Biotechnology and Genetic Engineering Reviews - Vol. 24, 107-116 (2007)

The SIGEX Scheme: High Throughput Screening of Environmental Metagenomes for the Isolation of Novel Catabolic Genes

TAKU UCHIYAMA[1] AND KAZUYA WATANABE[2]*

[1]Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba, Ibaraki, Japan and [2]Laboratory of Applied Microbiology, Marine Biotechnology Institute, 3-75-1, Heita, Kamaishi, Iwate, Japan

Introduction

Microorganisms represent the largest reservoir of genetic diversity on our planet. They occupy every part of the earth and manage its geochemistry, cycling of elements and breakdown of wastes. The prestige of microorganisms is ascribable to their diverse metabolic capacities, which also infers great potential for their application to biotechnological purposes, e.g., production of pharmaceuticals, enzymes, polymers and microbial agents. Nowadays, many institutes and companies have established the collection and harvesting of microorganism cultures isolated from a variety of environments. These microorganisms have been used for screening for their abilities to produce medicinally relevant metabolites and industrially useful enzymes. However, we have at present also known that only a small fraction of natural microorganisms can be recovered by our cultivation techniques. To cite an instance, scientists have estimated that cultivable microorganisms form less than 1% of natural microorganisms (Amann *et al.,* 1995; Cowan, 2000), implying that the great genetic diversity in the environment has as yet been unexplored.

*To whom correspondence may be addressed (kazuya.watanabe@mbio.jp)

Abbreviations used: SIGEX, substrate-induced gene expression screening; PCR, polymerase chain reaction; IAN-PCR, inverse affinity nested PCR; FACS, fluorescence active cell sorting; GFP, green fluorescence protein; IPTG, Isopropyl ß-D-Thiogalactoside; LB, Luria-Bertani; dLB, diluted Luria-Bertani; ORF, open reading frame; RFLP, restriction fragment length polymorphism; BLAST, Basic Local Alignment Search Tool.

Screening of environmental 'metagenomes' (a mixture of genomes of multiple organisms, particularly mixed microbial genomes extracted from an environmental sample, is referred to as metagenome) without isolating individual microorganisms has recently been recognized as an attractive approach for obtaining novel biocatalysts from natural genetic resources, since this method facilitates access to the large genetic diversity of uncultured microorganisms in the environment (Handelsman *et al.,* 1998; Lorenz *et al.,* 2002; Ferrer *et al.,* 2005). Current metagenome approaches employ either of two screening methods for the isolation of catabolic genes, namely, enzyme activity-based screening and nucleotide sequence-based screening. In terms of their intrinsic limitations (as described below), however, these methods can still be considered to access only small portions of the natural genetic diversity.

Considering the limitations associated with the current screening methods, we have recently developed a third option for the screening of environmental metagenome libraries for isolating novel catabolic genes, *i.e.*, substrate-induced gene expression screening (SIGEX) (Uchiyama *et al.,* 2005). In this review, we outline the SIGEX method in order to illustrate how SIGEX can contribute to the metagenomic approaches.

Traditional metagenome screening methods and SIGEX

Two major screening approaches have been utilized for producing novel biocatalysts from environmental metagenomes; one approach is *phenotype-based screening*, while the other is *genotype-based screening*. In the first approach, a shotgun library is constructed from an environmental metagenome and used for screening for clones that acquire a new phenotype, paricularly with regard to enzymatic activity and production of chemical compounds (Henne *et al.,* 1999; Henne *et al.*, 2000; Rondon *et al.,* 2000). A merit of this approach is that a positive clone will certainly contain whole genes necessary for expressing the acquired phenotype. It is however generally difficult to express an enzyme in a heterogeneous host organism as an active form, particularly when it requires supporting substances for its activity, for example a prosthetic group such as heme and an electron transport mechanism. This is one reason why easy-to-express enzymes such as lipases (Brady and Clardy, 2003) and esterases (Henne *et al.*, 2000) have thus far been those to have been produced from metagenome libraries. It is also worth pointing out at this juncture that in many cases it is very laborious to subject large metagenomic libraries to phenotype screening. For example, screenings of over 900,000 clones (total insert size was estimated to be 6 Gbp) were needed to obtain just five positive clones of 4-hydroxybutyrate-degrading enzymes (Henne *et al.,* 1999) and four positive clones of lipolytic enzymes (Henne *et al.*, 2000). Development of a high throughput screening strategy would therefore seem essential for facilitating phenotype screening of environmental metagenome libraries.

