Targeted metagenomics of active microbial populations with stable-isotope probing

Sara Coyotzi[a], Jennifer Pratscher[b], J. Colin Murrell[b], Josh D. Neufeld[a*]

[a] Department of Biology, University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

[b] School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

*Josh D. Neufeld: +1(519)888-4567 ext. 38344, jneufeld@uwaterloo.ca

Summary

The ability to explore microbial diversity and function has been enhanced by novel experimental and computational tools. The incorporation of stable isotopes into microbial biomass enables the recovery of labeled nucleic acids from active microorganisms, despite their initial abundance and culturability. Combining stable-isotope probing (SIP) with metagenomics provides access to genomes from microorganisms involved in metabolic processes of interest. Studies using metagenomic analysis on DNA obtained from DNA-SIP incubations can be ideal for the recovery of novel enzymes for biotechnology applications, including biodegradation, biotransformation, and biosynthesis. This chapter introduces metagenomic and DNA-SIP methodologies, highlights biotechnology-focused studies that combine these approaches, and provides perspectives on future uses of these methods as analysis tools for applied and environmental microbiology.

Introduction to metagenomics

Metagenomics involves the direct capture and analysis of deoxyribonucleic acid (DNA) mixtures from microbial communities, allowing access to the encoded functional traits and identities of its members. An important advantage of metagenomics is the ability to recover new enzymes through function-based analyses. In contrast to a purely sequence-based screening of metagenomic libraries, screening or selecting for function is, theoretically, more likely to discover completely novel enzymes given that prior genetic information is not important for the recovery of gene targets. Examples of recovering targets based on function include discoveries of lipases from an Atlantic rainforest soil [1] and intertidal flat sediments [2,3], and an esterase from South China Sea sediments [4], all of which were assigned to novel enzyme families.

One limitation of metagenomic studies is that sequencing of bulk DNA or capturing of DNA in large-insert libraries focuses primarily on genomic sequences from relatively abundant microorganisms. For low-diversity microbial habitats [5,6], including engineered environments [7], metagenomics can achieve relatively complete genomic coverage in metagenomic libraries with available sequencing capacity. However, although a 40,000-fosmid library might capture around 400 genomes of an average size of 4 Mbp [8], this falls short of capturing the majority of taxa in a species-rich environment, such as soil, which has potentially greater than 10^5 different genomes per g of material [8]. Because of the potential limitations associated with high diversity, metagenomic studies can be preceded by careful selection of environmental samples in order to target a particular set of microorganisms and enzymes with specific activity profiles that may be enriched. Such careful sample selection can increase the likelihood of identifying genes and gene products possessing the characteristics that are desired for industrial processes, but may overlook many novel organisms and enzymes found naturally yet at low relative abundance within highly

diverse aquatic and terrestrial habitats. Alternative methods that filter for genomes of interest can help with accessing novel species and their encoded functions. Indeed, in addition to possessing important biogeochemical roles and activities of biotechnological significance, the rare biosphere [9] of many terrestrial and aquatic habitats must be accessed through such alternative targeted approaches, such as DNA stable-isotope probing (DNA-SIP).

Introduction to DNA-SIP

Since its development, DNA-SIP [10,11] has been a key technique linking the assimilation of particular growth substrates with specific populations of microorganisms, regardless of our ability to cultivate these microbes in the laboratory. The recovery of genomes from microorganisms that were active under selected study conditions is a major benefit of this method. The DNA-SIP technique involves incubating an environmental sample with a stable-isotope-labeled substrate. Although ^{13}C-labeled substrates have been most commonly used, other isotopes have also been used, including ^{15}N and multiple forms of labelled water (e.g., $H_2{}^{18}O$ or $^{2}H_2O$). Importantly, incubation conditions enable assimilation of the target substrate and incorporation of isotope into genomes (and lipids, proteins, and RNA) of active microorganisms during replication and cellular metabolism. After incubation, DNA is isolated from the environmental sample and isopycnic centrifugation is used to separate the isotope-labeled DNA ("heavy DNA") from that of the inactive community ("light DNA") for downstream analyses [12].

The versatility of DNA-SIP is reflected in the wide range of assimilatory processes studied by this technique since its inception. Examples demonstrating the wide-range of applications of DNA-SIP include methane assimilation in landfill cover soil [13] and peatlands [14], methanol

assimilation in coastal sea water [15] and soda lake sediments [16], and acetate assimilation in arsenic-polluted sediments [17]. DNA-SIP has also been used to identify active microorganisms that degrade and assimilate polycyclic aromatic hydrocarbons (PAHs) [18-21], pentachlorophenol [22], vinyl chloride [23], ethylbenzene [24], and plant-derived compounds such as cellulose [25,26] and mixed plant residues [27-29]. The method has also enabled analysis of microbial food webs [30], plant-microbe interactions [31], and the effect of earthworms in a landfill bacterial community [32]. Although heavy DNA fractions recovered from SIP incubations are often analyzed by 16S rRNA gene amplicon sequencing in order to characterize the taxonomic affiliations of the stable isotope labeled microorganisms, many studies leverage the availability of genome sequence information to target functional gene markers. Example studies have focused on *pmoA* and *mmoX* for methanotrophs [13], *mxaF* and *mauA* for methylotrophs [15], *arrA* for bacterial respiration of As(V) [17], *nirK* and *nirS* for denitrifiers [33], *etnE* for vinyl chloride assimilation [23], and *amoA* for ammonia-oxidizing bacteria [34].

