Biotechnology Advances xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Biotechnology Advances

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

Research review paper

Metagenomics for the discovery of pollutant degrading enzymes

Lisa Ufarté, Élisabeth Laville, Sophie Duquesne, Gabrielle Potocki-Veronese *

Université de Toulouse, INSA, UPS, INP, LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France
INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France
CNRS, UMR5504, F-31400 Toulouse, France

ARTICLE INFO

Article history:
Received 17 July 2015
Received in revised form 20 October 2015
Accepted 22 October 2015
Available online xxxx

Keywords:
Pollutants
Metagenomics
Enzyme discovery
Screening

ABSTRACT

Organic pollutants, including xenobiotics, are often persistent and toxic organic compounds resulting from human activities and released in large amounts into terrestrial, fluvial and marine environments. However, some microbial species which are naturally exposed to these compounds in their own habitat are capable of degrading a large range of pollutants, especially poly-aromatic, halogenated and polyester molecules. These microbes constitute a huge reservoir of enzymes for the diagnosis of pollution and for bioremediation. Most are found in highly complex ecosystems like soils, activated sludge, compost or polluted water, and more than 99% have never been cultured. Meta-omic approaches are thus well suited to retrieve biocatalysts from these environmental samples. In this review, we report the latest advances in functional metagenomics aimed at the discovery of enzymes capable of acting on different kinds of polluting molecules.

Contents

1. Introduction

Large amounts of pollutants are released into the air, water and soil as a result of industrial, agricultural and domestic activities. Organic pollutants are synthetic compounds in the form of herbicides, dyes, pesticides, plastics and drugs (Rieger et al., 2002). Most are aromatic molecules, polymers of ring shaped molecules or planar molecules, and are therefore among the most stable and persistent molecules. Those that most frequently accumulate in natural environments are polycyclic aromatic hydrocarbons (PAH) especially chlorinated hydrocarbons, steroids (phenols, phthalates) and dyes, but also organocyanides including nitriles, long chain aliphatics found in numerous plastics and insulating materials like polyurethanes, or organophosphates and pyrethroid herbicides and pesticides.

Abbreviations: ABTS, 2,2′-Azino-bis(3-ethylbenzothiazoline-6-sulfonic acid); AMD, acid mine drainage; DBP, dibutyl phthalate; DGGE, denaturing gradient gel electrophoresis; GC–MS, gas chromatography–mass spectrometry; HC, hydrocarbon; HPLC, high performance liquid chromatography; PAH, polycyclic aromatic hydrocarbons; PBS, poly-butylene succinate; PBSA, poly-butylene succinate-co-adipate; PCB, polychlorinated biphenyl; PCL, polycaprolactone; PCR, polymerase chain reaction; PES, poly-ethylene succinate; PET, polyethylene terephthalate; PHB, polyhydroxybutyrate; PLA, polylactic acid; POP, persistent organic pollutant; PU, polyurethane; SIP, stable isotope probing; TCP, 3,5,6-trichloro-2-pyridinol.

* Corresponding author at: Equipe de Catalyse et Ingénierie Moléculaire Enzymatiques, Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés, INSA - UMR INRA 792 - UMR CNRS 5504, 135 avenue de Rangueil, 31077 Toulouse cedex 4, France.
E-mail addresses: ufarte@insa-toulouse.fr (L. Ufarté), laville@insa-toulouse.fr (É. Laville), duquesne@insa-toulouse.fr (S. Duquesne), veronese@insa-toulouse.fr (G. Potocki-Veronese).

http://dx.doi.org/10.1016/j.biotechadv.2015.10.009
0734-9750/

The toxicity of these compounds for the environment and for mammalian organisms results from their resistance to natural degradation, which is linked to their stable structures. They are grouped under the name of persistent organic pollutants (POP), some of them being potent carcinogens or mutagens, and/or having endocrine disrupting properties (Ballschmiter et al., 2002).

Cleaning up contaminated soils, wastewater, groundwater and marine environments is a challenging task. Physical and chemical technologies (electrochemical treatments, oxidising agents, activation by ultraviolet rays, adsorption of pollutants, membrane filtration, ion exchange, electrokinetic coagulation, etc.) can be combined to reduce contamination to a safe and acceptable level. Many of these processes have drawbacks, such as formation of by-products, high sludge production, and high processing costs (Riser-Roberts, 1998; Robinson et al., 2001). In this context, recent decades have seen the development of biological processes based on the breakdown of pollutant organic compounds into stable, innocuous end-products by bacteria (Watanabe, 2001; Yam et al., 2010), fungi (Cerniglia, 1997; Bertrand et al., 2015), algae or their enzymes (Sutherland et al., 2004). Since bioremediation may result in the complete metabolisation of pollutants, it is considered to be a highly effective and environmentally friendly strategy (Colleran, 1997). Microorganisms have indeed developed a wide range of aerobic and anaerobic catabolic strategies to degrade the huge range of organic compounds present in the ecosystems they colonise. Because pollutant molecules are often structurally similar to natural molecules, one can assume that there are always some organisms in contaminated ecosystems that are able to metabolise pollutants, which serve as their main carbon source.

