FEMS Microbiology Letters, 366, 2019, fny285

doi: 10.1093/femsle/fny285
Advance Access Publication Date: 5 December 2018
Perspective

PERSPECTIVE – Biotechnology & Synthetic Biology

Decoding the ocean's microbiological secrets for marine enzyme biodiscovery

Manuel Ferrer[1,*], Celia Méndez-García[1], Rafael Bargiela[2], Jennifer Chow[3], Sandra Alonso[1], Antonio García-Moyano[4], Gro E.K. Bjerga[4], Ida H. Steen[5], Tatjana Schwabe[6], Charlotte Blom[7], Jan Vester[7], Andrea Weckbecker[8], Patrick Shahgaldian[9], Carla C.C.R. de Carvalho[10], Rolandas Meskys[11], Giulio Zanaroli[12], Frank O. Glöckner[13,14], Antonio Fernández-Guerra[13], Siva Thambisetty[15], Fernando de la Calle[16], Olga V. Golyshina[2,17], Michail M. Yakimov[18,19], Karl-Erich Jaeger[20], Alexander F. Yakunin[2,17,21], Wolfgang R. Streit[3], Oonagh McMeel[22], Jan-Bart Calewaert[22], Nathalie Tonné[22], Peter N. Golyshin[2,17] and The INMARE Consortium

[1]Department of Applied Biocatalysis, Institute of Catalysis, Consejo Superior de Investigaciones Científicas, Marie Curie 2, 28049 Madrid, Spain, [2]School of Natural Sciences, Bangor University, Deiniol Road, LL57 2UW Bangor, United Kingdom, [3]Biozentrum Klein Flottbek, Mikrobiologie & Biotechnologie, Universität Hamburg, Ohnhorststr. 18, 22609 Hamburg, Germany, [4]NORCE Environment, NORCE Norwegian Research Centre AS, Thormøhlens gate 55, 5008 Bergen, Norway, [5]Department of Biological Sciences and KG Jebsen Centre for Deep Sea Research, University of Bergen, Thormøhlensgt 53A/B, 5020 Bergen, Norway, [6]CLIB2021 - Cluster industrielle Biotechnologie, Voelklinger Str. 4, 40219 Düsseldorf, Germany, [7]Novozymes A/S, Krogshoejvej 36, 2880 Bagsvaerd, Denmark, [8]evoxx technologies GmbH, Alfred-Nobel-Str. 10, 40789 Monheim am Rhein, Germany, [9]School of Life Sciences, University of Applied Sciences and Arts Northwestern Switzerland, Hofackerstrasse 35, CH-4132 Muttenz, Switzerland, [10]iBB - Institute for Bioengineering and Biosciences, Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais, 1049-001 Lisboa, Portugal, [11]Department of Molecular Microbiology and Biotechnology, Institute of Biochemistry, Life Sciences Center, Vilnius University, Sauletekio 7, LT-10257 Vilnius, Lithuania, [12]Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, via Terracini 28, 40131 Bologna, Italy, [13]Jacobs University Bremen gGmbH, Campus Ring 1, 28759 Bremen, Germany, [14]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, 28359 Bremen, Germany, [15]London School of Economics and Political Science, Houghton Street, WC2A 2AE London, United Kingdom, [16]Microbiology R&D Dpt., Pharma Mar, S.A., Avda. Los Reyes, 1, 28770 Colmenar Viejo, Madrid, Spain, [17]Centre for Environmental Biotechnology, Bangor University, Deiniol Road, LL57 2UW Bangor, United Kingdom, [18]Institute for Biological Resources and Marine Biotechnology, IRBIM-CNR, Spianata S. Raineri 86, 98122

Received: 7 November 2018; **Accepted:** 4 December 2018

Messina, Italy, [19]Institute of Living Systems, Immanuel Kant Baltic Federal University, Nevskogo 14[a], 236016 Kaliningrad, Russia, [20]Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf and Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Strasse, 52428 Jülich, Germany, [21]Department of Chemical Engineering and Applied Chemistry, University of Toronto, 200 College Street, Toronto, M5S 3E5 Ontario, Canada and [22]Seascape Belgium bvba, Kindermansstraat 14/19, 1000 Brussels, Belgium

*Corresponding author: Institute of Catalysis, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain. Tel: +34 915854872; E-mail: mferrer@icp.csic.es

One sentence summary: The success of using marine resources for the development of industrial products and processes depends on the manner they are screened, produced and tested on a larger scale.

