **2018**, *49*, 429–442. [[CrossRef](#)] [[PubMed](#)]
41. Pang, H.; Zhang, P.; Duan, C.-J.; Mo, X.-C.; Tang, J.-L.; Feng, J.-X. Identification of cellulase genes from the metagenomes of compost soils and functional characterization of one novel endoglucanase. *Curr. Microbiol.* **2009**, *58*, 404–408. [[CrossRef](#)] [[PubMed](#)]
42. Zhang, L.; Fan, Y.; Zheng, H.; Du, F.; Zhang, K.-Q.; Huang, X.; Wang, L.; Zhang, M.; Niu, Q. Isolation and characterization of a novel endoglucanase from a *Bursaphelenchus xylophilus* metagenomic library. *PLoS ONE* **2013**, *8*, e82437. [[CrossRef](#)] [[PubMed](#)]
43. Kim, H.W.; Ishikawa, K. Structure of hyperthermophilic endocellulase from *Pyrococcus horikoshii*. *Proteins Struct. Funct. Bioinforma.* **2009**, *78*, 496–500. [[CrossRef](#)]
44. Suleiman, M.; Schröder, C.; Klippel, B.; Schäfers, C.; Krüger, A.; Antranikian, G. Extremely thermoactive archaeal endoglucanase from a shallow marine hydrothermal vent from Vulcano Island. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 1267–1274. [[CrossRef](#)]

45. Tallapragada, P.; Dikshit, R. Chapter 11—Microbial Production of Secondary Metabolites As Food Ingredients. In *Microbial Production of Food Ingredients and Additives*; Handbook of Food Bioengineering; Academic Press: Cambridge, CA, USA, 2017; pp. 317–345.
46. Yeoman, C.J.; Han, Y.; Dodd, D.; Schroeder, C.M.; Mackie, R.I.; Cann, I.K.O. Thermostable enzymes as biocatalysts in the biofuel industry. *Adv. Appl. Microbiol.* **2010**, *70*, 1–55.
47. Molina, G.; Contesini, F.J.; de Melo, R.R.; Sato, H.H.; Pastore, G.M. Chapter 11— β -Glucosidase from *Aspergillus*. In *New and Future Developments in Microbial Biotechnology and Bioengineering*; Gupta, V.K., Ed.; Elsevier: New York, NY, USA, 2016; pp. 155–169. ISBN 9780444635051.
48. Dombrowski, N.; Teske, A.P.; Baker, B.J. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nat. Commun.* **2018**, *9*, 4999. [[CrossRef](#)]
49. Seitz, K.W.; Dombrowski, N.; Eme, L.; Spang, A.; Lombard, J.; Sieber, J.R.; Teske, A.P.; Ettema, T.J.G.; Baker, B.J. Asgard archaea capable of anaerobic hydrocarbon cycling. *Nat. Commun.* **2019**, *10*, 1822. [[CrossRef](#)] [[PubMed](#)]
50. Álvarez-Cao, M.E.; Cerdán, M.E.; González-Siso, M.I.; Becerra, M. Bioconversion of beet molasses to alpha-galactosidase and ethanol. *Front Microbiol.* **2019**, *10*, 405. [[CrossRef](#)] [[PubMed](#)]
51. Wissuwa, J.; Stokke, R.; Fedoy, A.E.; Lian, K.; Smalas, A.O.; Steen, I.H. Isolation and complete genome sequence of the thermophilic *Geobacillus* sp. 12AMOR1 from an Arctic deep-sea hydrothermal vent site. *Stand Genom. Sci.* **2016**, *11*, 16. [[CrossRef](#)] [[PubMed](#)]
52. Marteinson, V.T.; Birrien, J.L.; Reysenbach, A.L.; Vernet, M.; Marie, D.; Gambacorta, A.; Messner, P.; Sleytr, U.B.; Prieur, D. *Thermococcus barophilus* sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int. J. Syst. Bacteriol.* **1999**, *49*, 351–359. [[CrossRef](#)] [[PubMed](#)]
53. Kao, M.R.; Kuo, H.W.; Lee, C.C.; Huang, K.-Y.; Huang, T.-Y.; Li, C.-W.; Chen, C.W.; Wang, A.H.-J.; Yu, S.M.; Ho, T.-H.D. *Chaetomella Raphigera* β -glucosidase D2-BGL has intriguing structural features and a high substrate affinity that renders it an efficient cellulase supplement for lignocellulosic biomass hydrolysis. *Biotechnol. Biofuels* **2019**, *12*, 258. [[CrossRef](#)] [[PubMed](#)]
54. Rodrigues-Dutra, T.; Monteze-Guimarães, V.; Mascarenhas-Varela, E.; da Silva-Fialho, L.; Ferreira-Milagres, A.M.; Falkoski, D.L.; Zanuncio, J.C.; de Rezende, S.T. A *Chrysoporthe cubensis* enzyme cocktail produced from a low-cost carbon source with high biomass hydrolysis efficiency. *Sci. Rep.* **2017**, *7*, 3893. [[CrossRef](#)]
55. Srivastava, N.; Rathour, R.; Jha, S.; Pandey, K.; Srivastava, M.; Thakur, V.K.; Sengar, R.S.; Gupta, V.K.; Mazumder, P.B.; Khan, A.F.; et al. Microbial beta glucosidase enzymes: Recent advances in biomass conversion for biofuels application. *Biomolecules* **2019**, *9*, 220. [[CrossRef](#)]
56. Gusakov, A.V.; Salanovich, T.M.; Antonov, A.I.; Ustinov, B.B.; Okunev, O.N.; Burlingam, R.P.; Emalfarb, M.A.; Baez, M.A.; Sinitsyn, A.P.; Arkady, P. Construction of Highly Efficient Cellulase Compositions for Enzymatic Hydrolysis of Cellulose. U.S. Patent 8916363, 23 December 2014.
57. Maki, M.; Leung, K.T.; Qin, W. The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *Int. J. Biol. Sci.* **2009**, *5*, 500–516. [[CrossRef](#)]