Despite the widespread use of DNA-SIP for targeting active microorganisms, the limitations of this method include the potential for cross-feeding of label from primary consumers to other microbes, low recovery of heavy DNA, non-availability and high costs associated with the purchase of isotopically labeled substrates, and the requirement for relatively high concentrations of substrate for efficient labeling of DNA [35]. In order to help circumvent these weaknesses, careful experimental design with different incubation times can help detect and minimize cross-feeding [36,37]. Although the use of multiple displacement amplification (MDA) can help overcome low yields of labeled DNA, such DNA amplification approaches can introduce bias and chimeras during amplification [36,38], which can affect the fidelity of DNA-SIP data by changing the relative abundance of amplified genomes or modify collected genomic

data in ways that can be difficult to detect. Although subjecting amplified DNA to enzymatic reactions with S1 nuclease and DNA polymerase can help reduce bias and chimera formation [38], a more desirable recent direction is to use small amounts of labelled DNA as template for sequencing, without a MDA step [39]. When designing DNA-SIP experiments, it is necessary to use incubation conditions that are similar to *in situ* conditions in the environment and consideration should be given to substrate concentrations used, temperature, pH, and addition of supplemental nutrients. Rates of substrate utilization, when monitored, will also provide useful information to plan length of incubations; time course DNA-SIP experiments are recommended where possible [37]. Although *in situ*-like nutrient concentrations are important for studies with a microbial ecology focus, high substrate concentrations might not be a major concern when seeking to recover genomes from novel microorganisms that possess enzymes of interest for industrial or biotechnology applications [39].

Combining DNA-SIP and metagenomics

An important characteristic of DNA-SIP is its function as a filter, or "sieve", restricting the recovered genomes to those species of interest, which can then provide a targeted and more efficient search for "hits" within metagenomic libraries (Figure 1). Recently, DNA-SIP has been coupled with metagenomics (previously reviewed in [40]) for the recovery of novel enzymes [39,41], operons [42,43], and functional groups [44]. Additional ecological findings obtained from DNA-SIP based metagenomic approaches include the identification of methane monooxygenases distribution in selected environments. For example, DNA-SIP and metagenomics identified the presence of particulate methane monooxygenase and absence of soluble methane monooxygenase in oil sands tailing ponds [43]. The potential influence of the

environment on the metabolic capacity of active microorganisms has also been observed by combining these techniques. This is exemplified by unusual electron acceptor flexibility present in a *Desulfuromonas* strain in an oil spill-affected tidal flat sediment [45]. Another achievement is finding previously unknown metabolic capabilities in bacterial taxa, such as the discovery of dimethylsulfide degrading activity in the family *Methylophilaceae* [44]. The power of combining DNA-SIP and metagenomics has also been demonstrated in a recent study by Grob *et al.* [46] who used DNA-SIP to recover a near-complete genome of an uncultured *Methylophaga* species, which can metabolize methanol in the marine environment. Although many published studies have coupled isotopic labeling with metagenomic analyses (Table 1), recent examples of combined DNA-SIP and metagenomic studies, with a focus on biotechnology, are highlighted below.

The first description of the sequential implementation of DNA-SIP and metagenomic-library functional screening recovered DNA from taxa that had assimilated ^{13}C-labeled carbon from glycerol [47]. After screening a metagenomic library derived from heavy DNA, the authors detected coenzyme B_{12}-dependent glycerol dehydratases, their targeted enzyme, at a frequency that ranged between 2.1- and 3.8-fold higher than observed in a metagenome from a conventional enrichment culture with the native substrate. This study provided an initial demonstration of the potential value of employing DNA-SIP for targeted enzyme recovery. Although enrichment culture techniques could potentially recover some target populations labelled by DNA-SIP, an important advantage of isotope labeling is that isopycnic separation of labeled DNA provides an additional physical filter that helps maximize target populations for screening metagenomes with increased efficiency.