Editor: Michael Sauer

ABSTRACT

A global census of marine microbial life has been underway over the past several decades. During this period, there have been scientific breakthroughs in estimating microbial diversity and understanding microbial functioning and ecology. It is estimated that the ocean, covering 71% of the earth's surface with its estimated volume of about 2×10^{18} m^3 and an average depth of 3800 m, hosts the largest population of microbes on Earth. More than 2 million eukaryotic and prokaryotic species are thought to thrive both in the ocean and on its surface. Prokaryotic cell abundances can reach densities of up to 10^{12} cells per millilitre, exceeding eukaryotic densities of around 10^6 cells per millilitre of seawater. Besides their large numbers and abundance, marine microbial assemblages and their organic catalysts (enzymes) have a largely underestimated value for their use in the development of industrial products and processes. In this perspective article, we identified critical gaps in knowledge and technology to fast-track this development. We provided a general overview of the presumptive microbial assemblages in oceans, and an estimation of what is known and the enzymes that have been currently retrieved. We also discussed recent advances made in this area by the collaborative European Horizon 2020 project 'INMARE'.

Keywords: biodiversity; enzyme; biotechnology; marine diversity; omics

INTRODUCTION

The ocean is one of the greatest unexplored frontiers to humankind. From extremely high pressures in the deepest parts of the ocean to normal atmospheric pressure on the surface, from temperatures above 300°C in hydrothermal vents to sub-zero temperatures in sea ice, and from low-salinity conditions to salt-saturated brines, this environment, providing a range of conditions often hostile to many life forms, accommodates diverse communities of microorganisms that have adapted to such challenging settings. The metabolic diversity of these microorganisms promises a source of enzymes able to perform uniquely in industrial settings where harsh physical and chemical conditions are encountered (Harrison *et al.* 2013).

The market for industrial enzymes for non-therapeutic uses (foods, detergents, textiles, pulp and paper) has doubled over the past 15 years and has been recently estimated to be in the region of US$4 billion per annum (Evans 2013). In 2019, a turnover of about US$10 billion is expected with the production of industrial enzymes through the application of bulk enzymes in different market sectors with growing demand. In addition to this, up to 40% of the chemical synthesis processes that require environmentally damaging bulk organic solvents and high-energy inputs could employ enzymatic catalysis by 2030, creating a huge demand for new industrially relevant enzymes (Martínez-Martínez, Bargiela and Ferrer 2017). Rapid technological advances in omics and bioinformatics have revolutionised the exploration of nature's microbial diversity for industrially relevant enzymes and biochemical processes (for a recent example using complementary techniques, see Stumberger *et al.* 2016). However, until very recently, the isolation of marine microbes and their gene products using classical cultivation methods and biochemistry was work-intensive and time-consuming. Today, molecular technology has reached a high level of sophistication and has enabled a quantum leap in our knowledge of the marine microbial world (Salazar and Sunagawa 2017). The establishment of international consortia and ocean sampling campaigns to explore marine microbial diversity has contributed significantly to this technological breakthrough. In addition to our growing knowledge on the scales of marine microbial diversity, we are also gaining a clearer insight into the basis for the molecular adaptation of microorganisms to diverse conditions. This is associated with our increasing knowledge on the diversity of microbial gene products and how they contribute to regulating the metabolic state under diverse conditions (Glöckner *et al* 2012). This can include, for example, the recruitment of alternate pathways for carbon cycling and innovations across metabolic sub-systems and the tree of life by maximising growth rate, efficiency and evolutionary progress, favouring genome minimisation and streamlining (Braakman and Smith 2014). From a more applied point of view, microbial enzymes represent tools for biotechnological processes, some of which may find their way to industry (Costessi *et al.* 2018) and commercialisation (Sherkow 2017). In this article, we discuss the accumulated knowledge on marine microbial assemblages; the factors that have limited their screening and use as a source of novel enzymes and other gene products; and the importance of improving our understanding of how this diversity can be used, understood and sustainably exploited.

