

## MINI-REVIEW

# Insights from quantitative metaproteomics and protein-stable isotope probing into microbial ecology

Martin von Bergen<sup>1,2,3,9</sup>, Nico Jehmlich<sup>1,4,9</sup>, Martin Taubert<sup>1,5</sup>, Carsten Vogt<sup>6</sup>, Felipe Bastida<sup>7</sup>, Florian-Alexander Herbst<sup>1</sup>, Frank Schmidt<sup>4</sup>, Hans-Hermann Richnow<sup>6</sup> and Jana Seifert<sup>1,8</sup>

<sup>1</sup>Department of Proteomics, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany;

<sup>2</sup>Department of Metabolomics, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany;

<sup>3</sup>Department of Biotechnology, Chemistry and Environmental Engineering, Aalborg University, Aalborg, Denmark; <sup>4</sup>Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany; <sup>5</sup>School of Environmental Sciences, University of East Anglia, Norwich, UK;

<sup>6</sup>Department of Isotope Biogeochemistry, UFZ—Helmholtz Centre for Environmental Research, Leipzig, Germany; <sup>7</sup>Department of Soil and Water Conservation, Group of Soil Enzymology and Bioremediation and Organic Wastes CEBAS-CSIC, Campus Universitario de Espinardo, Espinardo, Murcia, Spain and

<sup>8</sup>Institute of Animal Nutrition, University of Hohenheim, Stuttgart, Germany

**The recent development of metaproteomics has enabled the direct identification and quantification of expressed proteins from microbial communities *in situ*, without the need for microbial enrichment. This became possible by (1) significant increases in quality and quantity of metagenome data and by improvements of (2) accuracy and (3) sensitivity of modern mass spectrometers (MS). The identification of physiologically relevant enzymes can help to understand the role of specific species within a community or an ecological niche. Beside identification, relative and absolute quantitation is also crucial. We will review label-free and label-based methods of quantitation in MS-based proteome analysis and the contribution of quantitative proteome data to microbial ecology. Additionally, approaches of protein-based stable isotope probing (protein-SIP) for deciphering community structures are reviewed. Information on the species-specific metabolic activity can be obtained when substrates or nutrients are labeled with stable isotopes in a protein-SIP approach. The stable isotopes (<sup>13</sup>C, <sup>15</sup>N, <sup>36</sup>S) are incorporated into proteins and the rate of incorporation can be used for assessing the metabolic activity of the corresponding species. We will focus on the relevance of the metabolic and phylogenetic information retrieved with protein-SIP studies and for detecting and quantifying the carbon flux within microbial consortia. Furthermore, the combination of protein-SIP with established tools in microbial ecology such as other stable isotope probing techniques are discussed.**

*The ISME Journal* (2013) 7, 1877–1885; doi:10.1038/ismej.2013.78; published online 16 May 2013

**Subject Category:** Integrated genomics and post-genomics approaches in microbial ecology

**Keywords:** quantitative metaproteomics; label-free quantification; spectral counting; stable isotope probing; protein-SIP; functional microbial community

## Introduction

Contemporary microbial ecology studies use one or several ‘omics’ approaches to gain a deeper insight into the structure and function of microbial communities. Among these approaches, metaproteomics provides the direct link between genetic potential

and actual functionality of the community. In recent years, a number of authors reviewed the steadily increasing number of metaproteomic studies and the ongoing technical progress, and almost all concluded that there is a need for quantitative input of metaproteomic studies (Keller and Hettich, 2009; Wilmes and Bond, 2009; Verberkmoes *et al.*, 2009a; Schneider and Riedel, 2010). The rationale behind this is the often found redundancy of functionalities in complex communities, requiring quantitative data to decide which species is mainly responsible for some metabolic functions. Quantitation can be achieved by employing label-free approaches (Old *et al.*, 2005; Neilson *et al.*, 2011), chemical labeling

Correspondence: J Seifert, Department of Proteomics, UFZ—Helmholtz Centre for Environmental Research, Permoserstrasse 15, D-04318 Leipzig, Germany.  
E-mail: jana.seifert@ufz.de

<sup>9</sup>These authors contributed equally to this work.

Received 12 November 2012; revised 8 April 2013; accepted 13 April 2013; published online 16 May 2013

such as isobaric tags (iTRAQ, TMT) (Thompson *et al.*, 2003; Ong and Mann, 2005) or dynamic incorporation of stable isotopes into proteins based on metabolic activity (Ong *et al.*, 2003; Beynon and Pratt, 2005; Jehmlich *et al.*, 2008b). The incorporation of heavy isotopes, for example,  $^{13}\text{C}$ , in proteins can be detected on peptide level using a protein-stable isotope probing (protein-SIP) approach (Jehmlich *et al.*, 2008b) and poses an additional way to unravel functionalities in relation to some specific processes such as hydrocarbon metabolism or degradation of organic xenobiotics. Diverse quantification methods were previously performed for pure cultures, but only label-free and metabolic labeling approaches were adopted for metaproteomic studies, and thus we will focus on them in the following sections.

### Label-free approaches

Label-free quantitation has been frequently used to study natural communities and reliable results depend on standardized and reproducible liquid chromatography (LC) conditions (Neilson *et al.*, 2011). Quantitation was performed using either signal intensity (area under the curve, AUC) or the number of fragmentation spectra (spectral counts, SC) (for review, see Neilson *et al.* (2011)). The AUC method is based on the detection of peptide ion abundances at specific retention times within given detection limits (Podwojski *et al.*, 2010). In the SC quantitation, the numbers of identified fragment ion spectra per protein are counted, and calculation is based on the assumption that the number of fragment spectra for a peptide correlates with its abundance (Liu *et al.*, 2004). This approach may lead to an overestimation of proteins of high molecular weight and peptide intensities. However, SC is recently becoming a common method for quantitative metaproteomic studies, as it is easy to implement for one-dimensional and two-dimensional proteomic data sets: Delmotte *et al.* (2009) used SC to roughly estimate protein abundances of 2257 proteins from phyllosphere microbiota, and Morris *et al.* (2010) evaluated the abundance of 2273 membrane proteins from microbial and viral communities of South Atlantic surface waters. The quantitation allowed them to clearly allocate molecular functions to specific phylotypes, which would not have been possible based on solely metagenome data. An improved SC approach is the normalized spectral abundance factor (NSAF), which considers the length of the quantified protein (Zybailov *et al.*, 2007). The NSAF was used to analyze an acidophilic biofilm describing specific protein functions and ecological behaviors (Denef *et al.*, 2010; Mueller *et al.*, 2010; Jiao *et al.*, 2011). Jiao *et al.* (2011) revealed that the abundance of 20 proteins account for more than the half of the 1351 identified proteins. The relative abundance of 6296 proteins

