Research review paper

Stable isotope probing in the metagenomics era: A bridge towards improved bioremediation

Ondrej Uhlik a, Mary-Cathrine Leewis b, Michal Strejcek a, Lucie Musilova a, Martina Mackova a,1, Mary Beth Leigh b, Tomas Macek a,*

a Institute of Chemical Technology Prague, Faculty of Food and Biochemical Technology, Department of Biochemistry and Microbiology, Technicka 3, 166 28 Prague, Czech Republic
b Institute of Arctic Biology, University of Alaska Fairbanks, 902 N. Koyukuk Dr., Fairbanks, AK 99775-7000, USA

Abstract

Microbial biodegradation and biotransformation reactions are essential to most bioremediation processes, yet the specific organisms, genes, and mechanisms involved are often not well understood. Stable isotope probing (SIP) enables researchers to directly link microbial metabolic capability to phylogenetic and metagenomic information within a community context by tracking isotopically labeled substances into phylogenetically and functionally informative biomarkers. SIP is thus applicable as a tool for the identification of active members of the microbial community and associated genes integral to the community functional potential, such as biodegradative processes.

The rapid evolution of SIP over the last decade and integration with metagenomics provide researchers with a much deeper insight into potential biodegradative genes, processes, and applications, thereby enabling an improved mechanistic understanding that can facilitate advances in the field of bioremediation.

© 2012 Elsevier Inc. All rights reserved.

Contents

1. Introduction .............................................................. 155
2. Principles of stable isotope probing (SIP) ................................................ 155
3. The metagenomics era ......................................................... 155
4. Integration of SIP with metagenomics .................................................. 158
4.1. SIP and metagenomics to study biodegradation of ecologically significant C1 compounds ............................ 158
4.2. SIP and metagenomics to study biodegradation of anthropogenic compounds ............................................ 158
5. Potential contributions of SIP and metagenomics to bioremediation .................................... 159
5.1. Assessing bioremediation potential ................................................ 159
5.2. Bioaugmentation ................................................................ 160
5.3. Understanding the mechanisms underlying biostimulation and phytoremediation technologies ................................................. 160
5.4. Carbon flow through contaminated systems ........................................ 160
6. Limitations .............................................................. 160
7. New frontiers of SIP .......................................................... 161
8. Alternatives to SIP ........................................................... 161
9. Conclusions .............................................................. 162
Acknowledgements ............................................................. 162
References ................................................................. 162

* Corresponding author. Tel: +420 220445140; fax: +420 220445167.
E-mail addresses: ondrej.uhlik@vscht.cz (O. Uhlik), mcleewis@alaska.edu (M.-C. Leewis), michal.strejcek@vscht.cz (M. Strejcek), lucie.musilova@vscht.cz (L. Musilova), mbleigh@alaska.edu (M.B. Leigh), tomas.macek@vscht.cz (T. Macek).

1 In loving memory of Prof. Martina Macková (May 7, 1965–August 2, 2012).

0734-9750/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.biotechadv.2012.09.003
1. Introduction

In addition to their critical role in global biogeochemical cycling, microbes play an essential role in the degradation, mineralization, or transformation of environmental pollutants and thus are potentially capable of restoring contaminated sites (Diaz, 2004; Geoffrey, 2010; Liu and Sufiita, 1993). The natural biodegradative processes occurring in contaminated sites are known as intrinsic bioremediation or natural attenuation (Jørgensen, 2007; Mulligan and Yong, 2004). These microbial biodegradative processes can often be accelerated using various strategies, known as bioremediation. Together with phytoremediation, microbial mediated bioremediation is generally considered an environmentally friendly, inexpensive, publicly accepted and promising means to remove contaminants from the environment (Macek et al., 2000; Schnoor et al., 1995).

When designing a strategy for the bioremediation of any contaminated site, an understanding of the indigenous microbial community can be highly valuable. Information integral to the success of bioremediation strategies can include: (i) identification of the microorganisms present in the contaminated site, (ii) investigation of their metabolic capabilities, and (iii) understanding of potential microbial community shifts in response to changing environmental factors (Lovley, 2003). Much of this information, however, has been difficult to elucidate and is rarely achieved with traditional microbiological techniques. Until recently, knowledge of microbes involved in bioremediation processes had been based mostly on culture-dependent studies, which do not account for the fact that laboratory conditions differ substantially from the environment. Cultivation-based techniques have been shown to target only about 1% of microbes occurring in the environment (Luzupone and Knight, 2008; Torsvik and Øvresæ, 2002; Zhang and Xu, 2008). More importantly, cultivation often fails to predict which microbes and which specific metabolic pathways will be active under realistic environmental conditions (Morales and Holben, 2011). The advent of molecular microbial ecology enabled culture-independent phylogenetic analyses of communities and functional genes, however, linking contaminant transformation to phylogenetic identity and specific genes/enzymes of metabolically active microbes remained a major challenge and still required cultivation. The development of stable isotope probing (SIP) was instrumental in circumventing the limitations of culture based investigations of biodegradation (Friedrich, 2006; Wellington et al., 2003). The coupling of SIP with the rapidly advancing field of metagenomics is further increasing its potential benefits to the field of bioremediation. This review aims to discuss the emerging trends in stable isotope probing techniques integrated with metagenomics, which can provide researchers with an unparalleled understanding of contaminant biodegradation and biotransformation processes in the environment.