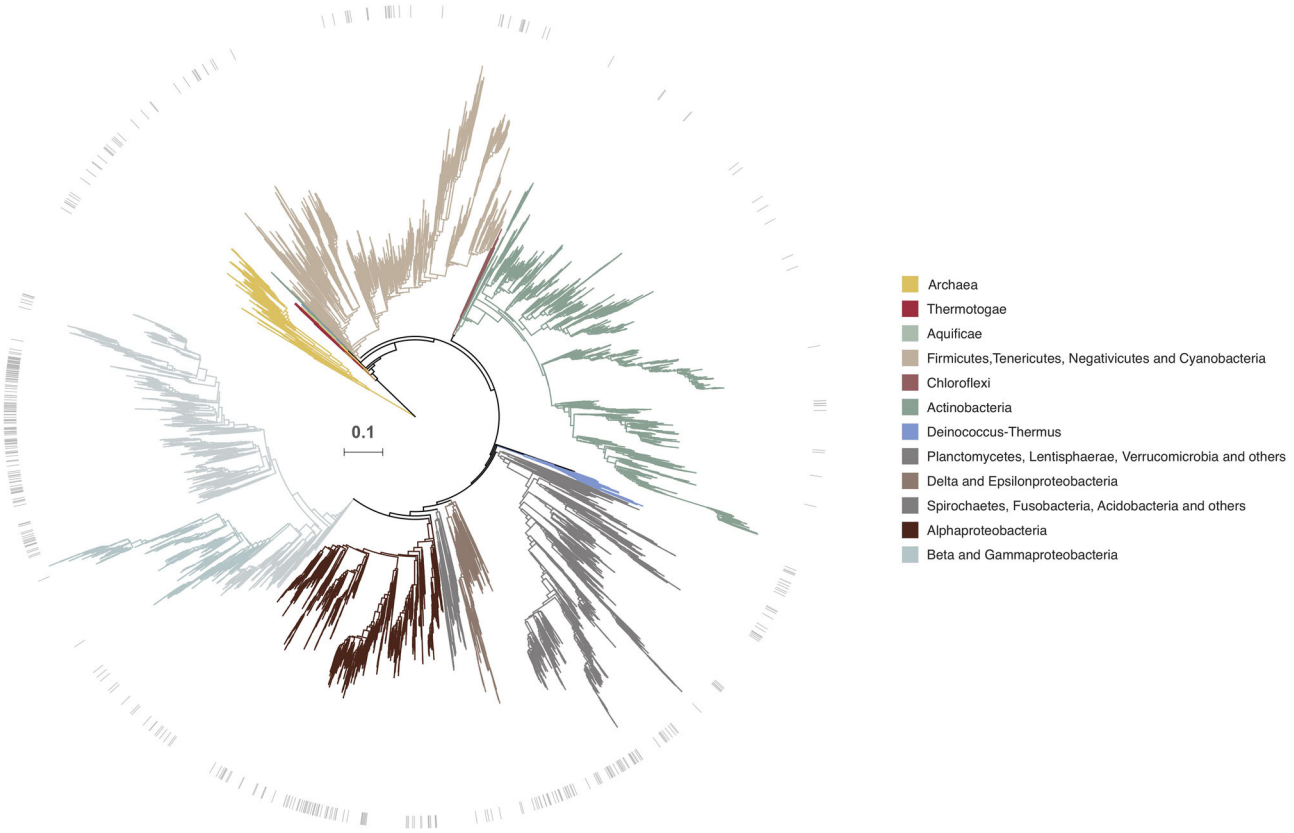

Figure 1. Phylogenetic tree representing the diversity of marine species spanning across the three domains of life. The figure represents the all-species living tree, release 132 (Yarza *et al.* 2010; Munoz *et al.* 2011), including all sequenced type strains belonging to the Archaea and Bacteria. The microorganisms of marine origin are shown in the outermost ring with grey lines. The tree comprises 13 903 leaves, from which 469 belong to microbial species isolated from marine environments. Major taxonomic clusters are highlighted by colours (see the legend). The scale bar represents the number of changes per site. The indications in the outer circle point the species from which full genome sequence, draft genome and/or with reviewed proteins in Swiss-Prot are available.