from 28 biofilm samples were analyzed by hierarchical clustering and nonmetric multi dimensional scaling (Mueller *et al.*, 2010). The NSAF of proteins within large multivariate data sets were compared and changes in the proteome of high and low abundant bacteria were discussed regarding the inhabiting biogeochemical zones. Microbial populations within distinct layers of a lake were investigated by Lauro *et al.* (2011) to compare 1824 proteins regarding their spatial distribution, and Schneider *et al.* (2012) studied the functionalities and microbial dynamics of litter degradation. Again, quantitation was the key for understanding the contribution of community members to the observed systemic activity. Recently, NSAF values were used to compare the effectiveness of different methods for the extraction of soil proteins (Keiblinger *et al.*, 2012). Signal intensity measurements or AUC have been rarely used in metaproteomic studies, as they require highly reproducible chromatographic runs and are not very amenable to highly fractionated samples before LC-MS/MS (Neilson *et al.*, 2011). A metaproteomic study described the biostimulation of an uranium-reducing aquifer community (Callister *et al.*, 2010). They detected the shift of protein abundances by the AUC method in microbial consortia originating from different groundwater wells temporarily stimulated with a supply of acetate and were capable to allocate them to selected species. The effect of biostimulation of naphthalene-degrading enrichment cultures was described by Guazzaroni *et al.* (2013). Protein abundances of the cultures with or without a prior biostimulation showed a broad metabolic activity of the first one and specialized naphthalene degradation potential for the latter.

In the recent years, the identification and quantification of proteins in the gut microbiota became a hot topic in microbial ecology. The human gut microbiota was investigated starting by a 2-DE study (Klaassens *et al.*, 2007), where a time-resolved analysis of protein spot changes was applied. Further on, gut microbiota studies were performed using 1-DE or label-free proteomic approaches. At first, Verberkmoes *et al.* (2009b) published an analysis employing relative protein abundances (normalized spectral) for a comparison of two fecal samples and two distinct sequence databases. Cantarel *et al.* (2011) used the same samples to test several computational strategies to improve protein identification. A new iterative workflow for protein identification leads to an increase of Clusters of Orthologous Groups (COG) categories (Rooijers *et al.*, 2011). Special emphasis has been placed on the role of the gut microbiota for metabolic syndromes, inflammation and obesity. Therefore, feces of obese and lean adolescents were comparatively investigated with the results that a core proteome of 505 proteins with equal abundance was identified (counted by AUC) (Ferrer *et al.*, 2013). Differences in the gene and protein composition between healthy

and Crohn's disease twin pairs were studied by metagenomic and metaproteomic analysis (Erickson *et al.*, 2012). Besides protein identification and quantitation, clear differences between noninflammation and inflammation in the ileum were observed. Furthermore, the alteration of the microbiota during an antibiotic treatment was studied in a multi-omics approach (Pérez-Cobas *et al.*, 2012). Besides metagenomics, metatranscriptomics and metametabolomics, a comprehensive analysis of the metaproteome was performed. Microbial populations showed oscillatory changes during the therapy, indicating an alteration of the gut microbiota and the strong interaction with the host. Feces analyses revealed only limited insight to the interface between host and microbiota. Thus, two studies focussed on the mucosa-associated microbiota in human and rats (Li *et al.*, 2011; Haange *et al.*, 2012). Representative, Haange *et al.* (2012) showed phylogenetic and functional differences along the intestinal tract (mucosa and gut content) of rats. The above-mentioned studies showed that metaproteomics goes clearly beyond mere identification because it allows functional classification and abundance correlation with consortia members.

## Metabolic labeling

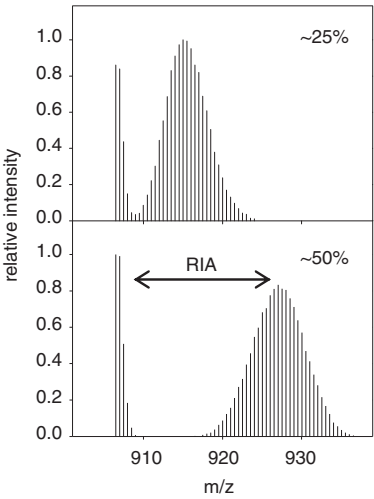
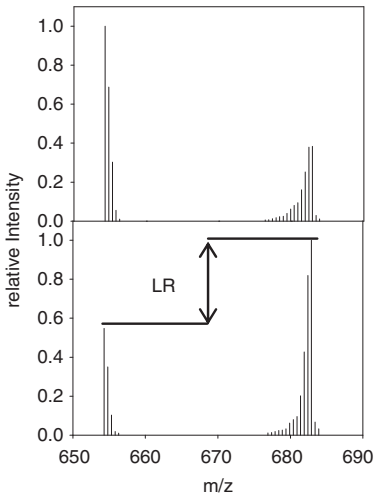
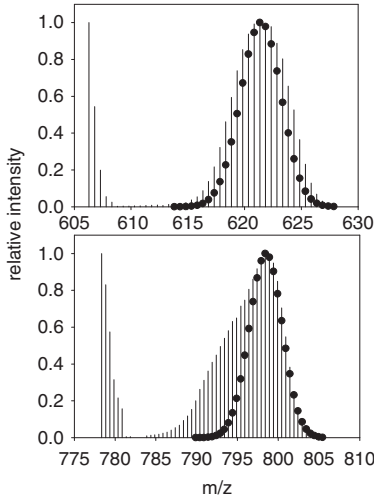
Besides label-free quantitation, protein abundances and expression patterns can be calculated by labeling strategies. Chemical labeling uses tags that are added after the cultivation; metabolic labeling uses the metabolic activity of the organisms to incorporate stable isotopes. In contrast to chemical labeling approaches, metabolic labeling techniques lead to the incorporation of the quantitation markers within the culture and thereby avoid artefacts from processing or chemical reactions. The most common approach for eukaryotes is stable isotope labeling with amino acids in cell culture (SILAC). There, isotope-labeled arginine and/or lysine are supplied and hence incorporated into the proteins during cultivation. After trypsin cleavage, they constitute the C-terminal amino acid and lead to at least one labeled amino acid per peptide. Combination of peptides and concomitant MS measurement allow the relative quantification of an unlabeled to a labeled sample in one measurement, hence avoiding any variability based on slight day-to-day differences in LC-MS/MS analysis. Despite being usually applied to organisms auxotrophic for arginine/lysine, the technique has also been used successfully for nematodes (Larance *et al.*, 2011) and bacteria such as *Staphylococcus aureus* (Schmidt *et al.*, 2010), which, although not being auxotrophic, were opportunistic enough to prefer the supplemented amino acids for peptide biosynthesis. Instead of amino acids, isotopically labeled substrates or nutrients can be applied that lead to a defined relative isotope abundance (RIA) in the biomass.

