Biotechnology Advances 30 (2012) 920–929

Contents lists available at SciVerse ScienceDirect

Biotechnology Advances

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

Research review paper

# Application of metagenomic techniques in mining enzymes from microbial communities for biofuel synthesis

Mei-Ning Xing [a,1], Xue-Zhu Zhang [b,1], He Huang [a,*]

[a] Department of Biological Engineering and Key Laboratory of Systems Bioengineering of Ministry of Education, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China
[b] The First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China

ARTICLE INFO

Available online 28 January 2012

*Keywords:*
Metagenomic sequencing
Biofuel synthesis
Critical enzymes
Genetic modification
Biomass

ABSTRACT

Feedstock for biofuel synthesis is transitioning to lignocelluosic biomass to address criticism over competition between first generation biofuels and food production. As microbial catalysis is increasingly applied for the conversion of biomass to biofuels, increased import has been placed on the development of novel enzymes. With revolutionary advances in sequencer technology and metagenomic sequencing, mining enzymes from microbial communities for biofuel synthesis is becoming more and more practical. The present article highlights the latest research progress on the special characteristics of metagenomic sequencing, which has been a powerful tool for new enzyme discovery and gene functional analysis in the biomass energy field. Critical enzymes recently developed for the pretreatment and conversion of lignocellulosic materials are evaluated with respect to their activity and stability, with additional explorations into xylanase, laccase, amylase, chitinase, and lipolytic biocatalysts for other biomass feedstocks.



**Contents**

1. Introduction . . . 921
2. Access to novel biocatalysts from the metagenome . . . 921
   2.1. Pretreatment of environmental samples . . . 921
   2.2. Extraction of nucleic acid . . . 921
   2.3. Construction of metagenomic library . . . 922
      2.3.1. Vector selection . . . 922
      2.3.2. Host selection . . . 922
   2.4. Metagenomic library screening . . . 922
      2.4.1. Biological activity screening . . . 923
      2.4.2. DNA sequence screening . . . 923
      2.4.3. Substrate-induced gene expression screening (SIGEX) . . . 923
      2.4.4. Compound configuration screening . . . 923
3. Research progress of metagenomic sequencing . . . 923
   3.1. Traditional sequencing and second-generation sequencing . . . 923
   3.2. Second-generation DNA sequencing . . . 923
      3.2.1. 454 pyrosequencing . . . 923
      3.2.2. Illumina Genome Analyzer . . . 924
      3.2.3. AB SOLiD . . . 924
4. Mining enzymes for biofuel production from metagenomes . . . 924
   4.1. Significance of metagenome-derived novel enzymes in biofuel production . . . 924
   4.2. Lignocellulose degrading enzymes for biofuel production . . . 925
      4.2.1. Lignin degrading enzymes . . . 925
      4.2.2. Xylanase . . . 925

* Corresponding author at: Department of Bioengineering, school of Biochemical Engineering, Tianjin University, Tianjin 300072, China. Fax: +86 22 2740 9598.
*E-mail address:* huang@tju.edu.cn (H. Huang).
[1] Mei-Ning Xing and Xue-Zhu Zhang are equal contributors.

0734-9750/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.biotechadv.2012.01.021

4.2.3. Cellulase . . . 926
4.2.4. Multifunctional enzymes . . . 926
4.3. Lipolytic enzymes for biodiesel production . . . 926
4.4. Other enzymes involved in biofuel production . . . 927
5. Conclusion . . . 927
Acknowledgements . . . 927
References . . . 927

## 1. Introduction

Derived from biomass, biofuels offer an environmentally benign and cost-effective solution for fossil fuel depletion. Of this alternative, renewable sources of energy, biodiesel and bioethanol have attracted growing attention from policy makers, industry and researchers for their economic, environmental and social benefits. Significant incentive exists for the development of efficient biofuel technology, with both the U.S. and members of the EU committing to increasing the proportion of renewable energy in their primary energy supply to 10% and 20% by 2010 and 2020, respectively. Other governments, such as Sweden's, have adopted even more ambitious targets, attempting to replace all fossil fuels with biofuels after 2020 to eliminate their dependence on oil.

Despite recent growth in global production of biofuels, significant technological bottlenecks still exist in the production processes to efficiently convert biomass into biofuels. While thermochemical conversion technologies can be applied for biomass conversion, industry is increasingly considering enzymes as a key technology for biofuels development and utilization, citing their efficiency and selectivity in the reaction chemistry (Jaeger et al., 1999). However, the enzymes currently employed for biomass conversion cannot meet the growing demand for economically viable biofuels due to their high cost, low activity and poor stability under the required operating conditions. Thus, continued development of novel enzymes for use in the production of advanced biofuels is required (Barnard et al., 2010).

