This document is the accepted manuscript version of the following article:
Achermann, S., Mansfeldt, C. B., Müller, M., Johnson, D. R., & Fenner, K. (2019). Relating Metatranscriptomic Profiles to the Micropollutant Biotransformation Potential of Complex Microbial Communities. Environmental Science and Technology. https://doi.org/10.1021/acs.est.9b05421

Relating Metatranscriptomic Profiles to the Micropollutant Biotransformation Potential of Complex Microbial Communities

Stefan Achermann,[1,2] Cresten B. Mansfeldt,[1] Marcel Müller,[1,3] David R. Johnson,[1] Kathrin Fenner*[,1,2,4]

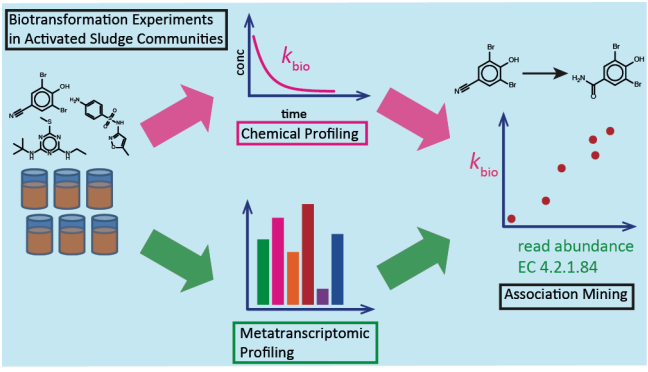
[1]Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600 Dübendorf, Switzerland. [2]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, 8092 Zürich, Switzerland. [3]Institute of Atmospheric and Climate Science, ETH Zürich, 8092 Zürich, Switzerland. [4]Department of Chemistry, University of Zürich, 8057 Zürich, Switzerland.

*Corresponding author (email: kathrin.fenner@eawag.ch)

S.A and C.B.M contributed equally to this work.

Abstract

Biotransformation of chemical contaminants is of importance in various natural and engineered systems. However, in complex microbial communities and with chemical contaminants at low concentrations, our current understanding of biotransformation at the level of enzyme-chemical interactions is limited. Here, we explored an approach to identify associations between micropollutant biotransformation and specific gene products in complex microbial communities, using association mining between chemical and metatranscriptomic data obtained from experiments with activated sludge grown at different solids retention times. We successfully demonstrate proportional relationships between the measured rate constants and associated gene transcripts for nitrification as a major community function, but also for the biotransformation of two nitrile-containing micropollutants (bromoxynil and acetamiprid) and transcripts of nitrile hydratases, a class of enzymes that we experimentally confirmed to produce the detected amide transformation products. Since these results suggest that metatranscriptomic information can indeed be quantitatively correlated with low abundant community functions such as micropollutant biotransformation in complex microbial communities, we proceeded to explore the potential of association mining to highlight enzymes likely involved in catalyzing less well-understood micropollutant biotransformation reactions. Specifically, we use the cases of nitrile hydration and oxidative biotransformation reactions to show that the consideration of additional experimental evidence (such as information on biotransformation pathways) increases the likelihood of detecting plausible novel enzyme-chemical relationships. Finally, we identify a cluster of mono- and dioxygenase fourth-level enzyme classes that most strongly correlate with oxidative micropollutant biotransformation reactions in activated sludge.



Introduction

Biotransformation by microbial communities in natural and engineered systems serves to reduce chemical contaminant loads in the environment.[1, 2] Considerable efforts have been undertaken to better understand the influence of various environmental parameters (e.g., temperature[3], redox conditions[4] or pH[5]) or operational parameters of wastewater treatment facilities[6] (e.g., solids or hydraulic retention times) on the biotransformation capacity of involved microbial communities. However, at the mechanistic level of interactions between enzymes and chemicals, our current knowledge remains limited, leaving major parts of observed variability in the extent of biotransformation across study conditions unexplained.[7] A more in-depth understanding of the causative agents of contaminant biotransformation (i.e., specific bacterial strains or enzymes) would not only support efforts to rationalize the influence of individual parameters, but, more generally, help to develop tools able to predict biotransformation pathways and half-lives, which is of interest for environmental risk assessment.[8]