2. Principles of stable isotope probing (SIP)

SIP tracks the incorporation of heavy stable isotopes, mainly $^{13}$C (Chen and Murrell, 2010; Dumont and Murrell, 2005; Madsen, 2010; Neufeld et al., 2007b; Radajewski et al., 2003; Uhlik et al., 2009a), $^{15}$N (Bell et al., 2011; Buckley et al., 2007a,b; Roh et al., 2009), or rarely $^{18}$O and $^2$H (Aanderud and Lennon, 2011; Woods et al., 2011), from specific substrates into phylogenetically informative biomarkers associated with microbes that assimilate the substrate. After stable isotopes have been pulsed into the environment and metabolically active cells have incorporated the label into their biomass, biomarkers are recovered and analyzed (Dumford and Neufeld, 2010; Neufeld et al., 2007b) (Fig. 1). Therefore, SIP is an approach that can identify microbial populations with a defined function. The first biomarkers introduced in SIP studies were phospholipid-derived fatty acids (PLFA) (Boschker et al., 1998), followed by DNA (Radajewski et al., 2000), and rRNA (Manefeld et al., 2002). Although it is the most sensitive of the three, PLFA-SIP is restricted to the classification of broad taxonomic groups (Table 1). Analyzing nucleic acids (RNA or DNA) can be far more informative taxonomically, with rRNA being a more responsive biomarker than DNA. This is due to the fact that the rates of RNA synthesis are much higher than those of DNA, and rRNA labeling is independent of cellular replication (Manefeld et al., 2002, 2004; Whiteley et al., 2006). In addition, reverse transcription and sequence analysis of 16S rRNA provides equal resolution for phylogenetic identification as sequence analysis of 16S rRNA genes (DNA, Table 1). Ribosomal RNA-based analyses, however, cannot provide access to the functional genes responsible for the metabolic capabilities of the community, and generally $^{13}$C-labeled mRNA is challenging to isolate in sufficient quantities for SIP (Neufeld et al., 2007a; Simon and Daniel, 2009). However, success has been achieved with mRNA-SIP for detection of naphthalene dioxygenase transcripts following SIP incubations with naphthalene (Huang et al., 2009). Further research employing DNA-, mRNA-, and rRNA-SIP (Dumont et al., 2011) demonstrated higher rates of labeling of functional mRNAs than their genes, and hence higher sensitivity of mRNA-SIP compared to DNA-SIP (Table 1). In addition, protein-SIP has recently been developed, taking advantage of proteins as a combined indicator for a specific metabolic activity as well as for obtaining phylogenetic information (Table 1, Jehmlich et al., 2008a,b, 2009).

To date, when functional gene analyses coupled to specific assimilatory processes are desired, DNA-SIP has been most often employed, due to the relative ease of detecting functional genes in $^{13}$C-DNA compared to mRNA. The use of either DNA-SIP or RNA-SIP can enable both the phylogenetic identification of those microbes performing the process as well as key metabolic genes possessed by the active populations likely to be involved in the process. When coupled with the rapidly expanding field of metagenomics, DNA-SIP has the potential for a focused, in-depth analysis of the collective genomes of the community active in a particular biodegradative process. This provides definitive information about the populations active in specific assimilatory processes and also requires considerably less sequencing effort than a full metagenomic analysis of the total community.

3. The metagenomics era

Metagenomics (also known as ecological genomics, community genomics, or environmental genomics) is a discipline that uses genomic methods to analyze natural ecological communities, namely the collective genomes in an environmental community (Handelsman et al., 1998; Riesenfeld et al., 2004). The major goal of metagenomics is to explicate the genomes of uncultured microbes, thereby permitting investigation of the broad diversity of taxonomically and phylogenetically relevant genes, individual catabolic genes, and whole operons (Schmeisser et al., 2007). Metagenomics itself was initially recognized for its potential to aid in discovery of novel biomolecules for biotechnological applications (Riesenfeld et al., 2004). Although the basic concept of metagenomics was first introduced at the end of last century (Handelsman et al., 1998), early forms of metagenomics had begun to emerge previously, with one example being the phylogenetic analyses of environmental microbial communities (Pace et al., 1985).

The approach introduced by Handelsman et al. (1998) involves extraction of the metagenome (genomic DNA from all organisms inhabiting the environment), its fragmentation, cloning, transformation, and subsequent screening of the constructed metagenomic library. The primary aim is to screen environmental communities for a specific biological activity and identify genes or gene clusters associated with it; also referred to as function-based screening (Yun and Ryu, 2005). The advent of high-throughput next generation DNA sequencing (e.g. 454 pyrosequencing, Illumina (Solexa) sequencing, SOLID sequencing) gave rise to another approach to metagenomics,
sequence-based screening. This method was first demonstrated by environmental genome shotgun sequencing of the Sargasso Sea (Venter et al., 2004), showing the potential of revealing the vast phylogenetic and metabolic diversity of microbial communities (Yun and Ryu, 2005). In addition to sequence-based screening of environmental metagenomic libraries, direct pyrosequencing of environmental communities is possible, bypassing cloning completely (Edwards et al., 2006). Because communities from habitats such as soils and sediments are often too diverse to permit screening in sufficient depth, even with the use of high-throughput sequencing, gene-targeted metagenomics techniques have emerged (Iwai et al., 2011), which are based on sequencing of PCR-generated amplicons (Fig. 2).