GENOMES AND ENZYMES OF MARINE MICROORGANISMS

The assessment of microbial diversity for all kingdoms of life on Earth, based on the World Register of Marine Species (WORMS) (www.marinespecies.org), predicts the existence of ~8.7 ± 1.3 million species globally, of which ~2.2 ± 0.18 million are marine species. In 2011, about 91% of the species (mostly eukaryotes) in the ocean still awaited description (Mora *et al.* 2011). Restricting this prediction to prokaryote taxa resulted in at least ~1.3 million species in the world's oceans, of which about half have been catalogued. The taxonomic census of all species present in the most recent release of the Living Tree Project database (Yarza *et al.* 2010) reflects the high diversity of cultured marine microorganisms (Fig. 1). The assessment of microbial diversity for all marine microbial genomes, regardless of the level of completeness based on MarDB (https://mmp.sfb.uit.no/databases/mardb/), may give an indication of marine microbial species with cultivable representatives for which genome sequence information is available. Understanding and analysing these genetic resources may not only help in understanding marine microbial ecology and adaptation mechanisms, but could also provide a source of novel enzymes.

The current version of the SILVA rRNA database project (release 132, www.arb-silva.de) contains sequence information for manually curated complete or partial genomes of about 8 900 cultured marine microorganisms (Fig. 1). Based on UniProt Knowledgebase (UniProtKB) (www.uniprot.org/), one can further gain estimates on those marine microbes with available genome sequences from which enzymatic biochemical knowledge has also been catalogued. This can be seen in Fig. 1, which shows that proteins from cultured and genome-sequenced marine microbial representatives, belonging to ~173 genera, have been catalogued in UniProtKB. This biochemical knowledge is also limited since in ~57% of those cases, less than 10 enzymes have been retrieved. This represents a global picture of a general paucity of biochemical knowledge about marine microbes. Considering the estimated diversity of prokaryotic species in the global ocean, approximately 1.3 million, only few microbes have been screened for their enzymes, whose number is also low. Microbial and enzymatic technologies are currently not a limiting factor in the advancement of new applications of marine microbial products in industrial biotechnology. The critical constraint on the path of making progress is that we are still far from cataloguing our microbial and, more important, our enzymatic bioresources to find those that can be further investigated for use in industry.

MARINE METAGENOMES

One of the most extensive global surveys of microbial diversity to date was carried out by the Earth Microbiome Project (http://www.earthmicrobiome.org/); it has gained information on at least 5.6 million non-redundant Operational Taxonomic

Units (OTUs) present in at least 15 000 samples. Extrapolating this genetic information to the level of proteins has been less successful, with almost no information available about enzymes associated with those OTUs. A similar situation was observed after examining available data compiled from marine circumnavigation ocean sampling expeditions. Some examples are as follows: the Sargasso Sea Expeditions (https://www.mbari.org/at-sea/expeditions/sargasso-sea-expeditions), the Sorcerer II Global Ocean Sampling Expedition (https://www.jcvi.org/gos), Tara Oceans (https://www.embl.de/tara-oceans/start/research/index.html), Malaspina (http://www.expedicionmalaspina.es/), Galathea 3 (http://www.galathea3.dk/uk.html), Global Ocean Sampling Expedition (https://www.ncbi.nlm.nih.gov/books/NBK6855/), as well as many other smaller scale expeditions (https://www.ebi.ac.uk/metagenomics/). These expeditions have collected and analysed hundreds of thousands of samples, resulting in a profusion of information about the 16S rRNA gene sequences and expansion of our appreciation of microbial biodiversity. However, associated information about the inherent functional biochemistry is rather scarce. This disparity between taxonomic and biochemical information can be explained mainly by recent advances in sequencing technology, allowing for large-scale, inexpensive and fast sequencing of uncultured biodiversity, with the most rate-limiting step in this process being the bioinformatic analysis. Unlocking the biochemical or biotechnological potential of these sequences, however, requires laborious wet lab work, including extensive cloning of genes of interest, followed by the expression and characterisation of enzymes. This is not a trivial exercise, given that not all genes in a metagenome assembly are full length or can be successfully cloned and expressed. The choice of methodology for enzyme discovery between genomic data predictions and naïve enzyme screening approaches remains open. Naïve screens allow identification of genes encoding enzymes with functions that were not known to humankind earlier, but sequencing followed by *in silico* predictions allows us to increase the speed of discovery, regardless of the novelty of the enzymes. Of course, it should be stressed that novelty itself does not guarantee better enzymatic performance and better opportunities of commercialisation.