In the second approach, nucleotide sequences of genes coding for known enzymes are compared to find out conserved regions whose sequences can then be used for designing PCR primers or hybridization probes (Okuta, *et al.,* 1998; Eschenfeldt *et al.,* 2001; Bell *et al.,* 2002; Kubota, *et al.,* 2005). These primers are probes are used to directly amplify gene fragments from an environmental metagenome and to select clones from a metagenomic library, respectively. These schemes are highly selective

owing to the high variety of nucleotide sequences and are reliable in terms of their employment of well developed genetic techniques. Additionally, PCR is considered to be highly efficient, amplifying a small number of target genes to a sufficient quantity.

However, there exist two major drawbacks in these approaches. First, the design of primers or probes is dependent on known sequences of enzymes and proteins that have mostly been purified and/or cloned from easily cultured laboratory isolates of microorganisms. It is therefore likely that the designed primers and probes are biased toward genes that are similar to known genes and eliminate less homologous genes encoding novel enzymes. The second major drawback is that PCR recovers only partial fragments of target genes, requiring additional steps to recover whole genes necessary for expressing their activities.

Different from these two approaches, SIGEX is a method to screen an environmental metagenome library for obtaining catabolic genes whose expression is induced in response to environmental stimuli, such as the occurrence of a chemical compound (Uchiyama *et al.,* 2005). This is based on knowledge that catabolic gene expression is generally induced by relevant compounds (substrates or metabolites) and, in many cases, controlled by regulatory elements situated proximate to catabolic genes. In order for SIGEX to be high throughput, we have constructed an operon-trap *gfp* vector (designated p18GFP) suitable for shotgun cloning and used in combination with fluorescence activated cell sorting (FACS) for the selection of positive clones in liquid cultures. In this way, FACS selects for cells that harbor stimulus-responsive metagenome fragments in front of the *gfp* gene and become fluorescent in response to the stimulus. Similar approaches have been used for identifying genes in a genome of pathogenic bacteria, which were expressed in human and plant bodies (Valdivia and Falkow, 1997; Dunn *et al.,* 2003). However, quite distinct from those approaches, SIGEX is optimized for the metagenomic screening by virtue of several improvements as we will now review.

The total procedure of the SIGEX scheme is presented in *Figure 1*. This scheme is comprised of the following four steps.

(I) Construction of a metagenomic library in liquid culture. A metagenome is digested by a restriction enzyme, ligated in the operon-trap p18GFP vector and used for transformation of a host strain (e.g., *Escherichia coli*). The transformed cells are grown in a liquid culture and stored in glycerol stocks.

(II) Removal of clones containing self-ligation plasmids and those expressing GFP constitutively. The liquid-culture library contains cells containing self-ligation plasmids (without inserts) and those containing plasmids whose inserts have constitutive transcriptional activities. These cells are removed by FACS after cultivation and induction with isopropyl-beta-D-thiogalactopyranoside (IPTG). This procedure greatly improves the efficiency for selecting positive clones, since false-positive clones can almost completely be eliminated.

(III) Selection of clones expressing GFP in the presence of a target substrate. Cells selected in step II are grown in liquid cultures in the presence of a target substrate and subjected to FACS to select for fluorescent cells. For cultivation in this step, we recommend to use the dLB medium (1/10-strength LB medium), since higher concentrations of organics can inhibit the expression of catabolic genes.

(IV) Colony isolation of the sorted cells on agar plates. A sorted-cell fraction is spread on agar plates to isolate positive clones as colonies. Isolated colonies are then grown under the same conditions as step III and subjected to FACS in order to check if they really are positive. Positive clones are then used for sequence analyses and expression of enzymatic activities.