Physical separation of labeled DNA, in contrast to enrichment culture, may be particularly valuable with environmental samples possessing high diversity yet with unknown culturability of its constituent members. For example, to discover new enzymes for potential industrial applications, Verastegui and colleagues performed DNA-SIP and metagenomics [39], with a focus on the recovery of genomes encoding glycoside hydrolases from microorganisms involved in degrading plant-derived polymers. This study focused on Canadian soils from Arctic tundra, a temperate rainforest, and an agricultural field. Fully ^{13}C-labeled substrates included glucose, cellobiose, xylose, arabinose, and cellulose. Sequences of 16S rRNA gene amplicons retrieved from ^{13}C-labeled DNA revealed that many of the active species in these environmental samples remained unclassified. Of the taxa identified, *Actinomycetales* (*Salinibacterium*), *Rhizobiales* (*Devosia*), *Rhodospirillales* (*Telmatospirillum*), and *Caulobacterales* (*Phenylobacterium* and *Asticcacaulis*) were revealed to be the active indicator species (i.e., present in heavy DNA fractions and absent from light DNA fractions) when combining data from all soils and all substrates. Next, the authors sequenced the ^{13}C-labeled metagenomic DNA and the annotation highlighted the low database coverage of recovered environmental sequences (i.e., only ~20% of recovered sequences were annotated using the SwissProt database), and hence of the targeted study enzymes: glycoside hydrolases. Construction of a metagenomic cosmid-based library from the pooled ^{13}C-labeled DNA retrieved from the cellulose incubation of the three soil types was followed by functional screening in *Escherichia coli*. Activity screenings using several polymeric substrates resulted in the recovery of eight positive clones after screening 2,876 clones, which greatly exceeded the commonly encountered frequency of targets identified in comparable studies without a DNA-SIP incubation step. The authors only screened a small portion of the library (i.e., approximately 3.5% of the clones) and generated a single library from

the heavy DNA obtained from the cellulose substrate incubation, such that enormous functional potential remains untapped within the heavy DNA generated from this study.

High-throughput sequencing can provide an alternative approach for characterization of heavy DNA, circumventing functional screens (Figure 1). Indeed, sequencing, assembly, and annotation of the ^{13}C-labeled material from a DNA-SIP PAH degradation experiment subsequently helped with amplification of sequences coding for polycyclic aromatic hydrocarbon dioxygenases, directly from unfractionated DNA, instead of requiring comprehensive functional screens [41]. The strategy enabled primer design for targeted PCR amplification of the desired dioxygenase sequences. The result was the retrieval of complete gene sets encoding dioxygenase activity and enzyme substrate characterization through overexpression in *E. coli* and biochemical assays, which contributed to bioremediation-strategy development.

As exemplified above, combining DNA-SIP and metagenomics has enormous potential for the recovery of genes and genomes from microbes involved in substrate degradation for bioremediation purposes. Indeed, DNA-SIP combined with metagenomic analyses have resulted in the recovery of genes and/or genomes encoding biphenyl dioxygenase from river sediment [48], naphthalene biodegradation operons from contaminated groundwater [49], polycyclic aromatic hydrocarbons degradation activity from marine sediment [50] and soil [41], and toluene degradation capacity from marine sediment [45] and oil sands tailing ponds [51,52]. In this way, DNA-SIP can provide insight into the microbial ecology of an environment of interest while simultaneously enabling the recovery of genes and genomes encoding functions that can aid in the remediation of polluted terrestrial and aquatic sites. Identifying microbial community members that possess and express enzymes for transforming contaminants found *in situ* is an important first step for biodegradation research. For example, a recent study using DNA-SIP and

metagenomics to explore biodegradation potential in contaminated soil was able to recover DNA from active phenanthrene degraders and later express and characterize polycyclic aromatic hydrocarbon dioxygenases from these microorganisms [41]. This study evaluated and confirmed bioremediation potential in the community, thus paving the way for further biodegradation strategy designs involving bioaugmentation or biostimulation that could help reduce contaminant levels.

Future directions

Although metagenomic research has sampled from thousands of unique locations, only a small proportion of recovered genes, potentially encoding novel enzymes, have been characterized in detail from these studies [53]. Such an untapped reservoir of under-explored enzymatic potential suggests a need for faster and more efficient screening and characterization methods. For function-based screenings of metagenomic libraries generated via DNA-SIP (Figure 1), an obvious future approach is automation to achieve high-throughput analyses, as exemplified by the simultaneous detection of 14 enzymatic activities during the screening of a metagenomic library of 12,160 fosmid clones [54]. Enriched metagenomic libraries obtained via DNA-SIP incubations are ideal candidates for picodroplet-based functional screening approaches [55]. Because functional metagenomic approaches remain in their infancy, especially when combined with targeted approaches using DNA-SIP, better high-throughput screens will allow for more efficient enzyme discovery and may help assign function to many unknown polypeptides encoded in genome and metagenome databases.

Although recent studies highlighted above have used DNA-SIP to recover genomes from active microorganisms, this approach will likely become much more commonplace. One of the

challenges of assembling genomes from metagenomes is that the diversity of many microbial communities can make genome assembly difficult. The recent application of differential abundance binning can help circumvent this limitation by using information about shared increases or decreases in sequence datasets from similar samples to bin sequences prior to assembly [56-58]. Differential abundance binning is ideally suited to the analysis of metagenomes derived from DNA-SIP incubations. Because relatively few microbial community members become labeled, lower diversity and availability of multiple heavy fractions across replicate gradients should increase the likelihood that fully assembled genomes will be recovered from differential abundance binning efforts. To our knowledge, no published DNA-SIP studies have yet used differential abundance binning of metagenomic sequence data from heavy DNA prior to genome assembly. Thus, this direction is still open to proof-of-principle and widespread adoption for future studies, in addition to enabling further analysis of previously generated labeled DNA from past SIP experiments.