In addition to high-throughput sequencing, metagenomic analyses have recently been performed with the use of high-throughput microarrays (Fig. 2). These have been used to analyze microbial communities and monitor environmental biogeochemical processes. GeoChip microarrays (He et al., 2007, 2010) currently contain 83,992 50-meric sequences covering approximately 152,414 genes encoding for enzymes responsible for biogeochemical (C, N, P, S) cycling, metabolic processes, heavy metal resistance, antibiotic resistance, degradation of pollutants, and gyrB genes (Hazen et al., 2010; Lu et al., 2012). Marker gyrB encoding for the gyrase β-subunit is used instead of the more common 16S rRNA genes as probes for 16S rRNA usually do not provide resolution below genus level. gyrB
can be used to differentiate even closely related species (He et al., 2010). GeoChip microarrays can be therefore used to study structure, dynamics, and potential metabolic activity of microbial communities and their variations depending on different stimuli. Another type of microarray valuable to microbial ecology and contaminant biodegradation is the PhyloChip, which is used for high throughput
phylegetic analyses of microbial communities (Brodie et al., 2006), and has been used for a variety of applications including assessing microbial community responses to petroleum contamination (DeAngelis et al., 2011; Hazen et al., 2010).

With the first applications of metagenomic techniques, it became apparent that they enable the discovery of genomic and metabolic diversity that had not been previously imagined (Schloss and Handelsman, 2005). As research progressed, however, the main drawbacks of metagenomics were realized: the inability to link specific functions to individual populations and to achieve full sequence coverage in more complex communities (Vieites et al., 2009). By combining metagenomic techniques with SIP, these limitations can be significantly reduced. SIP experiments are designed to provide a targeted analysis of the active populations, omitting the inactive majority that is not the focus of the study (Fig. 2).

4. Integration of SIP with metagenomics

One of the main drawbacks of total community metagenomics is the unlikelihood of detecting particular genes of interest due to the extremely vast diversity and abundance of microbial genes occurring in most ecosystems, even when function-based metagenomic screening is used. Targeting metagenomics to specific subpopulations that are likely to contain the genes of interest, as with SIP, may overcome this obstacle (Schloss and Handelsman, 2003). This was demonstrated by Schwarz et al. (2006) who isolated genes encoding for coenzyyme B12-dependent glycerol dehydratases. The source for this key enzyme for the anaerobic dehydration of glycerol was the enrichment cultures of 13C-glycerol-fermenting microorganisms from a sediment sample. When metagenomic library construction was preceded by SIP, the frequency of clones bearing target genes was increased by almost four fold. However, expressing target genes successfully in vitro, such as in metagenomic libraries, can be very challenging in some instances. Using high-throughput sequencing technologies for direct shotgun sequencing of SIP-derived metagenomes can aid in overcoming this obstacle. Selective enrichment of targeted populations, whose diversity is much less complex than that of total communities, with subsequent isolation of the particular functional metagenome of interest increases the feasibility of achieving coverage and assembly of individual genomes with significantly reduced efforts and sequencing cost. Shotgun sequencing of SIP-derived metagenomes can also help ensure that portions of the microbial community that have low abundance but are integral to the metabolic processes of interest will not be overlooked (Wellington et al., 2003). The major drawback of this approach is usually the recovery of DNA in quantities too small to be sufficient for shotgun sequencing. Advances in multiple displacement amplification over the last years (Binga et al., 2008) have helped to overcome this limitation. Although no bioremediation studies have yet been published performing direct shotgun sequencing of SIP-derived metagenomes, they are very likely to arise in the near future.

4.1. SIP and metagenomics to study biodegradation of ecologically significant C1 compounds

One of the pilot studies integrating DNA-SIP with metagenomics revealed a complete methane monoxygenase operon in forest soils (Dumont et al., 2006). Methane and other one-carbon (C1) compounds are of global ecological significance because they can affect climate change, influence atmospheric and marine chemistry, and impact cloud formation. In the context of bioremediation, methanotrophs and/or methylotrophs have been implicated in the degradation of trichloroethylene and cis-1,2-dichloroethylene (Arai et al., 1999; Little et al., 1988; Shigematsu et al., 1999; Takeuchi et al., 2005), insecticides (Topp et al., 1993), nitro-substituted explosives (Van Aken et al., 2004), methyl halides (Warner et al., 2005), methyl tert-butyl ether (Kane et al., 2007; Nakatsu et al., 2006), and other xenobiotic compounds. Therefore, investigating the metabolism of C1 compounds is also potentially valuable for bioremediation.

Dumont et al. (2006) were the first to apply SIP in combination with function- and sequence-based metagenomic library screening in this field. After the incubation of a soil sample with 13CH4, 13C-DNA was used for the construction of a metagenomic library using a bacterial artificial chromosome (BAC). Subsequent screening of the library for key methylotrophy genes resulted in the discovery of a clone carrying a pmoABC operon, encoding for the particulate methane monoxygenase. A complete sequence of the operon was determined by shotgun sequencing. The sequence of the identified pmoA gene was almost identical to a Methylocystis sp. sequence which had been previously detected in this soil (Radajewski et al., 2002). Additionally, 12 other putative genes were detected on the same clone (Dumont et al., 2006), 3 of which take part in C1 metabolism. Labeled methylotroph populations were also analyzed phylogenetically by 16S rRNA gene DGGE fingerprints and subsequent sequencing of dominant DGGE bands. In addition to Methylocystis, the methanotrophic genera Methylobacter and Methyloccella were identified together with sequences similar to Bacteroidetes and γ-Proteobacteria.