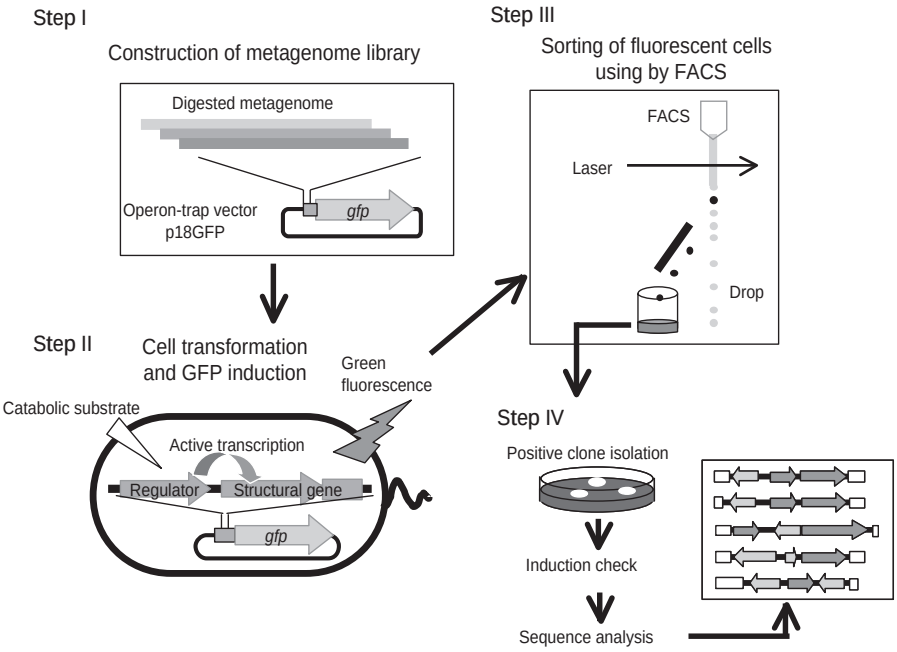


Figure 1. Schematic representation of the SIGEX scheme.

DNA fragments obtained by the SIGEX method are expected to contain a catabolic operon fragment for target compounds, i.e., ORFs coding for transcriptional regulator(s) and catabolic enzyme(s). When an operon thus obtained is partial, IAN-PCR (Uchiyama and Watanabe, 2006) is applicable to metagenome walking for obtaining a whole operon fragment.

Examples of the SIGEX screening

In a model experiment to demonstrate if SIGEX can be used for cloning of a catabolic-gene fragment, we had attempted to isolate a known phenol-degradative operon (the *pox* operon (Hino *et al.,* 1998)) from a phenol-degrading *Ralstonia eutropha* E2 genomic library. In this case, a *Ralstonia*-genomic library (not a metagenome library) constructed using p18GFP contained 26,000 clones with an average insertion size of 7 kb. The library was first grown in the LB medium supplemented with IPTG (step II) and subjected to FACS to recover fluorescence-negative clones (step III). The

clones in the recovered fraction were next grown in the dLB medium supplemented with phenol (step II), and fluorescence-positive clones were recovered by FACS (step III). Fluorescence-positive clones in step III were spread on agar plates containing LB plus ampicilin, and colonies obtained were grown in dLB with and without phenol (step IV). The two cultures from one colony were analyzed by FACS (but not fractionated) to know whether or not fluorescence expression was phenol dependent (step IV, induction check). Among positive clones (fluorescence was expressed only after cultivation in the presence of phenol), two clones were subjected to sequence analysis, showing that they contained a same *pox* fragment comprised of *poxRAB* (genes for a positive transcriptional regulator and two subunits of phenol hydroxylase). This example demonstrated the utility of SIGEX scheme for isolating a target catabolic-operon fragment from a genomic library (and possibly from a metagenome library).