ENZYMES: FROM DISCOVERY TO MANUFACTURING

Cataloguing and characterising microbial and enzymatic diversity is tedious and time-consuming. Thus, about 11 000 microbial species from cultivable representatives, including marine species, have been properly described and deposited in databases (Yarza *et al.* 2010, 2014). At the current rate of circa 600 new descriptions annually, it is expected that cataloguing all predicted microbial species will take at least another 1000 years. If these predictions are extrapolated to enzymes, the time frame will be even more extensive. At the present discovery rate, it would take about 1 million years to sequence all microbial genomes, retrieve the information of sequences encoding enzymes and further characterise them. This is without considering the analysis of enzymes encoded by sequences annotated as hypotheticals.

In addition, not all microbial products require the same time frame for discovery and marketing. Thus, based on the experience of Pharma Mar, S.A., the world's leading biopharmaceutical company in the discovery of marine-derived oncology drugs, it takes about 15–20 years to discover and commercialise one bioactive microbial product for clinical applications, from a starting number of several thousands of pre-selected analogues. The total cost of all the needed steps, including discovery, preclinical and clinical studies and approval, and sales, may reach up to ~US$802 million, with only two out of five molecules recovering the cost. In the case of bulk enzymes for applications in different market sectors, this time frame is shorter, but still challenging from an industrial perspective. Actually, it is difficult to estimate the time frame needed to discover and produce each marketable enzyme, but real examples on the web (http://www.biocatalysts.com/enzyme-development-manufacture/) suggest the following: 1 week for investigating whether a continuous supply of a novel enzyme at right price can allow running a manufacturing process; an undetermined time to find the right enzyme by applying bioinformatics, cloning and metagenomics technologies; 11 weeks for selecting and producing at a scale of 0.5–1 g best 20 candidates for testing; 8 to 12 weeks for delivering a sample of active and soluble enzyme with technical specifications; 8 to 12 weeks for proving that the enzyme performs well under manufacturing conditions so that it fits into commercial requirements; and 30 weeks for producing a standard enzyme product applying a validated, regulated and quality-assured routine manufacturing process. This accounts for at least 58–66 weeks plus the undetermined time, a priory challenging to be predicted, needed to screen the untapped biological diversity in the microbial world, including marine bioresources, to find the right enzymes. It is important to mention that the average cost of commercial enzymes is about US$400/kg, and that although estimations of the total cost for all the needed steps, including discovery, cloning, expression, characterisation, biocatalysis studies and approval, and sales, are difficult to make, a minimum cost of US$10 000 is expected.

INMARE: A EUROPEAN COLLABORATION PROVIDING A FASTER ROUTE TO IDENTIFICATION OF LEAD PRODUCTS

Research and development arising from large sampling expeditions, microbial cultivation efforts, and genome and metagenome sequencing have paved the way for future work in the field of enzymology and microbial-associated functions. However, there is a significant gap in knowledge related to the large-scale analysis of enzyme activities in microorganisms, including marine ones. There is a need to invest an effort to decode the enzyme content of marine microorganisms, independent of their amenability to cultivation. The understanding of cultured or uncultured marine microorganisms that encode enzymes with broad or narrow substrate spectra, higher or lower levels of chiral selectivity, solvent-resistance or solvent-activity and stability, and low- or high-pH/temperature activity is of increasing demand. Having this information framed in the context of some larger scale processes, such as carbon cycling and productivity, would help in articulating the relevance of enzymes in the context of microbial ecology and of biogeochemical cycles that underpin the functioning of marine ecosystems, and ultimately the Earth's biosphere. Additionally, compiling more information on marine enzymes will provide new biocatalysts for future developments and help in predicting enzyme properties from accumulated biochemical data. Advancing our knowledge in these areas can be most effectively addressed through the collaboration of academic/research institutes and industrial partners, particularly those already leading the process of streamlining and shortening pipelines of the gene product (enzymes and bioactives) discovery.

Table 1. Pipeline representing the technical solutions provided by INMARE to significantly shorten the time needed from discovery to application in different technological levels.