There are a number of well-established DNA manipulation techniques which can be used to construct new enzymes, such as mutagenesis, DNA gene shuffling, and chimeragenesis [59]. These techniques might also be coupled with isotopically labeled DNA from DNA-SIP experiments in the near future given that some metagenome-derived enzymes have already been improved in such ways [60,61]. Because naturally occurring genes in microbial communities already possess extensive polymorphisms [8,41], DNA retrieved directly from the environment contains an amazing wealth of genes with enormous potential for manipulation and application. Indeed, combining DNA-SIP and metagenomics is arguably the best methodological approach for accessing myriad genomes encoded within the rare biosphere of complex microbial communities for the discovery of new organisms and enzymes with novel properties.

Acknowledgements

JP acknowledges an Independent Research Fellowship (IRF) from the UK Natural Environment Research Council (NERC). JCM acknowledges funding by the Gordon and Betty Moore Foundation through Grant GBMF 3303 to the University of East Anglia (Murrell, J.C., principal investigator). JDN acknowledges Strategic Project and Discovery Grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

1. Faoro H, Glogauer A, Souza EM, Rigo LU, Cruz LM, Monteiro RA, Pedrosa FO: **Identification of a new lipase family in the Brazilian Atlantic Forest soil metagenome**. *Environ Microbiol Rep* 2011, **3**:750-755.
2. Lee M-H, Lee C-H, Oh T-K, Song JK, Yoon J-H: **Isolation and characterization of a novel lipase from a metagenomic library of tidal flat sediments: evidence for a new family of bacterial lipases**. *Appl Environ Microb* 2006, **72**:7406-7409.
3. Kim E-Y, Oh K-H, Lee M-H, Kang C-H, Oh T-K, Yoon J-H: **Novel cold-adapted alkaline lipase from an intertidal flat metagenome and proposal for a new family of bacterial lipases**. *Appl Environ Microb* 2009, **75**:257-260.
4. Fu C, Hu Y, Xie F, Guo H, Ashforth EJ, Polyak SW, Zhu B, Zhang L: **Molecular cloning and characterization of a new cold-active esterase from a deep-sea metagenomic library**. *Appl Microbiol Biotechnol* 2011, **90**:961-970.
5. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF: **Community structure and metabolism through reconstruction of microbial genomes from the environment**. *Nature* 2004, **428**:37-43.
6. Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, Gloeckner FO, Boffelli D, Anderson IJ, Barry KW, Shapiro HJ, et al.: **Symbiosis insights through metagenomic analysis of a microbial consortium**. *Nature* 2006, **443**:950-955.
7. Martin HG, Ivanova N, Kunin V, Warnecke F, Barry KW, McHardy AC, Yeates C, He S, Salamov AA, Szeto E, et al.: **Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities**. *Nat Biotech* 2006, **24**:1263-1269.
8. Vieites JM, Guazzaroni M-E, Beloqui A, Golyshin PN, Ferrer M: **Metagenomics approaches in systems microbiology**. *FEMS Microbiol Rev* 2009, **33**:236-255.
9. Lynch MD, Neufeld JD: **Ecology and exploration of the rare biosphere**. *Nat. Rev. Microbiol.* 2015, **13**:217-229.
10. Radajewski S, Ineson P, Parekh NR, Murrell JC: **Stable-isotope probing as a tool in microbial ecology**. *Nature* 2000, **403**:646-649.
11. Dumont MG, Murrell JC: **Stable isotope probing—linking microbial identity to function**. *Nature Rev Microbiol* 2005, **3**:499-504.
12. Neufeld JD, Vohra J, Dumont MG, Lueders T, Manefield M, Friedrich MW, Murrell JC: **DNA stable-isotope probing**. *Nat Protoc* 2007, **2**:860-866.
13. Cébron A, Bodrossy L, Chen Y, Singer AC, Thompson IP, Prosser JI, Murrell JC: **Identity of active methanotrophs in landfill cover soil as revealed by DNA-stable isotope probing**. *FEMS Microbiol Ecol* 2007, **62**:12-23.
14. Gupta V, Smemo KA, Yavitt JB, Basiliko N: **Active methanotrophs in two contrasting North American peatland ecosystems revealed using DNA-SIP**. *Microb Ecol* 2012, **63**:438-445.
15. Neufeld JD, Schafer H, Cox MJ, Boden R, McDonald IR, Murrell JC: **Stable-isotope probing implicates *Methylophaga* spp. and novel *Gammaproteobacteria* in marine methanol and methylamine metabolism**. *ISME J* 2007, **1**:480-491.
16. Antony CP, Kumaresan D, Ferrando L, Boden R, Moussard H, Scavino AF, Shouche YS, Murrell JC: **Active methylotrophs in the sediments of Lonar Lake, a saline and alkaline ecosystem formed by meteor impact**. *ISME J* 2010, **4**:1470-1480.