Next we constructed an environmental metagenome library (containing approximately 152,000 clones with an average insertion size of 7 kb) from a crude-oil contaminated groundwater microbial flora (Uchiyama *et al*., 2005). Molecular characterization of this microbial flora using rRNA phylogeny had already been described elsewhere (Watanabe *et al.,* 2000). The library was subjected to the SIGEX screening for the purpose of isolating aromatic-hydrocarbon catabolic operon fragments. In this experiment, benzoate and naphthalene were used as the induction substrates. Benzoate was primarily used, since this compound is a key intermediate metabolite in aromatic-hydrocarbon catabolic pathways and acts as an inducer for the expression of upper (Marques and Ramos, 1993) and lower (Parsek *et al.,* 1992) operons. Screening by SIGEX with the two substrates was accomplished within 4 days and yielded 62 positive clones (58 for benzoate and 4 for naphthalene). The restriction fragment length polymorphism (RFLP) analysis was conducted to know redundancy among the positive clones, showing that benzoate-positive clones were divided into 33 types, whilst naphthalene-positive clones were divided into 2 types. The large number of RFLP type was obtained with benzoate, probably because it is a central intermediate metabolite in a variety of degradative pathways. Relatively small numbers of positive clones for naphthalene may be ascribable to the presence of less abundant microbes able to metabolize this substrate.

In order to examine if SIGEX-selected metagenome fragments actually contained relevant catabolic genes, we then analyzed nucleotide sequences of 10 RFLP types (8 benzoate-positive and 2 naphthalene-positive RFLP types, see *Figure 2*). It was found that one benzoate-positive clone (BZO026) contained ORFs homologous to genes in benzoate-degradative (Cowles *et al.,* 2000) and catechol-degradative (Rothmel *et al.,* 1990) operons (*Figure 3*). Other clones whose sequence analyses suggested potential relevance to aromatic-hydrocarbon transformation included BZO032 (decarboxylase), BZO062 (dehydrogenase), BZO071 (cytchrome P450), BZO135 (tyrosine lyase) and NAP3 (hydroxylase). Sequence analysis showed that putative catabolic genes were enriched in the SIGEX-selected metagenome fragments, while putative ORFs coding for transporter proteins were also included. Most putative ORFs found in the SIGEX-selected metagenome fragments showed low homology to nucleotide-database sequences (Uchiyama *et al*., 2005), suggesting that they may have novel catabolic activities.

Among ORFs found in the SIGEX-selected fragments, we were particularly interested in characterizing the putative P450 in BZO071. In the "BLAST" search of

the GenBank database, it showed the highest homology (albeit 29%) to monoterpene-transforming P450 of *Pseudomonas* (Peterson *et al.,* 1992), whilst the relevance to aromatic-hydrocarbon transformation was not apparent in the sequence analysis. *E. coli* harboring putative P450 gene was first subjected to the transformation assay with a variety of benzoate derivatives; however, in those experiments we could not detect any transformation product. In addition, the typical P450 absorption spectrum was not observed in its cell-free extract. We therefore cloned this putative P450 gene into an expression plasmid and attempted to express putative P450 protein under the control of the T7 promoter. The resultant cell-free extract exhibited typical absorption spectra of P450, confirming that it was really P450. Next, a spectrophotometric substrate-binding assay was performed with benzoate relatives, showing that the absorption peak of P450 was shifted from 416 nm to around 395 nm (due to changes in the spin state of the heme) in the presence of 4-hydroxybenzoate but not benzoate, 2- and 3-hydroxylated derivatives. In order to identify transformation products, the P450 was mixed with recombinant electron-transfer proteins such as putidaredoxin and putidaredoxin reductase of *P. putida* (Peterson *et al.,* 1990), and the mixture was used for the transformation assay. The electron transfer between these proteins was confirmed by spectrophotometric analysis. Among the compounds tested, gas chromatography-mass spectrometry analysis detected products only from 4-hydroxybenzoate and identified it as protocatechuate. To our knowledge, we considered this to be the first P450 that specifically catalyzes 3-hydroxylation of 4-hydroxybenzoate.