Technologies to be implemented	Technology readiness levels (TRL)[1]	Principal activities
(i) Commercialisation (ii) Intellectual properties rights protection	TRL 9	(i) Novel enzymatic applications (ii) Novel biocatalytic processes
Business interaction	TRL 8	
Testing of candidates under real process conditions	TRL 7	Scale-up application and pilot-scale process
Enzyme collection: characterization and optimization	TRL 6	Enzyme candidates all-rounders
(i) High-quality crystallisation and structural analysis facilities (ii) Bioanalytical and bioprocess engineering facilities and expertise	TRL 5	Secondary screening and enzyme engineering
High-edge sequence annotation pipeline and bioinformatics resources	TRL 4	Pre-characterised enzyme library—candidate set
Innovative enzyme screening assays and platforms	TRL 3	Sequence and activity-based screening
State-of-art technologies for the construction of metagenomics libraries	TRL 2	Metagenomic and genomic libraries
Advanced technologies to access and sample unique marine biodiversity hotspots	TRL 1	Unique marine biodiversity resources

[1]TRL 9: system proven in operational environment; TRL 8: system completed and qualified; TRL 7: system prototype demonstration in operational environment; TRL 6: technology demonstrated in industrially relevant environment; TRL 5: technology validated in industrially relevant environment; TRL 4: technology validated in lab; TRL 3: experimental proof of concept; TRL 2: technology concept formulated; TRL 1: basic principles observed.

With these objectives, a pan-European consortium of partners was established in 2015 to work together on the EU project 'INMARE' (**In**dustrial Application of **Mar**ine **E**nzymes). Funded by the EU Horizon 2020 Research and Innovation Programme, INMARE aims to explore marine biodiversity in diverse environments to find novel enzymes to meet the needs of the growing industrial enzyme market. Led by Bangor University in Wales and comprising 24 multidisciplinary academic and industrial partners from 12 countries, the 4-year INMARE project adopted a multifaceted approach to achieve its aims, which is summarised in Table 1. It focused on samples isolated from already known or novel biodiversity hotspots (hypersaline, thermophilic, subsurface, deep-sea, low-pH vents, etc.). It has established *in vitro* and *in vivo* screening platforms to identify relevant gene products, as well as constructing sequence analysis pipelines to target enzymes of industrial relevance—thus streamlining the identification of genes of interest from sequence data. Associating unknown sequences with enzyme activities of interest was also a priority, given that most sequence databases are full of sequences encoding genes with unknown functions. Producing biocatalytic data was a key priority, which required expanding the spectrum of hosts for heterologous protein expression, thus shortening the cumbersome phase of enzyme optimisation by testing enzyme candidates under application conditions at the very early discovery phase. Finally, INMARE has increased our ability to analyse and compare biocatalytic data from highly diverse enzymes to predict more rapidly functions from sequencing data. This involved identifying enzymes with appropriate activity phenotypes by screening a large enzyme collection against arrays of a few hundred compounds representing challenging chemical steps in real applications, resulting in shortlisting priority targets, as well as understanding enzyme properties and ways to predict them. Engineering and immobilisation, which are important issues for industrial applications of enzymes, are the activities also covered within the project. The innovation comes not only by the technical developments but also by the application of a pipeline, which covers all steps from discovery to the end use. Only by screening a large diversity of marine microbes producing a large repertoire of marine enzymes, one can guarantee deciphering the ocean's microbial microbiome in a frame with biotechnology advances.

In its first phase, INMARE has contributed, through the development of advanced methods for biotechnological blueprinting of marine microbes, to the creation of one of the largest collections of genomic and metagenomic enzymes, with the focus on those most requested at the industrial level. Currently, this collection accounts for about 1000 enzymes, of which circa 94% are available in ready-to-use expression systems and 32% have been fully characterised (see, for recent examples, Popovic *et al*. 2017 and Martínez-Martínez *et al*. 2018).

These enzymes are quite diverse and non-redundant in terms of their amino acid sequences and geographical distribution and taxonomic origin (Fig. 2). They demonstrate a sequence identity of as low as 5% compared to sequences in databases and comprise enzymes from at least 193 geographically distinct sites, at least 283 known bacterial and archaeal genera, and approximately 607 microorganisms with unclear taxonomic affiliations. Condensing and streamlining the arduous phase of the optimisation process allowed the identification of lead products and prototypes with potential for IP protection and fast commercialisation. This can be achieved through subjecting enzyme candidates to multiple tests to assess their industrial applicability at the earliest stages of discovery, and to stabilisation by engineering and immobilisation and testing using conditions to mimic those required by industry.