One of the main criticisms associated with SIP is considered to be inappropriately high concentrations of labeled substrates introduced during incubations compared to concentrations that occur in situ. However, these high concentrations were necessary to achieve sufficient yields of labeled DNA for metagenomic analysis. Murrell and colleagues were the first to resolve this issue (Neufeld et al., 2008). Their strategy was the application of multiple displacement amplification to “bridge the gap” between the picogram quantities of labeled DNA and required micogram quantities for subsequent metagenomic analyses. Their study found Methylphaga spp. to be involved in oceanic methanol cycling and detected a 9-kb DNA fragment that encoded for the enzymes involved in methanol dehydrogenase synthesis, regulation, and assembly. Similarly, the techniques described were used for the analysis of as yet uncultured Methylocystis-related populations in acidic peatlands (Chen et al., 2008). These populations, which had been found to be dominant in the majority of acidic peatlands sampled, were further confirmed to be actively involved in oxidizing methane by SIP-based investigations. After 13CH4 had been assimilated, 13C-labeled DNA containing the genome of Methylocystis spp. was used for a construction of a metagenomic library and screened for the presence of key methylotrophic genes. Shotgun sequencing of a clone containing methanol dehydrogenase gene permitted the researchers to assemble a gene cluster encoding polypeptides involved in methanol utilization (mxaFGIRSAC). These reports (Chen et al., 2008; Dumont et al., 2006; Neufeld et al., 2008) were the first ones to show that retrieval of targeted genetic information can be achieved with minimal sequencing effort. At the same time, the authors proposed early sequencing of complete genomes of microbial populations directly from the environment (Dumont et al., 2006). Not long after, a nearly complete genome was obtained of a novel uncultured methylotrophic bacterium Methylothetema mobilis from the water and sediments of Lake Washington, WA, USA (Kaluzhnaya et al., 2008). This proof-of-principle study shows that genomes of ecologically relevant subpopulations can be reassembled after whole genome shotgun sequencing of stable isotopically labeled DNA. Additional results of this study revealed several clades of bacteria involved in C1 substrate metabolism; some were traditional methylotrophs such as Methylobacter, Methylophanas, Methyloversatilis, or Ralstonia, and others clustered with Verrucomicrobia, Nitrospira, and Planctomycetes, clades not commonly associated with methylotrophs.

4.2. SIP and metagenomics to study biodegradation of anthropogenic compounds

Some of the most widespread and environmentally significant xenobiotics are polychlorinated biphenyls (PCBs) (Breivik et al., 2002).
Correspondingly, studies combining SIP and metagenomics to study PCB-degrading bacteria are common. The first such study performed DNA-SIP integrated with GeoChip-mediated metagenomic analysis of bacteria in the root zone of an Austrian pine (*Pinus nigra* L.) growing naturally in PCB-contaminated soil using $^{13}$C-biphenyl, a PCB analog, as a substrate (Leigh et al., 2007). The findings of this study pointed to novel populations of biphenyl-utilizing bacteria, including *Pseudomonas*, *Nocardioidea*, *Kribella*, *Variovorax*, and *Polaromonas* in addition to previously known PCB-degrading *Sphingomonas* spp. GeoChip analyses of $^{13}$C-DNA detected 30 genes associated with organic contaminant degradation in the $^{13}$C-DNA, majority of which were associated with the degradation of aromatics, including biphenyl, benzoate, catechol, protocatechuate, naphthalene, phenol, dibenzofuran, and phenylpropionate. The presence of these genes in biphenyl-labeled populations suggests that they have the potential to degrade several aromatic substrates. In addition, genes of the $\beta$-ketoacylase pathway were detected indicating potential abilities of the populations to mineralize monoaromatic once biphenyl has been transformed into monoaromatic intermediates. Only four genes, however, were detected from the biphenyl upper pathway (*bph* operon) associated with rhodococci and bacilli, suggesting that only a tiny fraction of the actual diversity in the upper biphenyl pathway genes had been revealed. This hypothesis was supported by PCR amplification and sequence analyses of genes encoding for aromatic ring hydroxylating dioxygenases (ARHD), all of which shared homology but were not identical to those previously deposited in GenBank. Thus, these novel genes were undetected using the microarray. Moreover, some of the sequences did not cluster with any known ARHDs and represented a novel clade.

Sul et al. (2009) applied metagenomics directly to isolate a novel biphenyl dioxygenase (*bph*) gene from PCB-contaminated river sediment bacteria enriched by the incubation with $^{13}$C-biphenyl. Biphenyl dioxygenase (BphAE), a multicomponent enzyme catalyzing the activation of biphenyl ring by insertion of two oxygen atoms, is crucial for biodegradation of biphenyl. Degradation of PCBs is permitted by relaxed substrate specificity of the enzyme, which has been determined to be closely connected with its primary structure (Vézina et al., 2007, 2008). The dioxygenase sequence detected by Sul et al. (2009) was highly similar to that in *Pseudomonas* sp. Cam-1 and *Pseudomonas* pseudocaligenes KF707. Although in most laboratory PCB-degrading strains, genes *bphAE* are organized in operons with other enzymes for subsequent transformation of dihydroxylated biphenyl, this clone only contained *bphAE* genes. The authors ascribe this phenomenon to an acquisition of the genes from another microorganism during exposure of the sediment to PCBs, possibly by horizontal gene transfer. This hypothesis was supported by a different G+C content of *bphAE* than the average for the cloned fragment. The activity of BphAE was tested after expression of the genes *bphAE* from the cosmid clone along with *bphFCB* from *Burkholderia xenovorans* LB400. The spectrum of metabolized PCB congeners was similar to that of *P. pseudocaligenes* KF707, transforming only the congeners without chloro substitutions at the 2,3 positions.