Peptides derived from these experiments can then be mixed with a control and quantified on the basis of peak intensities of labeled and non-labeled peptide signals. This method has been described using  $^{15}\text{N}$  (Snijders *et al.*, 2005; Pan *et al.*, 2011) and also  $^{36}\text{S}$  (Jehmlich *et al.*, 2012). The incorporation of stable isotopes into proteins by metabolic labeling can be used to quantify protein abundances and to observe changes in the microbial community structure. For static comparisons,  $^{15}\text{N}$  was used as an isotopic metabolic label to detect proteome changes of an acidophilic biofilm (Belnap *et al.*, 2010, 2011; Mueller *et al.*, 2011). The community of interest was grown in the presence of  $^{15}\text{N}$ -labeled medium to enable a  $>95\%$   $^{15}\text{N}$  labeling of the proteins. It was first used to prove that laboratory cultures are comparable to natural samples and further to show a differentiation of the metaproteome after cultivation of the biofilm in a pH range of 0.85–1.45 (Belnap *et al.*, 2010, 2011). The spatial and temporal succession of the biofilm was studied using the  $^{15}\text{N}$ -labeling approach (Mueller *et al.*, 2011). Beside the sole description of community structure, the analysis of functionalities allocated to the various members of the community is of particular interest. This requires more detailed information about isotope label distribution than accessible by mere quantitative metaproteomics. Protein-SIP is a method likely able to provide this additional information.

## Features of protein-SIP to simultaneously detect metabolic activity and phylogeny

Protein-SIP is based on the metabolic incorporation of isotopically labeled substrates, where  $^{13}\text{C}$ ,  $^{15}\text{N}$  or  $^{36}\text{S}$  can be used (Jehmlich *et al.*, 2008a, b, 2012). The metabolism of the labeled substrates leads to assimilation of heavy isotopes and subsequent incorporation into the various classes of biomolecules, including proteins and can be detected by mass spectrometry of peptides as reviewed in Seifert *et al.* (2012). The amount of atoms replaced by their heavy isotopes changes the natural isotope composition and is defined as the RIA. The RIA can be determined by LC-MS/MS analysis at the peptide level (Table 1). The calculation of the RIA is either done by analysis of the distribution of different isotope patterns of the peptides or based on features of the peptide mass such as the relation of the parent mass to the first two digits (Jehmlich *et al.*, 2008b; Fetzer *et al.*, 2010; Jehmlich *et al.*, 2010; Taubert *et al.*, 2011a). For the first approach, the amino-acid sequence of the peptide is needed to calculate the elemental composition and the theoretical mass, thus this approach requires suitable sequence information. The calculation of the RIA includes the whole isotope pattern, which allows a highly accurate calculation of the RIA; enabling differentiation of deviations down to 0.1% with only 10 peptides (Taubert *et al.*, 2011a). The second

**Table 1** Numerical parameters (RIA, LR) and features (shape) of protein-SIP posing carbon and nitrogen fluxes in microbial communities.

RIA	LR	Shape of isotope pattern
		
<p>Percentage of stable isotope incorporation Calculated by the extent of the mass shift of the peptide peak Reflects the carbon and nitrogen, respectively, metabolization Limit of detection: 1% <math>^{13}\text{C}</math> and <math>^{15}\text{N}</math></p> <p>References: Taubert <i>et al.</i>, 2011a; Taubert <i>et al.</i>, 2013</p>	<p>Ratio between the intensity of the isotope pattern of the labeled peptide and the total intensity of isotope patterns of unlabeled and labeled peptide Reflects protein synthesis and turnover in time-resolved analysis</p> <p>References: Taubert <i>et al.</i>, 2011b</p>	<p>Incorporation of <math>^{13}\text{C}</math> in proteins (&lt;99%) yields Gaussian distribution of the isotope patterns of a peptide Ideal Gaussian distribution indicates direct substrate utilization Isotope pattern with negative skewness indicates indirect <math>^{13}\text{C}</math> substrate metabolisation</p> <p>References: Morris <i>et al.</i>, 2012; Taubert <i>et al.</i>, 2012</p>

Abbreviations: LR, labeling ratio; protein-SIP, protein-based stable isotope probing; RIA, relative isotope abundance.

RIA: MS spectra of peptide GIFAVGDVATWPLHSGGK from the benzene 1,2-dioxygenase ferredoxin-NAD(+) reductase subunit protein (*P. putida*) with a RIA of ~25% and ~50%; LR: MS spectra of peptide EQAFVM(ox)NPEVK of elongation factor Ts at an early (above) and later time point (below); shape: peptide SDKIDIYSDR of methyl-coenzyme M reductase, beta subunit from *M. thermophila* with a RIA of 59% (above), peptide LAENGMMFDMLQR of methyl-coenzyme M reductase, gamma subunit from *M. marisnigri* with a RIA of 59% and clear left-tailing (below).