OUTLOOK

Ocean's microbial diversity remains undersampled and poorly understood. Microbiology, molecular biology, biochemical and bioprocess technologies are diverse, but all need to be further implemented and integrated. INMARE is an example of how the

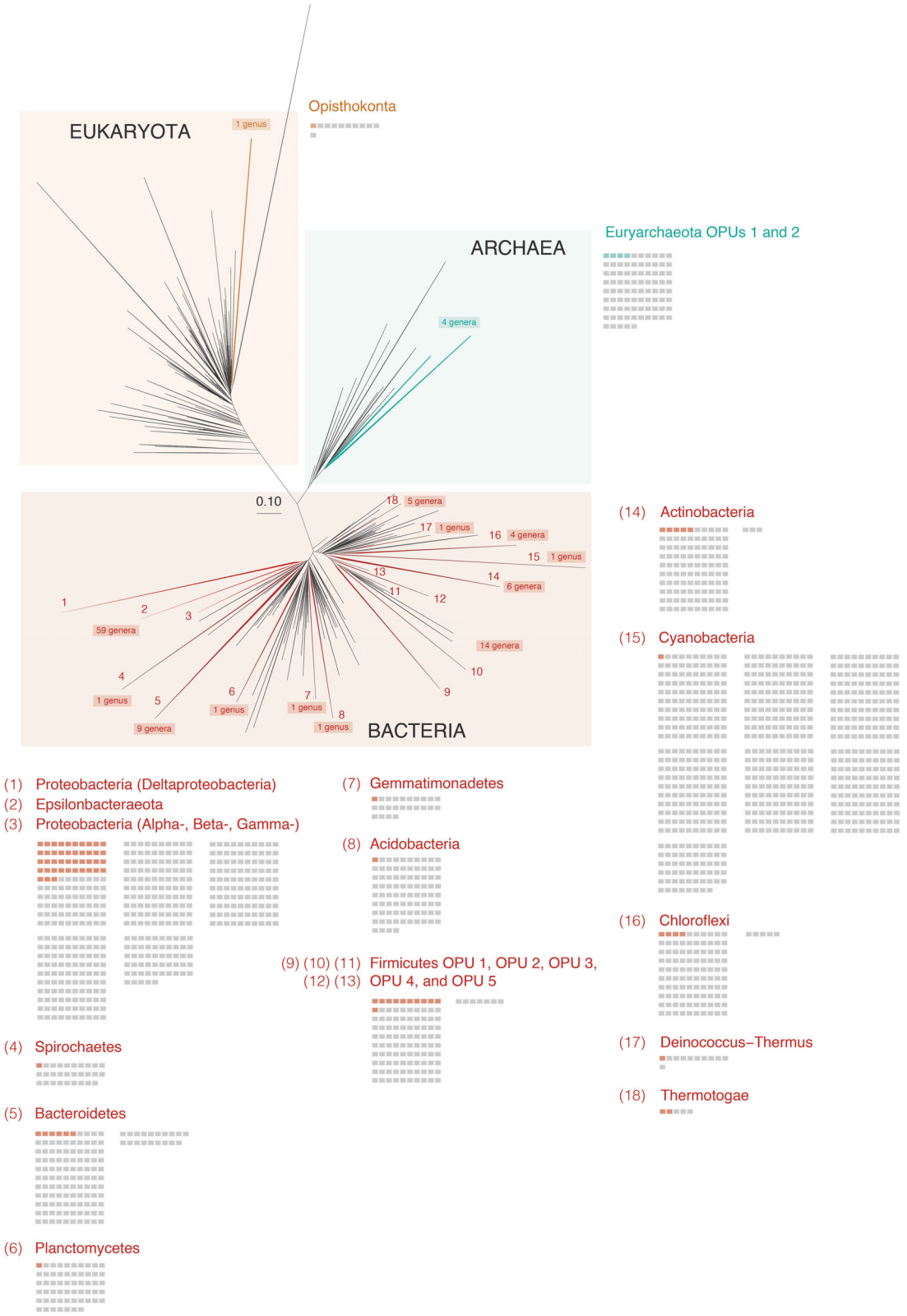

Figure 2. Phylogenetic tree representing the diversity of studied enzymes spanning the three domains of life. The database used for the reconstruction corresponds to the release number 132 from the non-redundant SILVA database for the small sub-unit of the ribosomal RNA (September 2018). Phyla (or corresponding rank for the Eukaryotes) containing the characterised enzymes are depicted in colour (yellow for Eukaryotes, blue for Archaea and red for Bacteria). The graphs represent the families for which enzymes have been characterised (coloured cells) and the total number of families within the phylum according to the SILVA taxonomy (level D4 for taxonomic family). Bigger squares contain 10 × 10 units to allow easier visualisation. The number of genera with INMARE enzymes indicated for each coloured clade. Scale bar represents substitutions per site. Abbreviations in figure: OPU, Operational Phylogenetic Unit.