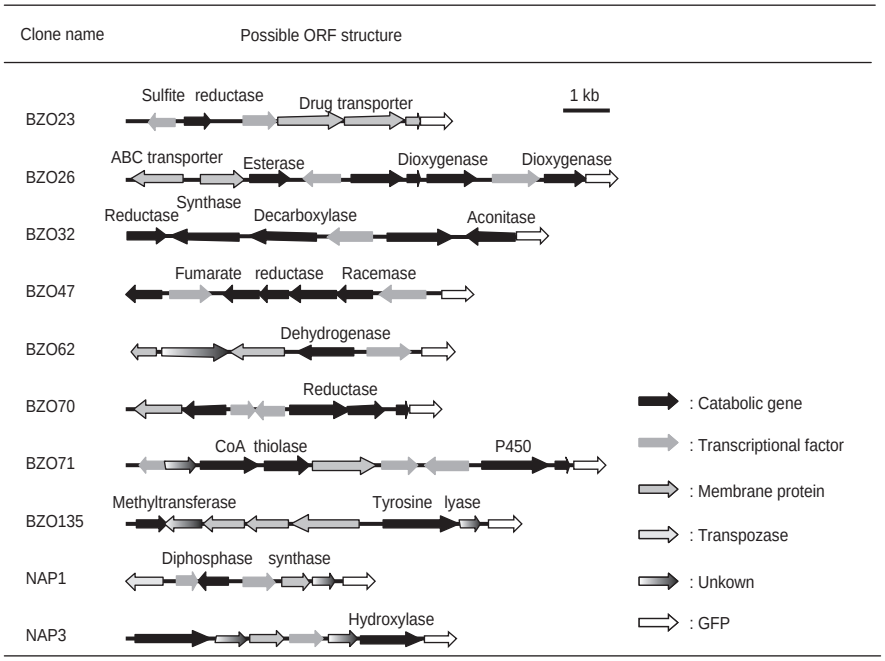
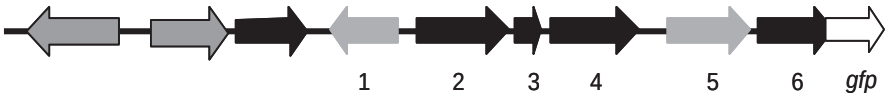


Figure 2. Sequence analysis of SIGEX-selected metagenome fragments. BZO and NAP fragments were obtained after induction with benzoate and naphthalene, respectively. Possible ORF structures and results of database search for putative ORFs are shown. For details, refer to a reference paper (Uchiyama *et al.,* 2005).



ORF	Similar proteins in databases	%AA identity	Organism
1	CatR; catBC operon regulator	71%	*P. putida*
2	CatB; muconate lactonizing enzyme I	74%	*P. putida*
3	CatC; muconolactone delta-isomerase	77%	*P. putida*
4	CatA; catechol 1,2-dioxygenase	65%	*P. putida*
5	BenR; benABC operon regulator	68%	*P. putida*
6	BenA; benzoate dioxigenase alpha subunit	85%	*P. putida*

Figure 3. Sequence analysis of BZO26 fragments.

The above example demonstrated the utility of the SIGEX scheme for isolating novel catabolic genes from an environmental metagenome library. As shown in the characterization of P450, SIGEX provides an opportunity for isolation and identification of a catabolic gene whose enzymatic activity is not easily expressed in the host cell. It was also shown that SIGEX facilitates prediction of substrates of newly obtained enzymes based on the induction substrate used in the SIGEX screening.

Merits and limitations of SIGEX

From the review of the experiments described above we can see how SIGEX appears to have two key advantages over other metagenome-screening methods that are currently available. Firstly, SIGEX can be considered to be high throughput with the aid of FACS, enabling cloning of many different catabolic genes in a relatively short time period. Secondly, as was the case of P450, SIGEX appears to enable us to access genes that are otherwise difficult to obtain.

Possible limitations of SIGEX include followings. First, the expression of catabolic genes in a SIGEX-selected genome fragment results from a cooperative work of a transcriptional regulator encoded in the SIGEX fragment and transcriptional factors in host cells, whilst their affinity may be low when the SIGEX fragment is derived from an organism distantly related to the host organism. This implies a possibility that SIGEX selection may bias itself toward genes in organisms closely related to the host. Second, regulatory genes and catabolic genes are not always proximate to each other; such catabolic genes are not obtained by SIGEX. Third, the expression of catabolic genes is not always induced by their substrates. And fourth, genes for catabolic enzymes obtained by SIGEX may sometimes be partial when they are situated at the end of a cloned genome fragment.