17. Lear G, Song B, Gault AG, Polya DA, Lloyd JR: **Molecular analysis of arsenate-reducing bacteria within Cambodian sediments following amendment with acetate**. *Appl Environ Microb* 2007, **73**:1041-1048.
18. Gutierrez T, Singleton DR, Berry D, Yang T, Aitken MD, Teske A: **Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP**. *ISME J* 2013, **7**:2091-2104.
19. Singleton DR, Powell SN, Sangaiah R, Gold A, Ball LM, Aitken MD: **Stable-isotope probing of bacteria capable of degrading salicylate, naphthalene, or phenanthrene in a bioreactor treating contaminated soil**. *Appl Environ Microbiol* 2005, **71**:1202-1209.
20. Singleton DR, Sangaiah R, Gold A, Ball LM, Aitken MD: **Identification and quantification of uncultivated Proteobacteria associated with pyrene degradation in a bioreactor treating PAH-contaminated soil**. *Environ Microbiol* 2006, **8**:1736-1745.
21. Song M, Jiang L, Zhang D, Luo C, Wang Y, Yu Z, Yin H, Zhang G: **Bacteria capable of degrading anthracene, phenanthrene, and fluoranthene as revealed by DNA based stable-isotope probing in a forest soil**. *J. Hazard. Mater.* 2016, **308**:50-57.
22. Li X, Lin Z, Luo C, Bai J, Sun Y, Li Y: **Enhanced microbial degradation of pentachlorophenol from soil in the presence of earthworms: Evidence of functional bacteria using DNA-stable isotope probing**. *Soil Biol. Biochem.* 2015, **81**:168-177.
23. Paes F, Liu X, Mattes TE, Cupples AM: **Elucidating carbon uptake from vinyl chloride using stable isotope probing and Illumina sequencing**. *Appl. Microbiol. Biotechnol.* 2015, **99**:7735-7743.
24. Jayamani I, Cupples AM: **Stable isotope probing and high-throughput sequencing implicate *Xanthomonadaceae* and *Rhodocyclaceae* in ethylbenzene degradation**. *Environ Eng Sci* 2015, **32**:240-249.
25. Li T, Mazeas L, Sghir A, Leblon G, Bouchez T: **Insights into networks of functional microbes catalysing methanization of cellulose under mesophilic conditions**. *Environ Microbiol* 2009, **11**:889-904.
26. Chapleur O, Mazeas L, Godon J-J, Bouchez T: **Asymmetrical response of anaerobic digestion microbiota to temperature changes**. *Appl. Microbiol. Biotechnol.* 2016, **100**:1445-1457.
27. Lee CG, Watanabe T, Sato Y, Murase J, Asakawa S, Kimura M: **Bacterial populations assimilating carbon from ^{13}C-labeled plant residue in soil: analysis by a DNA-SIP approach**. *Soil Biol Biochem* 2011, **43**:814-822.
28. Li Y, Lee CG, Watanabe T, Murase J, Asakawa S, Kimura M: **Identification of microbial communities that assimilate substrate from root cap cells in an aerobic soil using a DNA-SIP approach**. *Soil Biol Biochem* 2011, **43**:1928-1935.
29. Bernard L, Mougel C, Maron P-A, Nowak V, Lévêque J, Henault C, Haichar FeZ, Berge O, Marol C, Balesdent J, et al.: **Dynamics and identification of soil microbial populations actively assimilating carbon from ^{13}C-labelled wheat residue as estimated by DNA- and RNA-SIP techniques**. *Environ Microbiol* 2007, **9**:752-764.
30. Lueders T, Wagner B, Claus P, Friedrich MW: **Stable isotope probing of rRNA and DNA reveals a dynamic methylotroph community and trophic interactions with fungi and protozoa in oxic rice field soil**. *Environ Microbiol* 2004, **6**:60-72.
31. Lu Y, Lueders T, Friedrich MW, Conrad R: **Detecting active methanogenic populations on rice roots using stable isotope probing**. *Environ Microbiol* 2005, **7**:326-336.