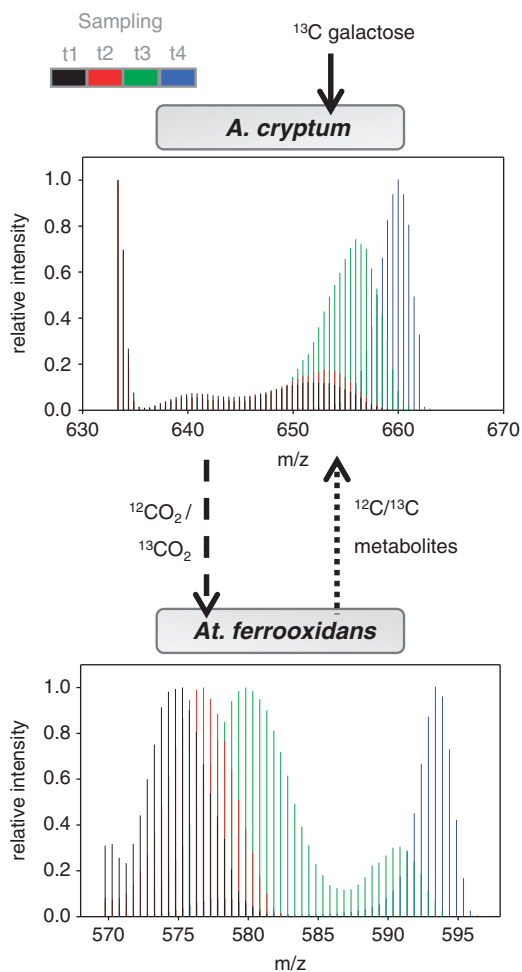
approach, called half decimal place rule, is sequence independent and the RIA calculation is based on the first two digits of the peptide mass. The major disadvantage of this method is the requirement of at least 40 peptides with equal RIA to achieve accuracy below 10% deviation (Jehlich *et al.*, 2010). Besides the amount of heavy isotopes within the molecules, the ratio between labeled and unlabeled molecules also contains valuable information. The 'labeling ratio' (LR) is based on the intensity ratio of the respective peptide MS signals (Table 1). The assessment of the LR allows, for example, the quantitation of protein biosynthesis as shown in a time series study by Taubert *et al.* (2011b), where protein induction by substrate shift was analyzed. Another valuable feature of protein-SIP is the possibility to detect differences in substrate usage by analysis of the shapes of isotope patterns in peptides. Direct metabolisation of isotopically labeled substrates leads to an isotope pattern closely resembling a normal distribution

(Table 1), whereas cross-feeding, that is, usage of labeled intermediates from degradation processes accumulating in culture, result in tailed distributions. Hence, the incorporation of isotopes into proteins of different groups of microorganisms can be used to track carbon fluxes within the studied consortia. The corresponding phylogenetic resolution depends on the available (meta-)genome sequences, the diversity of the phyla and the detection of valuable phylogenetic marker proteins such as chaperonin, ATPases and ribosomal proteins (Taubert *et al.*, 2012).

### Applications of protein-SIP in microbial ecology

The suitability of protein-SIP to detect key metabolic players within a community was first shown using a defined nitrate-reducing artificial mixed culture of the toluene-degrading *Aromatoleum aromaticum*

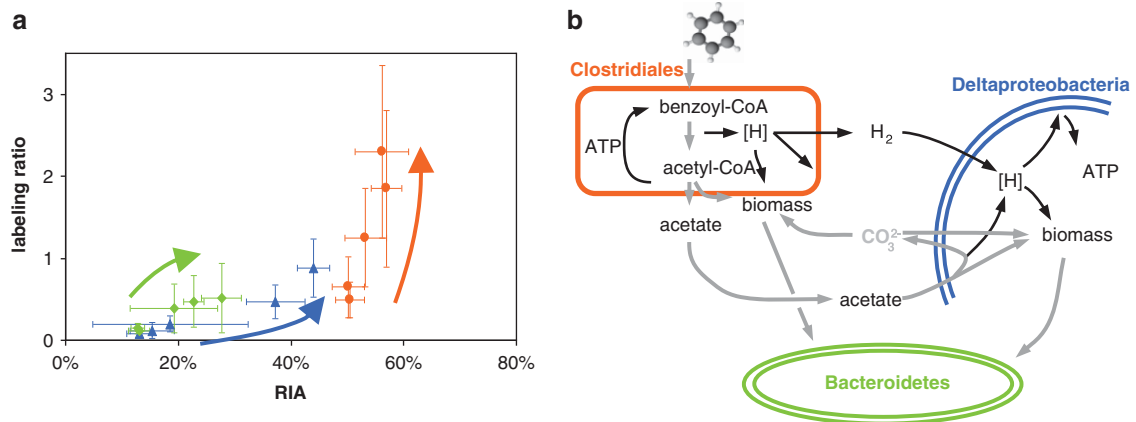




**Figure 1** Overview of a protein-SIP experiment studying the carbon flux between heterotrophic (*Acidiphilum cryptum*) and autotrophic (*Acidithiobacillus ferrooxidans*) organisms over four sampling points (t1–t4) (Kermer *et al.*, 2012). Labeled galactose is metabolized by *A. cryptum* to labeled CO<sub>2</sub>. In early time points, only a low amount of labeled CO<sub>2</sub> is present and is fixed by the autotrophic strain, leading to a low RIA. In later time points, the metabolic activity of *A. cryptum* resulted in a strong increase in <sup>13</sup>CO<sub>2</sub>, leading to a higher RIA in *A. ferrooxidans*. The <sup>13</sup>C-flux was reconstructed by the changes of the isotopologue patterns (color code, see sampling bar in the figure). Representative mass spectra showing the peptide ion mass distribution of peptide AGGLPAVIGELIR of dihydroxy-acid dehydratase (*A. cryptum*, gi148259108) and peptide AFDGSSSIAGWK of glutamine synthetase, type I (*A. ferrooxidans*, gi198282766).