The first limitation can be overcome by preparing a repertoire of taxonomically distant host organisms, including *Streptomyces lividans* and *Bacillus satilis*, in addition to *E. coli* (Courtois *et al.,* 2003; Martinez *et al.,* 2004). Concerning the second

limitation, we consider that this is not a large obstacle in environmental metagenomic screening, since the great genetic diversity and rapid recombination of catabolic operons in the environment result in the evolution of catabolic operons that fit to the SIGEX criteria. To solve the third limitation, we have recommend using a series of related compounds as induction substrates of the SIGEX screening, so that a possibility of obtaining desired catabolic genes can be increased. Finally, when a gene of interest is partial, IAN-PCR (inverse affinity nested PCR) can be used for obtaining franking genome fragments. This PCR has been developed for walking of rare genome fragments in environmental metagenomes (Uchiyama and Watanabe, 2006), which includes the following four steps: (i) inverse PCR in which one primer is connected to an affinity tag; (ii) affinity purification of PCR products for removing background metagenome; and (iii) nested PCR to recover target flanking regions. We have applied IAN-PCR to obtain a franking region of the SIGEX fragment BZO071, and, in work yet to be published, this reveals that the cytochrome P450 gene is followed by genes coding for putative ferredoxin and ferredoxin reductase - and present in the same transcriptional unit.

Conclusions

Even though the metagenomic approach has been proven promising for finding novel genes from environmental genetic resources, a large part of the environmental genetic diversity still remains unaccessed. Since each screening strategy has its own territory (in terms of the types of genes they can reach), the development of a new screening strategy should be the key to accessing hitherto unexploited genes with a high degree of novelty. According to the advent of SIGEX, we are now able to obtain catabolic genes whose enzyme activities are not easily expressed (e.g., cytochrome P450). Work so far therefore provides the platform for the next required step of attempting to develop other groundbreaking screening methods so as to expand the range of accessible genes in environmental metagenomes. In addition, we need to understand the utility and limitations of each screening method in order to facilitate the selection of the appropriate one for a desired gene or set of genes.

Acknowledgements

This work was supported by the New Energy and Industrial Technology Development Organization (NEDO).

References

AMANN, R., LUDWIG, W. AND SCHLEIFER, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological Reviews* **59,** 143-169.

BELL, P. J. L., SUNNA, A., GIBBS, M. D., CURACH, N. C., NEVALAINEN, H. AND BERGQUIST, P. L. (2002). Prospecting for novel lipase genes using PCR. *Microbiology* **148,** 2283-2291.

Brady, S. F., and Clardy, J. (2003). Synthesis of long-chain fatty acid enol esters isolated from an environmental DNA clone. *Org. Lett.* **5**, 121–24

Courtois, S., Cappellano, C. M., Ball, M., Francou, F. X., Normand, P., Helynck, G., Martinez, A., Kolvek, S. J., Hopke, J., Osburne, M. S., August, P. R., Nalin, R., Guerineau, M., Jeannin, P., Simonet, P. and Rernodet, J. L. (2003). Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Applied and Environmental Microbiology* **69,** 49-55.

Cowan, D. A. (2000). Microbial genomes – the untapped resource. *Trends in Biotechnology* **18,** 14-16.

Cowles, C. E., Nichols, N. N. and Harwood, C. S. (2000). BenR, a XylS homologe, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. *Journal of Bacteriology* **182,** 6339-6346.

Dunn, A. K., Klimowicz, A. K. and Handelsman, J. (2003). Use of a promoter trap to identify *Bacillus cereus* genes regulated by tomato seed exudate and a rhizosphere resident, *Pseudomonas aureofaciens*. *Applied and Environmental Microbiology* **69**, 1197-1205.

Eschenfeldt, W. H., Stols, L., Rosenbaum, H., Khambatta, Z. S., Quaite-Randall, E., Wu, S., Kilgore, D. C., Trent, J. D. and Donnelly, M. I. (2001). DNA from uncultured organisms as a source of 2,5-Diketo-D-Gluconic acid reductases. *Applied and Environmental Microbiology* **67,** 4206-4214.

Ferrer, M., Martinez-Abarca, F. and Golyshin, P. N. (2005). Mining genomes an 'metagenomes' for novel catalysts. *Current Opinion in Biotechnology* **16,** 1-6.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. and Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* **5**, 245-249.

