Genomics Proteomics Bioinformatics 13 (2015) 290–295

HOSTED BY

ELSEVIER

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

REVIEW

Marine Metagenome as A Resource for Novel Enzymes

Amani D. Alma'abadi [a], Takashi Gojobori [*,b], Katsuhiko Mineta [c]

Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

Received 2 October 2015; revised 19 October 2015; accepted 19 October 2015
Available online 10 November 2015

Handled by Fangqing Zhao

KEYWORDS

Microbial diversity;
Culture-independent studies;
Catalysis;
Lipase;
Biotechnology

Abstract More than 99% of identified prokaryotes, including many from the marine environment, cannot be cultured in the laboratory. This lack of capability restricts our knowledge of microbial genetics and community ecology. Metagenomics, the culture-independent cloning of environmental DNAs that are isolated directly from an environmental sample, has already provided a wealth of information about the uncultured microbial world. It has also facilitated the discovery of novel biocatalysts by allowing researchers to probe directly into a huge diversity of enzymes within natural microbial communities. Recent advances in these studies have led to a great interest in recruiting microbial enzymes for the development of environmentally-friendly industry. Although the metagenomics approach has many limitations, it is expected to provide not only scientific insights but also economic benefits, especially in industry. This review highlights the importance of metagenomics in mining microbial **lipases**, as an example, by using high-throughput techniques. In addition, we discuss challenges in the metagenomics as an important part of bioinformatics analysis in big data.

Introduction

Recent developments in catalysis have led to a renewed interest in the use of enzymes for the environmentally-friendly industry. Most industrially-relevant enzymes are of microbial origin [1]. Identification and isolation of microbial enzymes are thus important steps in improving industrial processes, although only less than 1% of environmental bacteria can be cultivated in the laboratory [2–4].

The current challenging questions have arisen regarding the discovery, identification, and function validation of the uncultured microorganisms. Metagenomics study, which usually starts from the isolation of environmental DNAs without culture, has emerged as an excellent means to study biodiversity and biotechnological applications in certain conditions such as marine environments (**Figure 1**) [5–7]. It provides insights into the genomic pool of microorganisms that are recovered directly from environmental sources. Thus, metagenomics can be used for not only exploring ecological and environmental puzzles, but also finding unique biocatalysts with promising

* Corresponding author.
E-mail: takashi.gojobori@kaust.edu.sa (Gojobori T).
[a] ORCID: 0000-0001-8895-5423.
[b] ORCID: 0000-0001-7850-1743.
[c] ORCID: 0000-0002-4727-045X.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2015.10.001

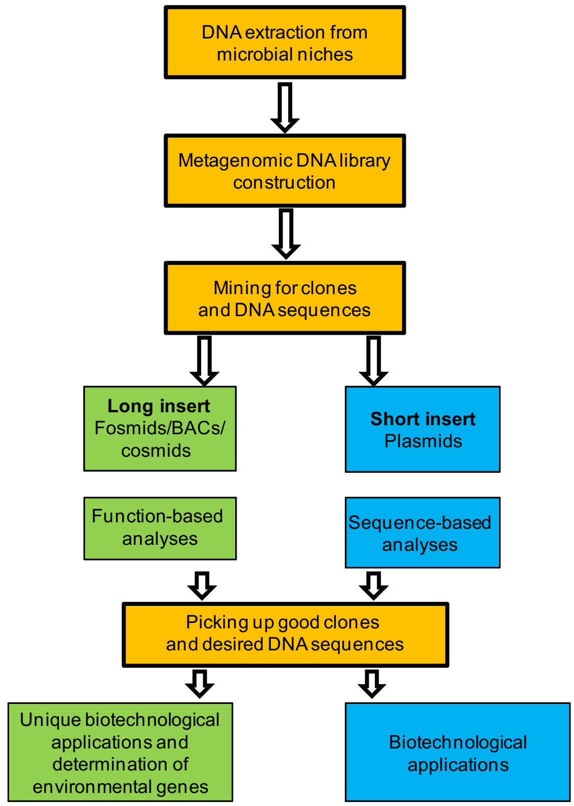

Figure 1 The process of functional metagenomics of marine microbes from environmental samples
This flowchart illustrates how metagenome is analyzed with the emphasis on the four important processes. BAC, bacterial artificial chromosome.

characteristics for biotechnological applications [8–10]. In particular of the biotechnological applications, metagenome libraries could be screened based on either protein function or nucleotide sequences.