32. Héry M, Singer AC, Kumaresan D, Bodrossy L, Stralis-Pavese N, Prosser JI, Thompson IP, Murrell JC: **Effect of earthworms on the community structure of active methanotrophic bacteria in a landfill cover soil**. *ISME J* 2008, **2**:92-104.
33. Osaka T, Yoshie S, Tsuneda S, Hirata A, Iwami N, Inamori Y: **Identification of acetate- or methanol-assimilating bacteria under nitrate-reducing conditions by stable-isotope probing**. *Microb Ecol* 2006, **52**:253-266.
34. Chen Y, Wu L, Boden R, Hillebrand A, Kumaresan D, Moussard H, Baciu M, Lu Y, Colin Murrell J: **Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave**. *ISME J* 2009, **3**:1093-1104.
35. Pinnell LJ, Charles TC, Neufeld JD: **Stable isotope probing and metagenomics**. In *Stable Isotope Probing and Related Technologies*. Edited by Murrell JC, Whiteley AS: American Society of Microbiology; 2011:97-114.
36. Neufeld JD, Chen Y, Dumont MG, Murrell JC: **Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics**. *Environ Microbiol* 2008, **10**:1526-1535.
37. Neufeld JD, Dumont MG, Vohra J, Murrell JC: **Methodological considerations for the use of stable isotope probing in microbial ecology**. *Microb Ecol* 2007, **53**:435-442.
38. Chen Y, Dumont MG, Neufeld JD, Bodrossy L, Stralis-Pavese N, McNamara NP, Ostle N, Briones MJI, Murrell JC: **Revealing the uncultivated majority: combining DNA stable-isotope probing, multiple displacement amplification and metagenomic analyses of uncultivated *Methylocystis* in acidic peatlands**. *Environ Microbiol* 2008, **10**:2609-2622.
39. Verastegui Y, Cheng J, Engel K, Kolczynski D, Mortimer S, Lavigne J, Montalibet J, Romantsov T, Hall M, McConkey BJ, et al.: **Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial communities**. *mBio* 2014, **5**:e01157-01114.
40. Chen Y, Murrell JC: **When metagenomics meets stable-isotope probing: progress and perspectives**. *Trends Microbiol* 2010, **18**:157-163.
41. Chemerys A, Pelletier E, Cruaud C, Martin F, Violet F, Jouanneau Y: **Characterization of novel polycyclic aromatic hydrocarbon dioxygenases from the bacterial metagenomic DNA of a contaminated soil**. *Appl Environ Microb* 2014, **80**:6591-6600.
42. Dumont MG, Radajewski SM, Miguez CB, McDonald IR, Murrell JC: **Identification of a complete methane monooxygenase operon from soil by combining stable isotope probing and metagenomic analysis**. *Environ Microbiol* 2006, **8**:1240-1250.
43. Saidi-Mehrabad A, He Z, Tamas I, Sharp CE, Brady AL, Rochman FF, Bodrossy L, Abell GC, Penner T, Dong X, et al.: **Methanotrophic bacteria in oilsands tailings ponds of northern Alberta**. *ISME J* 2013, **7**:908-921.
44. Eyice O, Namura M, Chen Y, Mead A, Samavedam S, Schafer H: **SIP metagenomics identifies uncultivated *Methylophilaceae* as dimethylsulphide degrading bacteria in soil and lake sediment**. *ISME J* 2015, **9**:2336-2348.
45. Kim S-J, Park S-J, Cha I-T, Min D, Kim J-S, Chung W-H, Chae J-C, Jeon CO, Rhee S-K: **Metabolic versatility of toluene-degrading, iron-reducing bacteria in tidal flat sediment, characterized by stable isotope probing-based metagenomic analysis**. *Environ Microbiol* 2014, **16**:189-204.
46. Grob C, Taubert M, Howat AM, Burns OJ, Dixon JL, Richnow HH, Jehmlich N, von Bergen M, Chen Y, Murrell JC: **Combining metagenomics with metaproteomics and stable**

isotope probing reveals metabolic pathways used by a naturally occurring marine methylotroph. *Environ Microbiol* 2015, **17**:4007-4018.