Henne, A., Daniel, R., Schmitz, R. A. and Gottschalk, G. (1999). Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate. *Applied and Environmental Microbiology* **65,** 3901- 3907.

Henne, A., Schmitz, R. A., Bomeke, M., Gottschalk, G. and Daniel, R. (2000). Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Applied and Environmental Microbiology* **66,** 3113-3116.

Hino, S., Watanabe, K. and Takahashi, N. (1998). Phenol hydroxylase cloned from *Ralstonia eutropha* strain E2 exhibits novel kinetic properties. *Microbiology* **144,** 1765-1772.

Kubota, K., Nodate, M., Yasumoto-Hirose, M., Uchiyama, T., Kagami, O., Shizuri, Y. and Misawa, N. (2005). Isolation and functional analysis of cytochrome P450 CYP153A genes from various environments. *Bioscience, Biotechnology, and Biochemistry* **69,** 2421-2430.

Lorenz, P., Liebeton, K., Niehaus, F. and Eck, J. (2002). Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Current Opinion in Biotechnol*ogy **13**, 572-577.

Marques, S. and Ramos, J. L. (1993). Transcriptional control of the *Pseudomonas putida* TOL plasmid catabolic pathways. *Molecular Microbiology* **9**, 923-929.

Martinez, A., Kolvek, S. J., Yip, C. L. T., Hopke, J., Brown, K. A., MacNeil, I. A. and Osburne, M. S. (2004). Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and

detecting heterologous natural products in multiple expression hosts. *Applied and Environmental Microbiology* **70,** 2452-2463.

OKUTA, A., OHNISHI, K. AND HARAYAMA, S. (1998). PCR isolation of catechol 2,3-dioxygenase gene fragments from environmental samples and their assembly into functional genes. *Gene* **212,** 221-218.

PARSEK, M. R., SHINABARGER, D. L., ROTHMEL, R. K. AND CHAKRABARTY, A. M. (1992). Roles of CatR and cis,cis-muconate in activation of the catBC operon, which is involved in benzoate degradation in *Pseudomonas putida*. *Journal of Bacteriology* **174**, 7798-7806.

PETERSON, J. A., LORENCE, M. C. AND AMARNEH, B. (1990). Putidaredoxin reductase and putidaredoxin. *Journal of Biological Chemistry* **265,** 6066-6073.

PETERSON, J. A., LU, J. Y., GEISSELSODER, J., GRAHAM-LORENCE, S., CARMONA, C., WITNEY, F. AND LORENCE, M. C. (1992). Cytochrome P-450terp. Isolation and purification of the protein and cloning and sequencing of its operon. *Journal of Biological Chemistry* **267,** 14193-14203.

RONDON, M. R., AUGUST, P. A., BETTERMANN, A. D., BRADY, S. F., GROSSMAN, T. H., LILES, M. R., LOIACONO, K. A., LYNCH, B. A., MACNEIL, I. A., MINOR, C., TIONG, C. L., GILMAN, M., OSBURNE, M. S., CLARDY, J., HANDELSMAN, J. AND GOODMAN, R. M. (2000). Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology* **66,** 2541-2547.

ROTHMEL, R K., ALDRICH, T. L., HOUGHTON, J. E., COCO, W. M., ORNSTON, L. N. AND CHAKRABARTY, A. M. (1990). Nucleotide sequencing and characterization of *Pseudomonas putida catR*: a positive regulator of the *catBC* operon is a member of the LysR family. *Journal of Bacteriology* **172,** 922-931.

UCHIYAMA, T., ABE, T., IKEMURA, T. AND WATANABE, K. (2005). Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. *Nature Biotechnology* **23,** 88-93.

UCHIYAMA, T. and WATANABE, K. (2006). Improved inverse PCR scheme for metagenome walking. *BioTechniques* **41,** 183-188.

VALDIVIA, R. H. AND FALKOW, S. (1997). Fluorescence-based isolation of bacterial genes expressed within host cells. *Science* **277**, 2007-2011.

WATANABE, K., KODAMA, Y., SYUTSUBO, K. AND HARAYAMA, S. (2000). Molecular characterization of bacterial populations in petroleum-contaminated groundwater discharged from underground crude-oil-storage cavities. *Applied and Environmental Microbiology* **66,** 4803-4809.