innovation through industrially relevant research and development can be fast-tracked. The project has achieved more than it aimed to in terms of developing innovative screening and processing methodologies, the use of which can and will drastically shorten the laborious industrial enzyme screening and optimisation steps. The INMARE success is clearly illustrated by the fact that in less than 3 years, the INMARE partnership has catalogued and characterised the most comprehensive collection of marine genomic and metagenomic enzymes worldwide, some of which currently outperform the best commercial prototypes. INMARE demonstrates the importance and effectiveness of the European collaboration and the role of the European Union's Horizon 2020 Research and Innovation Programme in driving innovation.

ACKNOWLEDGEMENTS

The authors acknowledge the members of the INMARE Consortium (http://www.inmare-h2020.eu/).

FUNDING

This project received funding from the European Union's Horizon 2020 research and innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant agreement no. [634486] (project acronym INMARE). M.F. acknowledges also the Spanish Ministry of Science, Innovation and Universities (former Spanish Ministry of Economy, Industry and Competitiveness) [grant numbers BIO2014–54494-R and BIO2017–85522-R] and CCCRC acknowledges FCT ['FCT Investigator 2013' programme; Portugal].

Conflict of interest. None declared.

REFERENCES

Braakman R, Smith E. Metabolic evolution of a deep-branching hyperthermophilic chemoautotrophic bacterium. *PLoS One* 2014;**9**:e87950.

Costessi A, van den Bogert B, May A *et al*. Novel sequencing technologies to support industrial biotechnology. *FEMS Microbiol Lett* 2018;**365**:1–13.

Evans J. Environmental markets for biotechnology. *Report* IO019C. BCC 2013.

Glöckner FO, Stal LJ, Sandaa R-A *et al*. Marine microbial diversity and its role in ecosystem functioning and environmental change. In: Calewaert JB, McDonough N (eds). *Marine Board Position Paper 17*. Ostend, Belgium: Marine Board-ESF, 2012.

Harrison JP, Gheeraert N, Tsigelnitskiy D *et al*. The limits for life under multiple extremes. *Trends Microbiol* 2013;**21**:204–12.

Mora C, Tittensor DP, Adl S *et al*. How many species are there on Earth and in the ocean? *PLoS Biol* 2011;**9**:e1001127.

Munoz R, Yarza P, Ludwig W *et al*. Release LTPs104 of the all-species living tree. *Syst Appl Microbiol* 2011;**34**:169–70.

Martínez-Martínez M, Bargiela R, Ferrer M. Chapter 7 – metagenomics and the search for industrial enzymes. In: Goutam B. (ed.). *Biotechnology of Microbial Enzymes: Production, Biocatalysis and Industrial Applications*. Chennai, India: Academic Press, 2017, 167–84.

Martínez-Martínez M, Coscolín C, Santiago G *et al*. Determinants and prediction of esterase substrate promiscuity patterns. *ACS Chem Biol* 2018;**13**:225–34.

Popovic A, Hai T, Tchigvintsev A *et al*. Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. *Sci Rep* 2017;**7**:44103.

Salazar G, Sunagawa S. Marine microbial diversity. *Curr Biol* 2017;**27**:R489–94.

Sherkow JS. Patent protection for microbial technologies. *FEMS Microbiol Lett* 2017;**364**:20.

Sturmberger L, Wallace PW, Glieder A *et al*. Synergism of proteomics and mRNA sequencing for enzyme discovery. *J Biotechnol* 2016;**235**:132–8.

Yarza P, Ludwig W, Euzéby J *et al*. Update of the all-species living tree project based on 16S and 23S rRNA sequence analyses. *Syst Appl Microbiol* 2010;**33**:291–9.

Yarza P, Yilmaz P, Pruesse E *et al*. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Micro* 2014;**12**:635–45.