47. Schwarz S, Waschkowitz T, Daniel R: **Enhancement of gene detection frequencies by combining DNA-based stable-isotope probing with the construction of metagenomic DNA libraries**. *World J Microb Biot* 2006, **22**:363-368.
48. Sul WJ, Park J, Quensen JF, Rodrigues JLM, Seliger L, Tsoi TV, Zylstra GJ, Tiedje JM: **DNA-stable isotope probing integrated with metagenomics for retrieval of biphenyl dioxygenase genes from polychlorinated biphenyl-contaminated river sediment**. *Appl Environ Microb* 2009, **75**:5501-5506.
49. Wang Y, Chen Y, Zhou Q, Huang S, Ning K, Xu J, Kalin RM, Rolfe S, Huang WE: **A culture-independent approach to unravel uncultured bacteria and functional genes in a complex microbial community**. *PLoS ONE* 2012, **7**:e47530.
50. Kim S-J, Park S-J, Jung M-Y, Kim J-G, Madsen EL, Rhee S-K: **An uncultivated nitrate-reducing member of the genus *Herminiimonas* degrades toluene**. *Appl Environ Microb* 2014, **80**:3233-3243.
51. Laban NA, Dao A, Foght J: **DNA stable isotope probing of oil sands tailings pond enrichment cultures reveals different key players for toluene degradation under methanogenic and sulfidogenic conditions**. *FEMS Microbiol Ecol* 2015, **91**:fiv03.
52. Abu Laban N, Tan B, Dao A, Foght J: **Draft genome sequence of uncultivated toluene-degrading *Desulfobulbaceae* bacterium Tol-SR, obtained by stable isotope probing using [$^{13}C_6$]toluene**. *Genome Announc* 2015, **3**:e01423-01414.
53. Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina OV, Golyshin PN: **Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends**. *Microb Biotechnol* 2015:doi: 10.1111/1751-7915.12309.
54. Nyyssonen M, Tran HM, Karaoz U, Weihe C, Hadi MZ, Martiny JB, Martiny AC, Brodie EL: **Coupled high-throughput functional screening and next generation sequencing for identification of plant polymer decomposing enzymes in metagenomic libraries**. *Front Microbiol* 2013, **4**:282.
55. Colin P-Y, Kintses B, Gielen F, Miton CM, Fischer G, Mohamed MF, Hyvonen M, Morgavi DP, Janssen DB, Hollfelder F: **Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics**. *Nat. Commun.* 2015, **6**.
56. Imelfort M, Parks D, Woodcroft BJ, Dennis P, Hugenholtz P, Tyson GW: **GroopM: an automated tool for the recovery of population genomes from related metagenomes**. *PeerJ* 2014, **2**:e603.
57. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH: **Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes**. *Nat. Biotechnol.* 2013, **31**:533-538.
58. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C: **Binning metagenomic contigs by coverage and composition**. *Nature Methods* 2014, **11**:1144-1146.
59. Moussard H, Smith TJ, Murrell JC: **DNA stable isotope probing and gene mining**. In *Stable Isotope Probing and Related Technologies*. Edited by Murrell JC, Whiteley AS: American Society of Microbiology; 2011:259-276.
60. Yin X, Ma L, Pei X, Du P, Li C, Xie T, Yu L, Yu L, Wang Q: **Creation of functionally diverse chimerical α-glucosidase enzymes by swapping homologous gene fragments retrieved from soil DNA**. *Indian J Microbiol* 2015, **55**:114-117.

61. Qian C, Liu N, Yan X, Wang Q, Zhou Z, Wang Q: **Engineering a high-performance, metagenomic-derived novel xylanase with improved soluble protein yield and thermostability**. *Enzyme Microb Technol* 2015, **70**:35-41.
62. Kalyuhznaya MG, Martens-Habbena W, Wang T, Hackett M, Stolyar SM, Stahl DA, Lidstrom ME, Chistoserdova L: ***Methylophilaceae*** **link methanol oxidation to denitrification in freshwater lake sediment as suggested by stable isotope probing and pure culture analysis**. *Environ Microbiol Rep* 2009, **1**:385-392.
63. Pinnell LJ, Dunford E, Ronan P, Hausner M, Neufeld JD: **Recovering glycoside hydrolase genes from active tundra cellulolytic bacteria**. *Can J Microbiol* 2014, **60**:469-476.
64. Abu Laban N, Tan B, Dao A, Foght J: **Draft genome sequence of uncultivated** ***Desulfosporosinus*** **sp. strain Tol-M, obtained by stable isotope probing using [$^{13}C_6$]toluene**. *Genome Announc.* 2015, **3**:e01422-01414.

*18: Used labeled PAHs to identify active marine oil spill degraders, including a comparison of ^{13}C-DNA to metagenomes from the same samples.
**39: Combining DNA-SIP and functional metagenomics accesses novel glycoside hydrolases from soil microorganisms, for a range of potential industry applications.
**41: Accessing novel ring-hydroxylating dioxygenases by leveraging information obtained via DNA-SIP incubations.
*44: Study demonstrates that DNA-SIP can identify bacteria that would otherwise be missed by traditional cultivation- or metagenomic-based analyses, including an observation that multiple-displacement amplification did not bias genomic DNA amplification detectably.
**46: Insight into identity and function of active microbial community members by exploring SIP-labeled nucleic acid and proteins.
*52: Good example of using DNA-SIP and metagenomics for genome assembly from heavy DNA.
*54: Example of an ideal methodological analysis for heavy DNA derived from DNA-SIP incubations.
**57: First demonstration of how abundance data can improve genome assemblies, which is ideally suited to the analysis of heavy DNA from SIP incubations

Figure legend

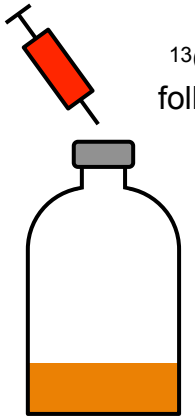
Figure 1. Stable-isotope probing (SIP) of nucleic acids coupled to downstream sequence analysis approaches.

Table 1. Examples of coupled DNA-SIP and metagenomic studies.