EbN1 and a gluconate-degrading enrichment culture not able to use toluene as substrate (Jehmlich *et al.*, 2008b). The mixed culture was fed with <sup>13</sup>C-labeled toluene and <sup>12</sup>C-gluconate as carbon sources. Although all members of the consortium grew under these conditions, <sup>13</sup>C-labeled proteins were exclusively assigned to strain EbN1, demonstrating the suitability of protein-SIP to detect and identify <sup>13</sup>C-labeled proteins in a mixed culture. Protein-SIP was further established as a suitable method to trace the carbon flux between a heterotrophic and an autotrophic bacterial species by Kermer *et al.* (2012). In the course of the labeling experiments, the development of the RIA and the LR of peptides revealed that

a carbon transfer occurred when CO<sub>2</sub> formed by heterotrophic metabolism was used by the autotrophic species as a carbon source (Figure 1). The mass shifts of the peptides caused by the alternating incorporation of <sup>13</sup>C are depicted as colored isotopologue patterns in Figure 1. Protein-SIP applications in more complex cultures and environmental samples were conducted first in a study on a MTBE-degrading culture (Bastida *et al.*, 2010). The microbial community structure was analyzed after growth on <sup>13</sup>C-MTBE revealing that proteins related to members of the *Comamonadaceae* family were not <sup>13</sup>C-labeled, whereas proteins related to *Methylobium petroleiphilum* PM1 showed an average <sup>13</sup>C RIA of 94.5 atom% <sup>13</sup>C. In this study, the phylogenetic information retrieved from proteins correlated well with the structure of the microbial community established by 16S rRNA gene libraries. In further studies, a relationship between RNA-SIP and protein-SIP data was published by Jechalke *et al.* (2013). The active fraction of an aerobic benzene-degrading microbial community was identified as members of Beta- and Gammaproteobacteria with RIAs of about 90 atom% <sup>13</sup>C. The detection of <sup>13</sup>C-labeled proteins related to toluene dioxygenase and catechol 2,3-dioxygenase suggested benzene degradation by a dihydroxylation pathway with subsequent *meta*-cleavage of the catechol formed. Recently, protein-SIP was used to characterize the functionality of a sulfate-reducing community with *m*-xylene as the sole source of carbon and energy (Bozinovski *et al.*, 2012). Out of the 331 identified proteins, 23% were not <sup>13</sup>C-labeled, whereas 77% showed an incorporation of 19–22 atom% <sup>13</sup>C. <sup>13</sup>C-labeled proteins were involved in anaerobic *m*-xylene biodegradation, in sulfate reduction, in the Wood–Ljungdahl pathway and in general house-keeping processes. Thirty-eight percent of the labeled proteins were affiliated to Deltaproteobacteria, which suggests a key role of this group in the assimilation of *m*-xylene. Morris *et al.* (2012) studied the carbon flow in an anaerobic hydrocarbon degrading consortium. The methanogenic culture was incubated either with an oil-containing or an oil-free solid matrix and amended with either <sup>13</sup>C-labeled *n*-fatty acid or *n*-alkane. Proteins were identified with an *in silico*-created metagenome consisting of public genomes of sulfate-reducing species, fatty acid-metabolizing species and methanogenic archaea. Beside yielding different <sup>13</sup>C RIAs of the metabolic groups, depending on the used substrate and solid matrix, protein-SIP allowed a differentiation of direct and indirect carbon utilization based on peptides' isotope patterns (Table 1). The authors posed that peptides belonging to *Methanoculleus marisnigri*, which metabolizes CO<sub>2</sub> and other C1 substrates, showed patterns typical for consumption of degradation products of a labeled substrate, while peptides from *Methanosaeta thermophila*, an acetoclastic methanogen, showed distributions resembling a direct usage of the labeled substrates. The results indicated that methanogenesis



**Figure 2** Data evaluation of a protein-SIP experiment to detect the carbon flux in an anaerobic benzene-degrading community (Taubert *et al.*, 2012). (a) Protein grouping based on RIA vs LR to define metabolic groups of bacterial organisms obtained by <sup>13</sup>C benzene metabolisation. Colored data points indicate different metabolic groups: group 1, Clostridiales (orange dots); group 2, Deltaproteobacteria (blue triangles); group 3, Bacteroidetes (green squares). Arrows indicate the chronological development between the RIA and LR. A low variability of the RIA and a strong increase in LR (as shown for group 1) indicates a direct metabolisation of the <sup>13</sup>C-labeled substrate and a high growth rate. A stronger change of RIA over time indicates an indirect metabolisation (like shown for group 2 and 3). (b) Model of the carbon flux with the bacterial community based on data of the protein-SIP experiment using <sup>13</sup>C-labeled benzene and <sup>13</sup>C-labeled carbonate.

in this consortium was driven by acetate metabolisation and only secondary by hydrogenotrophic CO<sub>2</sub> reduction. The carbon flow in a sulfate-reducing, benzene-degrading community was recently studied (Taubert *et al.*, 2012). The anaerobic aquifer community was set up either using <sup>13</sup>C-labeled benzene and <sup>12</sup>C carbonate or <sup>12</sup>C benzene and <sup>13</sup>C-labeled carbonate. Up to five samples were investigated over 300 days per experiment. According to the development of the RIA and the LR over time, three distinct groups of proteins could be defined, based on MS identifications from metagenomic sequences, thus underlining the importance of following a time course in protein-SIP experiments (Figure 2). Phylogenetic affiliation was performed by BLASTP-based annotation of the respective metagenome contigs. The metagenome was employed to identify proteins and the labeling experiment resulted in a reconstruction of the carbon flow in the microbial community. Species belonging to Clostridiales were defined as the primary benzene consumers and as heterotrophic CO<sub>2</sub>-fixing bacteria. <sup>13</sup>C-labeled CO<sub>2</sub> was also fixed by Deltaproteobacteria members and additionally they used <sup>13</sup>C-labeled metabolites released by the Clostridiales. <sup>13</sup>C-labeled secondary metabolites or dead biomass of both groups were probably used as the carbon sources of putative scavengers belonging to Bacteroidetes/Chlorobi organisms defined as the third metabolic group in the consortium (Figure 2b).