The identification of previously characterized organisms and genes by Sul et al. (2009) may be due to the long SIP incubation and enrichment of fast-growing organisms that would be amenable to detection using cultivation-based approaches. In the future, combining direct shotgun metagenomic sequencing or function-based screening with DNA-SIP could enable the discovery of highly novel degradative genes, rather than those with similarity to known sequences as were detected with Leigh et al. (2007) and Sul et al. (2009) using PCR-based or microarray-based detection. Although Sul et al. (2009) did not discover a dioxygenase with a broader substrate specificity than had been observed previously, this study contributes to our understanding of genomic features of degradative populations. This particular case suggests that the gene organization of *bph* genes in nature might be scattered rather than clustered in operons. The idea of catabolic genes being dispersed on chromosomes and plasmids was supported by another paper which discusses the organization of aromatic degradation pathway genes (Suenaga et al., 2009). Thirty-eight fosmid clones were analyzed carrying genes for extradiol dioxygenases, and only two of the clones contained complete degradation pathways that are commonly found in aromatic compound-utilizing isolates. The other clones contained only subsets of the pathway genes with novel gene arrangements.

Recent results (Uhlík et al., 2012) demonstrated the ability of biphenyl-metabolizing bacteria to utilize other aromatic compounds in contaminated soil; populations of *Rhodanobacter*, *Burkholderia*, *Pandorea*, *Dyella* and other *Proteobacteria* were observed to derive carbon from benzoate and naphthalene in addition to biphenyl. This study combined SIP with sequence analysis of 16S rRNA gene pyrotags amplified from $^{13}$C-DNA to identify taxa associated with the biodegradation of pollutants. Results of a few recently published bioremediation-related SIP studies reveal bacteria that had not been associated with utilization of the substrates before. Examples include newly associated populations of *Pseudomonas* or *Rhodanobacter* with the degradation of biphenyl (Lee et al., 2011; Uhlík et al., 2012) and *Thermimcola* with the degradation of toluene (Pilloni et al., 2011). In addition, many unclassified 16S rRNA gene sequences were retrieved from $^{13}$C-DNA labeled by different substrates pointing to novel yet-to-be described bacterial taxa involved in biodegradation of biphenyl, benzoate, naphthalene, or toluene (Lee et al., 2011; Pilloni et al., 2011; Uhlík et al., 2012).

5. Potential contributions of SIP and metagenomics to bioremediation

Fundamental research on microbial aspects of bioremediation improves understandings of processes and could potentially improve bioremediation technologies. Thorough analysis of the labeled metagenomes of bioremediative populations can provide valuable information for assessing bioremediation potential of autochthonous microorganisms as well as designing and monitoring engineered bioremediation strategies.

5.1. Assessing bioremediation potential

Before bioremediation strategies are applied to a contaminated site, the bioremediation potential of the indigenous microflora should be assessed. In this case, SIP can be valuable for determining whether organisms capable of metabolizing the contaminant are already present at the site. If so, then biostimulation would likely be a viable bioremediation strategy. For these purposes, SIP incubations can be performed either *in vitro* using microcosms constructed from field-collected samples (Leigh et al., 2007; Uhlík et al., 2009b; Winderl et al., 2010) or *directly in situ* (Bobbio et al., 2010; DeRito et al., 2005; Liou et al., 2008; Mahmood et al., 2005; Padmanabhan et al., 2003; Pumphrey and Madsen, 2008). Metagenomic functional gene analyses of SIP studies are particularly valuable in the case of biodegradation of xenobiotics that occur as mixtures, such as PCBs, since biphenyl dioxygenase enzymes vary widely in their capability to degrade different lower chlorinated PCB congeners (Erickson and Mondello, 1993; Mondello et al., 1997). Analyzing the SIP-labeled dioxygenase gene sequence in relation to known enzymes with known substrate specificities can help predict which of the congeners are likely to be degraded by the microbial community (Barriauskut et al., 2002; Vézina et al., 2008). The absence or inactivity of dioxygenases with appropriate congener specificity for the PCBs on-site might indicate that an anaerobic treatment to promote dehalogenation would be appropriate (Smidt and de Vos, 2004; Wiegel and Wu, 2000). In addition, metagenomic exploration of active populations can clarify the metabolic capabilities as well as regulatory mechanisms within microbes, such as through sequence-based detection of regulatory
elements, and these properties can be subjected to genetic manipulations with the aim of improving the efficacy of bioremediation through bioaugmentation.

5.2. Bioaugmentation

SIP is ideal for identifying microbes optimal for bioaugmentation for several reasons. First, SIP-metagenomics has the potential to reveal the identity of microbial species that are metabolically active under a variety of environmental conditions (type of matrix, temperature, moisture, oxygen levels, etc.) or that derive carbon from more than one contaminant. Second, once bioaugmentation with exogenous organisms has been applied on a contaminated site, SIP can help determine whether the added microorganisms have adapted to new conditions and survived and, more importantly, if they are actively biodegrading the contaminant(s).