During microbial degradation, all changes in the chemical structure of pollutants are due to the action of enzymes, whose specificity is often broad enough to accommodate several molecules of similar structures. Once identified and isolated, these enzymes can therefore be engineered by directed evolution to improve their stability or efficiency with respect to a particular compound (Festa et al., 2008; Theerachat et al., 2012). However, the greatest advances in recent decades have been made thanks to the development of next generation sequencing and other 'omics' technologies. These new technologies have enabled numerous insights into the genes and metabolic pathways involved in pollutant catabolism by bacteria and fungi, thereby improving our understanding of the genetic and molecular bases of the reactions involved in the degradation processes. Most knowledge has been obtained by functional genomics of pollutant-degrading microorganisms or model microbial communities. This allowed one, after genome sequencing and/or activity-based screening of genomic libraries, to identify genes encoding the enzymes that are directly involved in the breakdown of the targeted compound (Pieper et al., 2004). However, since more than 99% of the microbial species of the terrestrial (Pham and Kim, 2012) and aqueous (Zengler et al., 2002) ecosystems are uncultured, one can consider that most of the microorganisms that are capable of breaking down pollutants remain uncharacterised.

Functional metagenomics, which consists in assigning functions to proteins encoded by all genomes of a microbial community with no isolation and cultivation step, is a highly efficient way to boost the discovery of novel biocatalysts from the huge diversity of uncultured microbes. In this review, we present the various functional metagenomic approaches that require, or not, cloning of metagenomic fragments in an expression host in order to screen and experimentally validate gene functions, which have been used in recent years to identify enzymes and metabolic pathways involved in pollutant degradation.

2. Targeted metagenome sampling

Numerous microbial communities have the potential to degrade and metabolise most pollutants. This is even the case of those in natural, non-polluted ecosystems, as long as (i) their taxonomical and functional diversity is sufficient, and (ii) the structure of the targeted pollutant is similar to that of one of the natural substrates used as carbon source to support microbial growth. Therefore, several such natural microbiota that had never or only rarely been exposed to pollutants, were successfully mined for xenobiotic degrading enzymes. For example, as detailed later in subsections 3.2.2.1 and 3.2.2.2, organophosphate and 3,5,6-trichloro-2-pyridinol (TCP) degrading enzymes (Math et al., 2010), chlorpyrifos insecticides degrading esterases (Kambiranda et al., 2009) were discovered in the cow rumen microbiome. Likewise, in the human gut metagenomic gene catalogue, 15% of genes have been annotated as being involved in the degradation and metabolism of xenobiotics, like halogenated aromatic compounds (Qin et al., 2010).

Nevertheless, since in polluted environments the microbial communities are enriched in microorganisms able to thrive by degrading pollutant compounds, most metagenomic studies have been conducted on polluted soils, activated sludge, sediment, or wastewater environments. To reduce the complexity of the sample, stable isotope probing (SIP) can be used to specifically label the genomes of pollutant metabolising species. For instance, Wang et al. (2012) amended microcosms from contaminated groundwater with ^{13}C-labelled naphthalene. The ^{13}C-labelled DNA was then easily separated from the unlabelled DNA, to be specifically sequenced and concurrently cloned before activity-based screening. This approach is suitable for focusing on active degraders, especially those which act on relatively simple pollutant structures. Indeed, when more complex compound structures are targeted, polyaromatics for instance, there is a risk of missing the first steps of degradation undertaken by microbes acting on the most recalcitrant molecules, with the simpler reaction products being metabolised and incorporated in the DNA of the end-user microbes.

3. Screening for pollutant degrading enzymes

Two complementary approaches can be used to identify pollutant degrading biocatalysts in microbial communities. One is guided by gene sequence analysis and functional annotation, based on the content of available sequence databases (Cardenas and Tiedje, 2008). The other is guided by the observation of pollutant degrading phenotypes, harboured by recombinant metagenomic clones (Fig. 1). In both cases, only the cloning of the targeted metagenomic sequences allows the gene function to be validated experimentally.

3.1. Sequence-based approaches to understand pollutant degradation mechanisms by microbial communities

Sequence-based metagenomics relies on whole-genome DNA extraction from microbial communities, shotgun sequencing, and read assembly (Fig. 1). The assembled sequences are subjected to bioinformatic analysis to assess the composition and functions of the microbial populations. Over the past 10 years many sequencing projects have targeted heavily polluted ecosystems to study the structure of microbial populations (Chouari et al., 2003; Guermazi et al., 2008), species interactions, metabolic pathways, and the genes involved in species survival in such environments (Yamada et al., 2012). Among these projects, we cite the culture-independent recovery of nearly complete genomes of five dominant members of an acid mine drainage (AMD), by in-depth metagenome sequencing (Tyson et al., 2004). Functional annotation has provided new insights into the ecological roles of individual members of this community, with ferrous iron oxidation playing a key role in the microbial and geochemical processes in AMD ecosystems. More recently, the metagenome of the activated biomass from a common effluent treatment plant was sequenced to retrieve novel oxygenase sequences (Jadeja et al., 2014). The sequences were compared to enzymes identified in previous metagenomes from different wastewater compositions and at different locations, including a sewage effluent treatment plant, an enhanced biological phosphorus treatment plant, and a tannery waste treatment plant (Albertsen et al., 2012; Yu and Zhang, 2012; Wang et al., 2013). In addition, samples of the activated

Please cite this article as: Ufarté, L., et al., Metagenomics for the discovery of pollutant degrading enzymes, Biotechnol Adv (2015), http://dx.doi.org/10.1016/j.biotechadv.2015.10.009