Overall, protein-SIP is capable of providing insights into microbial consortia beyond the mere identification and quantification of proteins. Protein-SIP can (i) accurately detect metabolic activity and thereby distinguish between key players and bystanders in a microbial community, (ii) differentiate between the type of substrate usage and (iii) deliver information on protein induction. Further applications of protein-SIP may allow the

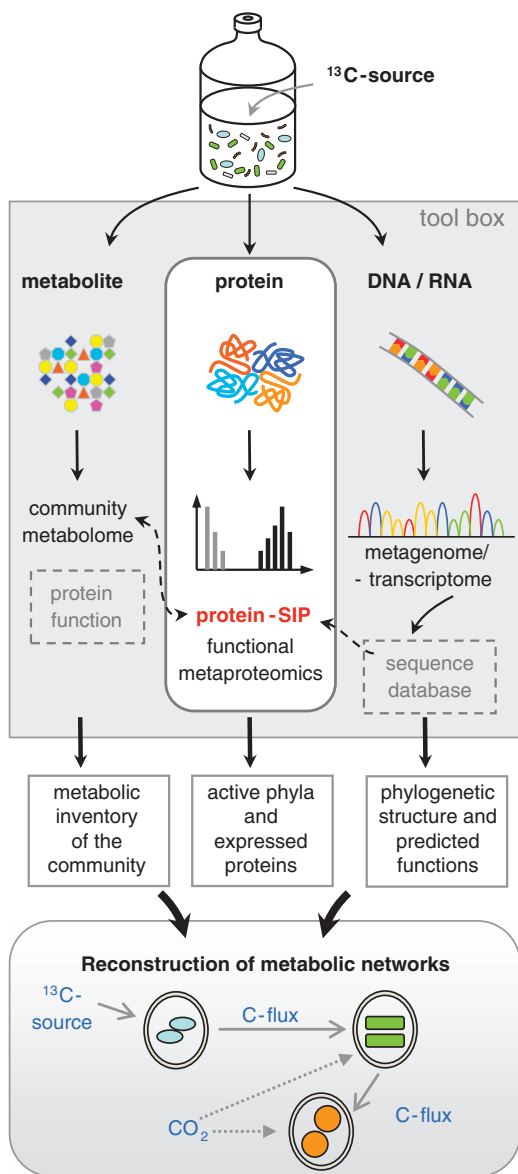
deciphering of the main pathways and key players involved in the biodegradation of newly appearing pollutants, in the global carbon and nitrogen cycling in aquatic and terrestrial ecosystems and in the interaction between human and their microbiota.

### Future development: automatic calculation of isotope labeling

To increase the throughput of mass spectra analysis for protein-SIP, an automated calculation of RIA and LR is needed. Currently, one automated pipeline is available and its application has been shown for <sup>15</sup>N-labeled proteins by transforming the whole *in silico* mass spectra from the sample-specific metagenome with increments of 1% in <sup>15</sup>N RIA to identify species-specific migration patterns from established into regrowing communities of an acid mine drainage biofilm (Pan *et al.*, 2011). Because of the stepwise transformation of a huge data set, the process required a relatively high amount of processing time in respect to more complex samples. In addition, a higher resolution with RIAs of 0.1% would even need more computational effort using the described bioinformatic pipeline. One alternative that has to be developed could be an automated tool, which identifies <sup>13</sup>C- or <sup>15</sup>N-labeled masses in correlation to the unlabeled ones within a certain retention time window and determines peptide sequences, RIAs and LR in one step.

### Integration of protein-SIP in the toolbox of microbial ecology and conclusion

The toolbox of microbial ecology comprises techniques that either use DNA, RNA, proteins and



**Figure 3** Integration of protein-SIP in the toolbox of microbial ecology and the combination of OMICS to reconstruct metabolic networks.

metabolites extracted from the whole community, or analyze intact single cells. The combination of SIP with these techniques leads to a holistic view into the metabolic networks (Figure 3). Single-cell analysis with nano secondary ion mass spectrometry is the most sensitive method to detect stable isotope incorporation. However, the gain of information about the phylogeny and function of the single cells depends highly on the used probes, and high-throughput experiments are not applicable due to the required machine time (Musat *et al.*, 2012). This problem might have been solved by the development of Chip-SIP (Mayali *et al.*, 2012). Raman microspectroscopy has a high spatial resolution but requires a label of at least 10 atom% for spectral shifts for identification of

isotope enrichment in labeling experiments (Huang *et al.*, 2007).

Within the SIP toolbox, protein-SIP is integrated as an interface between the nucleic acids-based techniques that provide the sequence database for MS analysis and the analysis of metabolites that indicate protein activities (Figure 3). In comparison to other SIP approaches three important features of protein-SIP are noteworthy: (i) minimal required incorporation of heavy isotope (about 0.1% of additional  $^{13}\text{C}$  incorporation) compared with DNA/RNA-SIP (50/10 atom% (Manfield *et al.*, 2002; Radajewski *et al.*, 2003)), (ii) dynamic range for quantification of isotope composition (two orders of magnitude) and (iii) high resolution of species and function (depending on the database) due to the fact that proteomic studies are not dependent on the amplification of target genes. The combination of multi-omics approaches to quantify metabolic groups and by using meta-metabolome-SIP would lead into a comprehensive tool to track the nutrient fluxes within microbial communities.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgements

We acknowledge the financial support provided by the European Community Project MAGICPAH (FP7-KBBE-2009-245226), the Deutsche Forschungsgemeinschaft (SPP1319, FOR 1530), BMBF IFB Adiposity Disease and by the Helmholtz Impulse and Networking Fund through Helmholtz Interdisciplinary Graduate School for Environmental Research (HIGRADE).

## References

- Bastida F, Rosell M, Franchini AG, Seifert J, Finsterbusch S, Jehmlich N *et al.* (2010). Elucidating MTBE degradation in a mixed consortium using a multidisciplinary approach. *FEMS Microbiol Ecol* **73**: 370–384.
- Belnap CP, Pan C, Denev VJ, Samatova NF, Hettich RL, Banfield JF. (2011). Quantitative proteomic analyses of the response of acidophilic microbial communities to different pH conditions. *ISMEJ* **5**: 1152–1161.
- Belnap CP, Pan C, VerBerkmoes NC, Power ME, Samatova NF, Carver RL *et al.* (2010). Cultivation and quantitative proteomic analyses of acidophilic microbial communities. *ISMEJ* **4**: 520–530.
- Beynon RJ, Pratt JM. (2005). Metabolic labeling of proteins for proteomics. *Mol Cell Proteomics* **4**: 857–872.
- Bozinovski D, Herrmann S, Richnow HH, von Bergen M, Seifert J, Vogt C. (2012). Functional analysis of an anaerobic *m*-xylene-degrading enrichment culture using protein-based stable isotope probing. *FEMS Microbiol Ecol* **81**: 134–144.
- Callister SJ, Wilkins MJ, Nicora CD, Williams KH, Banfield JF, Verberkmoes NC *et al.* (2010). Analysis of biostimulated microbial communities from two field experiments reveals temporal and spatial differences in proteome profiles. *Environ Sci Technol* **44**: 8897–8903.