5.3. Understanding the mechanisms underlying biostimulation and phytoremediation technologies

Biostimulation methods seek to address limitations in environmental conditions and/or bioavailability in order to enable indigenous microbial communities to more rapidly biodegrade the contaminant. One approach to biostimulation is the use of plants to promote biodegradative activity in the root zone, which is a form of phytoremediation known as rhizoremediation (Gerhardt et al., 2009; Kuiper et al., 2004; Macek et al., 2009; Macková et al., 2006). Plants can rhizostimulate contaminant biodegradation through a variety of mechanisms. Root exudation and root decay provides rhizosphere microorganisms with growth substrates and secondary compounds that may function as comestabolites or inducers of biodegradative pathways, or surfactants and phytochemicals that increase the bioavailability of poorly soluble pollutants (Bertin et al., 2003; Leigh et al., 2002, 2006; Singer et al., 2003, 2004; Toussaint et al., 2012; Yi and Crowley, 2007). In addition, the rhizosphere is also richer in oxygen essential for the activity of aerobic organisms and their dioxygenases and monoxygenases that are often involved in biodegradation processes of organic contaminants (Leigh et al., 2002). Understanding how the microbial community responds to biostimulation is important to improving bioremediation technologies. SIP and metagenomics approaches could provide novel and important insight into the mechanisms underlying rhizoremediation (Prosser et al., 2006) by identifying contaminant degraders active under different environmental and biostimulatory conditions.

Plants also host many bacteria in their endosphere. Emerging research indicates that they contribute to biodegradation of toxic organic compounds in contaminated soil and could have potential for improving phytoremediation (McGuinness and Dowling, 2009; Newman and Reynolds, 2005; Weyens et al., 2009, 2011). Although no studies have as of yet employed SIP-metagenomics to study microbial metabolism inside plants, this topic is likely to be investigated with respect to bioremediation in the future.

Some plants contribute to remediation more directly, by taking up and transforming contaminants or storing them in above-ground parts (Macek et al., 2000). Studies have been published showing that plants, however, have limited abilities to mineralize pollutants. In case of PCBs, for instance, monohydroxylated and/or dihydroxylated derivatives are formed without the aromatic ring being cleaved (Rezek et al., 2008). In connection with improving phytoremediation efficiency, novel genes revealed by SIP-metagenomics can be used in order to prepare genetically modified plants with abilities to cleave aromatic structures (Macek et al., 2008; Novákova et al., 2009; Sylvestre et al., 2009; Van Aken et al., 2010).

SIP appears to be an ideal tool for use during laboratory tests evaluating how different biostimulation strategies alter biodegradative microbial populations, and the resultant information could also potentially be useful for creating and monitoring engineered bioremediation treatments.

5.4. Carbon flow through contaminated systems

One of the major drawbacks of DNA-SIP is the detection of cross-feeding populations in addition to primary utilizers of a substrate (Neufeld et al., 2007c). Experimental data have shown a very rapid sequestration of carbon in some microbial communities (Luiders et al., 2006). In addition, mineralization of 13C-labeled substrates generates 13CO2, which can further be taken up by autotrophic bacteria and incorporated into their biomass biasing thereby the conclusions of a study. Therefore, different strategies have been suggested to minimize the effect of cross-feeding detection by DNA-SIP or to clearly distinguish it from primary utilization. These include time-course experiments (Leigh et al., 2007), shortening the incubation times and addition of 13C-carrier DNA in the density gradient separation (Gallagher et al., 2005), or combination of DNA-SIP with RNA-SIP (Dumont et al., 2011). Yet in the context of bioremediation, cross-feeding populations may also be important contributors. For instance, intermediate metabolites or incompletely-degraded substrates released by the biodegradative population may be removed from the environment by cross-feeding microbes. Monitoring cross-feeding also has the potential to identify all microorganisms essential to complete degradation of contaminants and to reveal the flow of carbon through the microbial food web in contaminated environments.

Recently, a new SIP methodology has been developed which is ideal for determining carbon flow through food webs. Known as stable isotope switching (SIS) (Maxfield et al., 2012), the approach is based on the incubation of the matrix with 13C-labeled substrate that is switched for an unlabeled substrate when full labeling is achieved. Such a setting allows for monitoring carbon uptake, turnover, release, and sequestration.

6. Limitations

Despite the great potential contribution of SIP and metagenomics to microbial ecology and biotechnology, one has to realistically evaluate the limitations associated with each of the approaches. A major challenge associated with SIP is a very limited availability and high cost of labeled substrates. SIP is also very labor-intensive and low-throughput using current techniques. These factors may prohibit performing a sufficient number of replicates to enable statistically valid comparisons among treatments. In the future, the use of robotics and automation could assist in increasing the throughput of SIP gradient fractionation and assessments. Another challenge of SIP is the need to combine many techniques and technologies spanning many disciplines, which may be an obstacle for researchers without experience in these techniques (reviewed by Chen and Murrell, 2010).

When designing SIP-based experiments, it is also important to use realistic substrate concentrations and incubation times that are appropriate duration to enable labeling yet minimize potential cross-feeding and over-enrichment (Neufeld et al., 2007a). Inappropriately high concentrations of a labeled substrate can lead to population shifts, such as opportunistic growth of some degraders, suppression of populations adapted to only low concentrations of the substrate, or shifts caused by accumulation of inhibitory or toxic intermediates.