Function-based screening is a direct way of identifying novel enzymes [2]. In this method, enzyme activities are assayed by harvesting a metagenomic library on agar plates enriched with substrates. Positive clones may then be recognized by visual screening for a clear zone called a halo [11]. As a result, function-based screening selects clones with functional activities, such as the synthetic and degradation activities. Unlike sequence-based approaches mentioned later, functional-based screening does not require identification of homologies to genes of known functions. It therefore contributes to nucleic acid and protein databases by adding novel functional annotations. However, this method often suffers from a number of limitations, such as a low hit rate of positive clones, low throughput, and time-consuming screening [11].

In contrast, in sequence-based screening, which involves metagenomic DNA sequencing using next-generation sequencing (NGS) technology, microbial enzymes and bioactive compounds can be explored from niches of interest [10]. However, sequence-based screening requires the detection of gene variants with conserved domain or motif of the known functions for enzymes identifications. This approach does not necessarily identify the novel genes.

In light of an increasing demand for enzymes such as carbohydrases, proteases, polymerases, nucleases, and lipases, it is becoming extremely difficult to ignore the importance of hydrolytic enzymes as potential biocatalysts in a wide variety of industries, including chemical processing, dairy, agrochemicals, paper, cosmetics, pharmaceuticals, surfactants, detergents, polymers, and biofuel synthesis [12,13]. For example, a lipase is often used at the consumable detergent, as it can hydrolyze fat from clothes and thus enhance its cleaning efficiency. Therefore, the hydrolytic enzymes have been used as promising environmentally-friendly biocatalysts in various industries.

According to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, enzymes are classified into six main classes (**Table 1**). One of the most important classes is hydrolases (E.C.3.-.-.-), which catalyze the hydrolytic cleavage of different types of chemical bonds. Many commercially-critical enzymes belong to this class, *e.g.*, proteases, amylases, acylases, lipases, and esterases [14]. Lipases are simply hydrolytic enzymes that catalyze hydrolysis and synthesis reactions by breaking down triacylglycerides into free fatty acids and glycerols, which act under aqueous conditions on the carboxyl ester bonds present in triacylglycerols to liberate fatty acids and glycerol [15–17]. Hydrolysis of glycerol esters carrying an acyl chain, which comprises less than 10

Table 1 Lipase and enzyme classification according to EC number

EC number	Enzyme
EC 1.-.-.-	Oxidoreductases
EC 2.-.-.-	Transferases
EC 3.-.-.-	Hydrolases
EC 3.1.-.-	Acting on ester bonds
EC 3.1.1.-	Carboxylic ester hydrolases
EC 3.1.1.3	**Triacylglycerol lipase (= lipase, in general)**
EC 3.2.-.-	Glycosylases
EC 3.3.-.-	Acting on ether bonds
EC 3.4.-.-	Acting on peptide bonds (peptide hydrolases)
EC 3.5.-.-	Acting on carbon–nitrogen bonds, other than peptide bonds
EC 3.6.-.-	Acting on acid anhydrides
EC 3.7.-.-	Acting on carbon–carbon bonds
EC 3.8.-.-	Acting on halide bonds
EC 3.9.-.-	Acting on phosphorus–nitrogen bonds
EC 3.10.-.-	Acting on sulfur–nitrogen bonds
EC 3.11.-.-	Acting on carbon–phosphorus bonds
EC 3.12.-.-	Acting on sulfur–sulfur bonds
EC 3.13.-.-	Acting on carbon–sulfur bonds
EC 4.-.-.-	Lyases
EC 5.-.-.-	Isomerases
EC 6.-.-.-	Ligases

Note: EC numbers and their descriptions are adapted from the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology. Lipase is highlighted in bold to show its position among the EC classification.