Habitat	Functional group (substrate)	Heavy DNA analyses	Key findings	Ref
Forest soil	Methanotrophs ($^{13}CH_4$)	BAC library for SBS and probe design for colony hybridization	Sequence obtained for particulate methane monooxygenase operon *pmoCAB*.	[42]
Wadden Sea sediment	Glycerol dehydratases (^{13}C-glycerol)	Plasmid library for SBS with PCR-amplified dehydratase-gene hybridization and FBS	Gene detection frequencies with DNA-SIP ~2.6-fold higher than with unlabeled DNA.	[47]
Lake sediment	Methylotrophs (^{13}C one carbon compounds)	Plasmid library for sequencing, assembly, and bioinformatic analyses	Reconstruction of nearly complete genome of novel methylotroph *Methylotenera mobilis*	[62]
Marine-surface water	Methylotrophs (^{13}C- methanol)	MDA-based fosmid library for SBS, followed by insert sequencing and assembly	*Methylophaga*-like phylotypes active. Methanol dehydrogenase gene cluster recovery.	[36]
European peatlands	Methanotrophs ($^{13}CH_4$)	MDA-based fosmid library for SBS, followed by insert sequencing and assembly	*Methylocystis*-related species active. Methanol utilization gene cluster recovered.	[38]
Contaminated sediment	Biphenyl dioxygenases (^{13}C-biphenyl)	Cosmid library for SBS	Recovery of biphenyl dioxygenase subunit genes *bphAE* with demonstrated activity.	[48]
Contaminated groundwater	Naphthalene degraders (^{13}C-naphthalene)	MDA for HTS, assembly, *in silico* SBS, and primer design for targeted recovery	*Acidovorax* sp. activity. Novel naphthalene degradation operon *nag2*. HGT detected.	[49]
Oilsands tailings ponds	Methanotrophs ($^{13}CH_4$)	MDA for HTS, assembly, and *in silico* SBS	*Methylocaldum* and *Methylomonas* spp. active. *pmoCAB* operon and *pmoA* genes.	[43]
Contaminated soil	PAH dioxygenases (^{13}C-phenanthrene)	HTS, assembly, and annotation, followed by specific primer design for gene recovery	4 ring-hydroxylating dioxygenase gene sets recovered, overexpressed, and characterized.	[41]
Oil spill-affected tidal flat sediment	IRB/aromatic hydrocarbon degradation (^{13}C-toluene)	HTS, assembly, and annotation. *In silico* aromatic hydrocarbon degradation SBS	*Desulfuromonas* active. Draft genome reveals capacity to use broad range of e^- acceptors.	[45]
Contaminated sediment	NRB/ aromatic hydrocarbon degradation (^{13}C-toluene)	HTS, assembly, and annotation. *In silico* aromatic hydrocarbon metabolism analyses	*Herminiimonas*-related bacterium capable of anaerobic aromatic hydrocarbon degradation.	[50]
Artic tundra	GHs (^{13}C- glucose, cellulose)	MDA for HTS, assembly, and annotation, followed by *in silico* GH analyses	Active populations vary with substrate. Target genes increased in SIP libraries.	[63]
Various soils	GHs (^{13}C-glucose, arabinose, cellobiose, xylose, cellulose)	MDA-based cosmid library for FBS	Positive GHs clones in higher proportion than in non DNA-SIP metagenomic studies.	[39]
Agricultural soil and lake sediment	Dimethylsulfide degraders (^{13}C-dimethylsulfide)	MDA for HTS, assembly, and annotation. *In silico* methylotrophy, sulfur metabolic SBS	*Methylophilaceae* involved in dimethylsulfide degradation, a novel function for this family.	[44]
Oil sands tailing pond	Toluene degradation/ methanogenic, sulfidogenic conditions (^{13}C-toluene)	HTS, assembly, and annotation	Genome of *Desulfosporosinus* sp. Tol-M and *Desulfobulbaceae* bacterium Tol-SR.	[51,52,64]
Coastal seawater	Methylotrophs (^{13}C-methanol)	MDA for HTS, followed by genome assembly	*Methylophaga* genome reveals methanol metabolism; confirmed by metaproteomics.	[46]

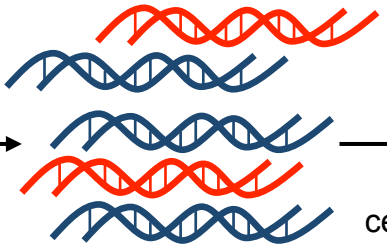
FBS=function-based screening, GH=glycoside hydrolase, HGT=horizontal gene transfer, HTS=high-throughput sequencing, IRB=Iron-reducing bacteria, MDA=Multiple displacement amplification, NRB=Nitrate-reducing bacteria, PAH=Polycyclic aromatic hydrocarbon, SBS=sequence-based screening.

Figure



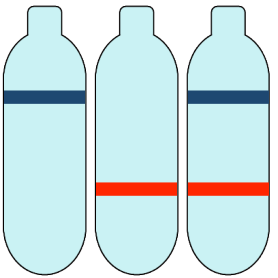
^{13}C substrate added, followed by incubation

Extraction of total DNA and/or RNA



^{12}C and ^{13}C DNA or RNA

Isopycnic centrifugation



Isolation of ^{13}C DNA or RNA by fractionation

^{13}C DNA or RNA

Phylogeny (e.g., using 16S rRNA gene transcripts or amplicons)

Metagenomics (e.g., functional or sequence-based)

Functional gene analysis (e.g., probes or amplicon sequencing)