It has been proposed earlier (Madsen, 2006) that, especially for bioremediation purposes, SIP assay conditions should ideally match the conditions in the field. Field conditions are constantly changing (temperature, humidity, etc.) and may be affected by larger scale processes such as living plant roots, and these factors are challenging to adequately represent under laboratory microcosms. Microcosms often create other unrealistic conditions associated with depletion
of oxygen, increased concentrations of CO₂, unrealistic substrate concentrations, or unrealistic availability of the substrate (Friedrich, 2006). Therefore, laboratory microcosms need to be recognized as model systems, with the caveat that they will not fully represent field conditions. Thorough bioremediation surveys require either performing SIP directly in situ (Bombach et al., 2010; DeRito et al., 2005; Liou et al., 2008; Mahmood et al., 2005; Padmanabhan et al., 2003; Pumphrey and Madsen, 2008) or a very careful design and monitoring of microcosm experiments.

One of the major limitations of SIP-metagenomics for bioremediation purposes is the restriction of its utility solely to assimilatory processes (Andreoni and Gianfreda, 2007). Yet many bioremediation processes are dissimilatory, for example, with the contaminant serving as the terminal electron acceptor. Many anthropogenic compounds are halogenated (polychlorinated dioxins, dibenzofurans, biphenyls, or chlorinated solvents), and their bioremediation relies on dehalorespiration (i.e., utilization of a halogenated compound as a terminal electron acceptor to yield energy) (Futagami et al., 2008). These processes can be monitored through different stable isotope analyses (for review, see Ruess and Chamberlain, 2010; Boecklen et al., 2011; Braeckevelt et al., 2012).

Metagenomics techniques are also not exempt from limitations. Rapidly advancing high-throughput sequencing technologies have substantially reduced the cost of sequencing efforts over traditional Sanger sequencing. Sequence-based screening of metagenomic data, however, relies on databases that are far from being completely or accurately annotated. The incompleteness of genetic databases means that about 30% of the sequences sequenced can remain unassigned (Harrington et al., 2007), which prevents metagenomics from achieving its full potential. Other potential problems are associated with the function-based screening of metagenomic libraries. These include the necessity of analyzing a large number of clones to recover positive ones (Sul et al., 2009), possible inabilities of the host cells (usually E. coli) to express foreign genes and to form active proteins, or problems associated with cloning of large gene clusters which requires high-molecular weight DNA of high purity (Daniel, 2005; Simon and Daniel, 2009). At the same time, problems associated with gene-targeted metagenomics have been identified. These include errors introduced during PCR by DNA polymerase, generation of chimeric sequences when heterologous templates are used, or errors introduced during pyrosequencing. As a result of these phenomena, diversity is overestimated (Kunin et al., 2010) unless appropriate tools are used for sequence processing. Introduction of several denoising algorithms (Quince et al., 2009, 2011; Reeder and Knight, 2010), improved operational taxonomic unit (OTU) clustering (Huse et al., 2010), or their combination (Schloss et al., 2011; Uhlík et al., 2012) have resulted in much more accurate diversity estimations. Additionally, the most commonly used phylogenetic marker gene is 16S rRNA, which is present in different copy numbers in different species, ranging from 1 to 15 (Pei et al., 2010). As a result, the relative abundance of 16S rRNA genes in a sequence library does not necessarily correspond to the number of cells, and may distort the apparent community structure.

7. New frontiers of SIP

Realizing the limitations of DNA-SIP and metagenomics gives rise to alternative approaches being considered. One of these is SIP-metatranscriptomics. mRNA-SIP benefits from high sensitivity and the fact that mRNA bears functional information. Studies carried out so far by Huang et al. (2009) and Dumont et al. (2011) focused on one transcript each—naphthalene dioxygenase and methane monooxygenase, respectively. The rapidly advancing field of metatranscriptomics (Fillitraut et al., 2011; Simon and Daniel, 2011; Su et al., 2012) seems to have overcome some of the major obstacles linked with mRNA, such as low recovery of high-quality environmental transcripts, instability of mRNA, or difficulties in separation of mRNA from other RNA species (Feike et al., 2012; Marchetti et al., 2012; Mou et al., 2011; Ottesen et al., 2011; Rinta-Kanto et al., 2012; Stewart et al., 2012). Therefore, it is only a matter of time before mRNA-SIP is combined with metatranscriptomic analyses to provide comprehensive information on the actively transcribed genes associated with the utilization of a specific substrate.

Another group of biomarkers bearing both phylogenetic and functional information are proteins. Protein-SIP was initially applied to track carbon flow in pure cultures (Jehmlich et al., 2008a,b), however the authors proposed its use for the analysis of phylogenetically diverse microbial communities as well, which was reported shortly thereafter (Bastida et al., 2010, 2011; Pan et al., 2011). More recently, protein-SIP was used for quantitative analysis of induced proteins in substrate shift experiments (Taubert et al., 2011). Protein-SIP is currently very challenging due to the inability to assign even a putative function to many proteins. Yet protein-SIP is very promising as it may provide a more substantial access to real microbial activity, since proteins are the most explicit indicators of metabolic activity. Therefore, in the future, with more complete databases, better modeling systems, and further development of metaproteomics techniques, protein-SIP is expected to widely expand and advance the field of microbial ecology.