Acceleration of the novel enzyme development process is primarily dependent on two factors: (1) efficiency and sensitivity of the screening strategy, and (2) diversity of candidate genes (microbial). The fact that traditionally enzymes could only be obtained from bacterial isolates was one of the main limitations to the widespread application of enzymes in industry (Leresche and Meyer, 2006). More than 99% of microorganisms from natural environments cannot be efficiently cultivated using current isolation and culture methods, severely reducing the microbial resources which can be utilized (Torsvik and Ovreas, 2002). Scientists have thus focused on the development of new methods capable of utilizing the genes of these microorganisms in biotechnology which are independent of routine culture techniques.

Metagenomics is an advanced methodology which emerged in the late 1990s, by means of extracting all microbial genomic DNAs in a certain environmental habitat, constructing metagenomic libraries, and screening to seek novel functional genes and/or biologically active compounds (Ferrer et al., 2005; Wang et al., 2009). Metagenomics overcomes the disadvantages of isolation and cultivation procedures of the traditional microbial method, and thus greatly broadens the space of microbial resource utilization. It has become one of the powerful research tools for microbiology, biotechnology, soil and environmental sciences, and a new field of genetic engineering.

At present, with the help of the rapid development of high-throughput sequencing methods, metagenomics has been employed to identify enzymes for use in biofuels production. Many novel enzymes have been found by means of this technology, including lignases, xylanase, endoglucanase, amylolytic enzymes, β-glucosidase for bioethanol, and lipolytic enzymes for biodiesel. Some of these have multiple functions and can catalyze a number of different reactions (Kim et al., 2008; Nam et al., 2009, 2010; Palackal et al., 2007; Zhao et al., 2010). Some exhibit high activities (Fang et al., 2009), specificities (Wong et al., 2010) and stability (Pottkamper et al., 2009), and can work under a wide range of pH (Duan et al., 2009), temperature (Sharma et al., 2010) or ionic conditions (Ilmberger and Streit, 2010). These enzymes may have potential for new application in biofuels production.

## 2. Access to novel biocatalysts from the metagenome

Research strategies for accessing novel biocatalysts from the metagenome include: pretreatment of genes of interest, extraction of nucleic acid, selection of vector and host system, and metagenomic library screening (Fig. 1). Multidisciplinary developments in the areas of microbiology, molecular biology and bioinformatics have enabled metagenomic technologies within each of these stages, contributing significantly to the development of novel biocatalysts.

### 2.1. Pretreatment of environmental samples

Pretreatment for nucleic acid extraction processing operations includes non-enrichment or enrichment of interested genes. In most experimental research, non-enriched methods are used due to their improved ability to maintain the diversity of microbial communities. However, enrichment methods are known to improve the specificity of a sample's genomic DNA, benefiting sequencing-based screening of novel genes. Stable-isotope probing (SIP) (Radajewski et al., 2002), suppression subtractive hybridization (SSH) (Galbraith et al., 2004), differential display (Liang, 2002), phage-display (Crameri and Suter, 1993), affinity capture (Demidov et al., 2000), and microarrays (Wu et al., 2001) are all methods of enrichment.

### 2.2. Extraction of nucleic acid

Construction of a metagenomic library requires a sufficient number of high quality DNA samples, making the extraction and purification of DNA from the environmental samples a critical step (Wilkinson et al., 2002). Two types of extraction methods are commonly applied, according to the size of target genes and different screening strategies: direct extraction and indirect extraction. Direct extraction methods use detergents and enzymes to process the test samples without the cultivation of microorganisms, followed by phenol or chloroform-based extraction and separation of the DNA. Although this method has a greater DNA recovery rate, the smaller extracted DNA fragments (general 1–50 kb) and elevated impurity content due to destructive mechanical forces makes this method inappropriate for constructing large inserts libraries (Desai and Madamwar, 2007). Nonetheless, direct extraction methods have been successfully used to extract DNA from microbial communities (Bey et al., 2010). Indirect extraction methods (cell separation and extraction method) employ physical means to separate the microorganisms from the sample followed by lysis extraction, thus obtaining larger DNA fragments by avoiding high mechanical strength actions directly on the DNA. The recovery rate of indirect extraction is 10–100 times lower than direct extraction (Parachin et al., 2010). Thus, in specific experiments, the extraction method should be selected by weighing the various requirements for product recovery, including: operational