- Cantarel BL, Erickson AR, VerBerkmoes NC, Erickson BK, Carey PA, Pan C *et al.* (2011). Strategies for metagenomic-guided whole-community proteomics of complex microbial environments. *PLoS ONE* **6**: e27173.
- Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R *et al.* (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci USA* **106**: 16428–16433.
- Denef VJ, Kalnejais LH, Mueller RS, Wilmes P, Baker BJ, Thomas BC *et al.* (2010). Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc Natl Acad Sci USA* **107**: 2383–2390.
- Erickson AR, Cantarel BL, Lamendella R, Darzi Y, Mongodin EF, Pan C *et al.* (2012). Integrated metagenomics/metaproteomics reveals human host-microbiota signatures of Crohn's disease. *PLoS ONE* **7**: e49138.
- Ferrer M, Ruiz A, Lanza F, Haange SB, Oberbach A, Till H *et al.* (2013). Microbiota from the distal guts of lean and obese adolescents exhibit partial functional redundancy besides clear differences in community structure. *Environ Microbiol* **15**: 211–226.
- Fetzer I, Jehmlich N, Vogt C, Richnow HH, Seifert J, Harms H *et al.* (2010). Calculation of partial isotope incorporation into peptides measured by mass spectrometry. *BMC Res Notes* **3**: 178.
- Guazzaroni ME, Herbst FA, Lores I, Tamames J, Pelaez AI, Lopez-Cortes N *et al.* (2013). Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation. *ISMEJ* **7**: 122–136.
- Haange SB, Oberbach A, Schlichting N, Hugenholtz F, Smidt H, von Bergen M *et al.* (2012). Metaproteome analysis and molecular genetics of rat intestinal microbiota reveals section and localization resolved species distribution and enzymatic functionalities. *J Proteome Res* **11**: 5406–5417.
- Huang WE, Stoecker K, Griffiths R, Newbold L, Daims H, Whiteley AS *et al.* (2007). Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environ Microbiol* **9**: 1878–1889.
- Jechalke S, Franchini AG, Bastida F, Bombach P, Rosell M, Seifert J *et al.* (2013). Analysis of structure, function and activity of a benzene degrading microbial community. *FEMS Microbiol Ecol* (doi:10.1111/1574-6941.12090).
- Jehmlich N, Fetzer I, Seifert J, Mattow J, Vogt C, Harms H *et al.* (2010). Decimal place slope, a fast and precise method for quantifying  $^{13}\text{C}$  incorporation levels for detecting the metabolic activity of microbial species. *Mol Cell Proteomics* **9**: 1221–1227.
- Jehmlich N, Kopinke FD, Lenhard S, Herbst FA, Seifert J, Lissner U *et al.* (2012). Sulphur- $^{36}\text{S}$  stable isotope labeling of amino acids for quantification (SULAQ). *Proteomics* **12**: 37–42.
- Jehmlich N, Schmidt F, Hartwich M, von Bergen M, Richnow HH, Vogt C. (2008a). Incorporation of carbon and nitrogen atoms into proteins measured by protein-based stable isotope probing (Protein-SIP). *Rapid Commun Mass Spectrom* **22**: 2889–2897.
- Jehmlich N, Schmidt F, von Bergen M, Richnow HH, Vogt C. (2008b). Protein-based stable isotope probing (Protein-SIP) reveals active species within anoxic mixed cultures. *ISMEJ* **2**: 1122–1133.
- Jiao Y, D'Haeseleer P, Dill BD, Shah M, Verberkmoes NC, Hettich RL *et al.* (2011). Identification of biofilm matrix-associated proteins from an Acid mine drainage microbial community. *Appl Environ Microbiol* **77**: 5230–5237.
- Keiblinger KM, Wilharitz IC, Schneider T, Roschitzki B, Schmid E, Eberl L *et al.* (2012). Soil metaproteomics—Comparative evaluation of protein extraction protocols. *Soil Biol Biochem* **54**: 14–24.
- Keller M, Hettich R. (2009). Environmental proteomics: a paradigm shift in characterizing microbial activities at the molecular level. *Microbiol Mol Biol Rev* **73**: 62–70.
- Kermer R, Hedrich S, Taubert M, Baumann S, Schlömann M, Johnson DB *et al.* (2012). Elucidation of carbon transfer in a mixed culture of *Acidiphilium cryptum* and *Acidithiobacillus ferrooxidans* using protein-based stable isotope probing. *JMIOMICS* **2**: 37–45.
- Klaassens ES, de Vos WM, Vaughan EE. (2007). Metaproteomics approach to study the functionality of the microbiota in the human infant gastrointestinal tract. *Appl Environ Microbiol* **73**: 1388–1392.
- Larance M, Bailly AP, Pourkarimi E, Hay RT, Buchanan G, Coulthurst S *et al.* (2011). Stable-isotope labeling with amino acids in nematodes. *Nat Methods* **8**: 849–851.
- Lauro FM, DeMaere MZ, Yau S, Brown MV, Ng C, Wilkins D *et al.* (2011). An integrative study of a meromictic lake ecosystem in Antarctica. *ISMEJ* **5**: 879–895.
- Li X, LeBlanc J, Truong A, Vuthoori R, Chen SS, Lustgarten JL *et al.* (2011). A metaproteomic approach to study human-microbial ecosystems at the mucosal luminal interface. *PLoS ONE* **6**: e26542.
- Liu H, Sadygov RG, Yates JR 3rd. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* **76**: 4193–41201.
- Manefield M, Whiteley AS, Ostle N, Ineson P, Bailey MJ. (2002). Technical considerations for RNA-based stable isotope probing: an approach to associating microbial diversity with microbial community function. *Rapid Commun Mass Spectrom* **16**: 2179–2183.
- Mayali X, Weber PK, Brodie EL, Mabery S, Hoepflich PD, Pett-Ridge J. (2012). High-throughput isotopic analysis of RNA microarrays to quantify microbial resource use. *ISMEJ* **6**: 1210–1221.
- Morris BEL, Herbst FA, Bastida F, Seifert J, von Bergen M, Richnow HH *et al.* (2012). Microbial interactions during residual oil and n-fatty acid metabolism by a methanogenic consortium. *Environ Microbiol Rep* **4**: 297–306.
- Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS, Rocap G. (2010). Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISMEJ* **4**: 673–685.
- Mueller RS, Denef VJ, Kalnejais LH, Suttle KB, Thomas BC, Wilmes P *et al.* (2010). Ecological distribution and population physiology defined by proteomics in a natural microbial community. *Mol Syst Biol* **6**: 374.
- Mueller RS, Dill BD, Pan C, Belnap CP, Thomas BC, Verberkmoes NC *et al.* (2011). Proteome changes in the initial bacterial colonist during ecological succession in an acid mine drainage biofilm community. *Environ Microbiol* **13**: 2279–2292.
- Musat N, Foster R, Vagner T, Adam B, Kuypers MMM. (2012). Detecting metabolic activities in single cells, with emphasis on nanoSIMS. *FEMS Microbiol Rev* **36**: 486–511.



- Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G *et al.* (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* **11**: 535–553.
- Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR *et al.* (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Mol Cell Proteomics* **4**: 1487–1502.
- Ong SE, Foster LJ, Mann M. (2003). Mass spectrometric-based approaches in quantitative proteomics. *Methods* **29**: 124–130.
- Ong SE, Mann M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nat Chem Biol* **1**: 252–262.
- Pan C, Fischer CR, Hyatt D, Bowen BP, Hettich RL, Banfield JF. (2011). Quantitative tracking of isotope flows in proteomes of microbial communities. *Mol Cell Proteomics* **10**: 006049.
- Peréz-Cobas AE, Gosalbes MJ, Friedrichs A, Knecht H, Artacho A, Eismann K *et al.* (2012). Gut microbiota disturbance during antibiotic therapy: a multi-omic approach. *Gut*, Doi:10.1136/gutjnl-2012-303184 (in press).
- Podwojski K, Eisenacher M, Kohl M, Turewicz M, Meyer HE, Rahnenfuhrer J *et al.* (2010). Peek a peak: a glance at statistics for quantitative label-free proteomics. *Expert Rev Proteomics* **7**: 249–261.
- Radajewski S, McDonald IR, Murrell JC. (2003). Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Curr Opin Biotechnol* **14**: 296–302.
- Rooijers K, Kolmeder C, Juste C, Dore J, de Been M, Boeren S *et al.* (2011). An iterative workflow for mining the human intestinal metaproteome. *BMC Genom* **12**: 6.
- Schmidt F, Scharf SS, Hildebrandt P, Burian M, Bernhardt J, Dhople V *et al.* (2010). Time-resolved quantitative proteome profiling of host-pathogen interactions: the response of *Staphylococcus aureus* RN1HG to internalisation by human airway epithelial cells. *Proteomics* **10**: 2801–2811.
- Schneider T, Keiblinger KM, Schmid E, Sterflinger-Gleixner K, Ellersdorfer G, Roschitzki B *et al.* (2012). Who is who in litter decomposition? Metaproteomics reveals major microbial players and their biogeochemical functions. *ISMEJ* **6**: 1749–1762.
- Schneider T, Riedel K. (2010). Environmental proteomics: analysis of structure and function of microbial communities. *Proteomics* **10**: 785–798.
- Seifert J, Taubert M, Jehmlich N, Schmidt F, Volker U, Vogt C *et al.* (2012). Protein-based stable isotope probing (protein-SIP) in functional metaproteomics. *Mass Spectrom Rev* **31**: 683–697.
- Snijders APL, de Koning B, Wright PC. (2005). Perturbation and interpretation of nitrogen isotope distribution patterns in proteomics. *J Proteome Res* **4**: 2185–2191.
- Taubert M, Baumann S, von Bergen M, Seifert J. (2011a). Exploring the limits of robust detection of incorporation of <sup>13</sup>C by mass spectrometry in protein-based stable isotope probing (protein-SIP). *Anal Bioanal Chem* **401**: 1975–1982.
- Taubert M, Jehmlich N, Vogt C, Richnow HH, Schmidt F, von Bergen M *et al.* (2011b). Time resolved protein-based stable isotope probing (Protein-SIP) analysis allows quantification of induced proteins in substrate shift experiments. *Proteomics* **11**: 2265–2274.
- Taubert M, Vogt C, Wubet T, Kleinstaub S, Tarkka MT, Harms H *et al.* (2012). Protein-SIP enables time-resolved analysis of the carbon flux in a sulfate-reducing, benzene-degrading microbial consortium. *ISMEJ* **6**: 2291–2301.
- Taubert M, von Bergen M, Seifert J. (2013). Limitations in detection of <sup>15</sup>N incorporation by mass spectrometry in protein-based stable isotope probing (protein-SIP). *Anal Bioanal Chem* **405**: 3989–3996.
- Thompson A, Schafer J, Kuhn K, Kienle S, Schwarz J, Schmidt G *et al.* (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **75**: 1895–1904.
- Verberkmoes NC, Denev VJ, Hettich RL, Banfield JF. (2009a). Systems biology: Functional analysis of natural microbial consortia using community proteomics. *Nature Rev Microbiol* **7**: 196–205.
- Verberkmoes NC, Russell AL, Shah M, Godzik A, Rosenquist M, Halfvarson J *et al.* (2009b). Shotgun metaproteomics of the human distal gut microbiota. *ISMEJ* **3**: 179–189.
- Wilmes P, Bond PL. (2009). Microbial community proteomics: elucidating the catalysts and metabolic mechanisms that drive the Earth's biogeochemical cycles. *Curr Opin Microbiol* **12**: 310–317.
- Zybailov BL, Florens L, Washburn MP. (2007). Quantitative shotgun proteomics using a protease with broad specificity and normalized spectral abundance factors. *Mol Biosyst* **3**: 354–360.