8. Alternatives to SIP

Although SIP is the main focus of this review, other technologies expanding the scope of isotope labeling experiments have been developed that can potentially link microbial phylogeny with metabolic activity (Gutierrez-Zamora and Manefield, 2010). Most of these methods involve fluorescence in situ hybridization (FISH), which is a technique commonly applied to taxonomically identify microbial cells using rRNA-targeted oligonucleotide probes (Wagner et al., 2003). FISH developed into a function-identity method after it was combined with (i) microautoradiography (FISH-MAR), which uses radioactive isotopes to monitor microbial uptake of labeled substrates (Lee et al., 1999); (ii) secondary ion mass spectrometry (FISH-SIMS) and nanometer-scale secondary ion mass spectrometry (FISH-nanoSIMS), which determines the isotopic composition of the FISH-identified targeted cells (Kuypers and Jørgensen, 2007; Li et al., 2008; Orphan et al., 2001); (iii) Raman microspectroscopy (Raman-FISH), which detects vibrational shifts of covalent bonds in molecules of FISH-labeled cells (Huang et al., 2007); or (iv) immunomagnetic cell capture (magneto-FISH), which permits targeted magnetic capture of FISH-labeled microorganisms and cell aggregates (Pernthaler et al., 2008). Behrens et al. (2008) reported a new combination of FISH and nanoSIMS for linking microbial phylogeny to metabolic activity at the single cell level by applying enhanced element labeling of microbial cells by FISH (EL-FISH). Such a modification allowed them to increase the sensitivity of cell detection and broadened the applicability of the methodology for environmental studies. Some of these techniques, mainly FISH-MAR, have been employed in bioremediation media (Hesselsoe et al., 2008; Yang et al., 2003).

In addition to FISH, other isotope-based techniques have been used such as isotope arrays (Adamczyk et al., 2003) or shotgun isotope arrays (Tobino et al., 2011), small subunit-isotope ratio mass spectrometry (SSU-IRMS) (MacGregor et al., 2002, 2006), or radioactive isotope probing (RIP) (Nikolausz et al., 2007). The advantages and limitations of these methods as well as their comparison to SIP are discussed elsewhere (Gutierrez-Zamora and Manefield, 2010; Neufeld et al., 2007c; Wagner et al., 2006). Recently some FISH-based methods have been combined with SIP in order to link microbial phylogeny to metabolic activity at the single-cell level. For example, Huang et al. (2009) combined rRNA and mRNA-SIP with single-cell Raman-FISH. The main advantage of
such a combined approach is that it allows for determination and quantification of in situ functions of a microbial community. RNA-SIP has also been combined with magnetic-bead capture hybridization (MacGregor et al., 2006; Miyatake et al., 2009). This approach is applicable for linking phylogeny with metabolic activity at the level of class or family while requiring approximately 10,000 times less stable isotope enrichment of RNA, allowing thus the use of environmentally relevant concentrations of the isotopically labeled target substrate. The principle of RNA-SIP combined with magnetic-bead capture hybridization lies in extracting RNA from the matrix after it has been stable isotope labeled, hybridizing the RNA with probes, capturing those hybrids with beads, and collecting beads with a magnet. Captured RNA is then released and analyzed by isotope ratio mass spectrometry (IRMS). This approach thus eliminates possible biases associated with the use of artificially high substrate concentrations during incubations relative to those in situ, which is one of the major criticisms of SIP (Dumont and Murrell, 2005; Neufeld et al., 2008).

9. Conclusions
Since the first development of SIP more than a decade ago (Boschker et al., 1998; Radajewski et al., 2000), this technique has progressed rapidly and has significantly broadened the field of microbial ecology, first by linking phylogenetic identity with function, including that of novel clades with no cultured representatives, and later by linking function with novel functional gene sequences. The recent integration of SIP with metagenomics has enabled a more comprehensive understanding of the functional community dynamics of entire microbial systems. An improved mechanistic understanding of microbial ecological function has the potential to enable new breakthroughs in bioremediation technologies such as biostimulation, bioaugmentation and phytoremediation. The great advantage of SIP is its ability to enable a focused detection and analysis of only the organisms active in the utilization of a specific substrate, either directly or indirectly through the food web. In the context of bioremediation, SIP-metagenomics is invaluable for revealing the identity of contaminant-degraders and their functional genes, and this information may be used to assess their response to biostimulation methods or to identify organisms and genes useful for bioaugmentation. As with any method, SIP has limitations, such as false positives, in which cross-feeding organisms may be mistaken for primary utilizers, and the potential for microbial community change during incubations. However if these factors are taken into account, SIP combined with metagenomics has the potential to provide novel insights to the intricate interactions within microbial ecosystems. Limitations like cross-feeding can also be advantageous, providing valuable new insight into the flow of carbon derived from contaminants through the microbial food web. As cultivation-independent methods like SIP and metagenomic studies advance, it is important to note that cultivation-based techniques are still a crucial means to verify and investigate the physiology and genetics of individual contaminant-degrading microorganisms, to facilitate bioaugmentation, and to enable improved annotation of metagenomic databases. The integration of SIP with the rapidly advancing field of metagenomics is opening new windows into microbial processes and interactions, which may enable major breakthroughs in the field of bioremediation.

Acknowledgements
The authors acknowledge the support of Czech Ministry of Education, Youth and Sports (grant ME09024), and the 7th Framework Programme (grant MINOTaurus, no: 265946). This publication was made possible by grants from the National Center for Research Resources (5P20RR16466-12) and the National Institute of General Medical Sciences (8P20GM103395-12) from the National Institutes of Health. MCL acknowledges the David L. Boren National Security Education Program (NSEP) Fellowship for support.

References