Conventionally, linkages between specific bacterial strains, genes or gene products and chemical contaminant biotransformation are established in pure or enriched cultures with the chemical in question serving as sole growth substrate (e.g., refs [9, 10]). However, increasing evidence indicates that such experiments are often not a good model of the ability of complex environmental communities to biotransform contaminants present at low substrate concentrations (i.e., micropollutants).[11, 12] Other approaches such as stable isotope probing (SIP)[13, 14] or microautoradiography coupled with fluorescence *in situ* hybridization (MAR-FISH)[15, 16] use isotope labeling to overcome the limitations of pure or enriched cultures. These methods allow probing for the microbial species in which labeled atoms are incorporated even in complex microbial communities. However, micropollutant concentrations are most likely often insufficient to meet maintenance demands of individual metabolizing cells.[17] The biotransformation of micropollutants is therefore assumed to result from a complex sequence of individual biotransformation reactions.[17-19] In that case, microorganisms that incorporate and metabolize residues from the original chemicals may be distinct from the ones responsible for the initial transformation reaction.[8] Therefore, data from experiments relying on label

incorporation may not always be helpful to identify the type of cells involved in the initial, rate-determining transformation reactions.

As an alternative approach to evaluate individual hypothesized relationships between genes or gene products and the biotransformation of micropollutants, enzyme inhibition[20] or reverse transcription quantitative PCR (RT-qPCR)[20, 21] have been applied. However, such targeted approaches are only feasible when there is robust knowledge on the enzyme(s) catalyzing the biotransformation of a specific contaminant. While such knowledge is indeed available for a number of legacy contaminants[22] and some plant protection products,[21] hardly any knowledge exists on the enzymes involved in the biotransformation of most other environmental contaminants, such as those contained in urban wastewater. Previously, Johnson *et al.*[23] suggested association mining between the rate constants of observed biotransformation reactions and meta-omics data as an untargeted approach to generate hypotheses about potential causal linkages between enzymes and micropollutant biotransformation. Specifically, they suggest using metatranscriptomic data to describe the active functions present across microbial communities, mainly for two reasons. First, in earlier work, metatranscriptomic data have repeatedly been shown to correlate with protein abundance levels,[24-26] and, second, as of yet, metatranscriptomic analysis has higher sensitivity for detecting low abundance gene transcripts when compared to the ability of metaproteomics to detect the expressed protein.[27] Application of metatranscriptomic analysis to better understand micropollutant biotransformation is largely unexplored, with only select studies providing examples for individual contaminants,[23, 28] likely because massively-parallel sequencing has only become more broadly accessible in the last few years.

In a recent study, Helbling *et al.*[29] were able to uncover a linear and proportional relationship between measured biotransformation rates and relative gene transcript abundances of a gene previously spiked into an activated sludge community, independent of whether the gene transcript abundances were quantified by metatranscriptomic analysis or RT-qPCR. Whereas these results demonstrated that gene transcript abundances of an exogenously added microbial function quantified using metatranscriptomic analysis could feasibly be used in quantitative correlation analysis and hence association mining, it remained to be shown whether the same was true for non-spiked, native

functions. Another potential limitation of association mining between gene transcript abundances and contaminant biotransformation rate constants, which has not been addressed in previous studies, is the potentially large number of false positives. If large numbers of candidate enzymes are considered across a (comparably small) number of microbial communities, then it becomes increasingly probable that strong correlations could emerge by chance, highlighting non-causal relationships.[23]

Recently, we described trends in rate constants and biotransformation pathways for 42 micropollutants along a gradient of solids retention time (SRT) comprising six activated sludge communities.[30] Notably, we found that trends along the SRT gradient varied but were rather consistent within groups of chemicals undergoing the same type of initial biotransformation reactions, suggesting that shared enzymes or enzyme systems that are similarly regulated catalyze biotransformation reactions within such groups. In parallel, we performed metatranscriptomic sequencing to allow for functional characterization of the microbial communities.[31] In the present study, we combine the information previously obtained from biological and chemical analyses to further explore the potential of association mining between metatranscriptomic and micropollutant biotransformation information to generate hypotheses about potential causal linkages between enzymes and micropollutant biotransformation. Specifically, we first ask whether metatranscriptomic information for gene transcripts encoding previously described reactions can indeed quantitatively predict the rate of the associated observed micropollutant biotransformation reaction in complex microbial communities. If this were the case, then it follows that association mining can plausibly be employed to identify enzyme candidates that catalyze other observed reactions. Given the large efforts and costs involved in generating both biotransformation and metatranscriptomic data for a large number of microbial communities, we then proceed to ask whether consideration of additional information on the type of enzymatic reaction can increase the probability of detecting plausible enzyme candidates and hence address the challenge of low sample sizes. Through these analyses, we aim to highlight the potential but also limitations of association mining to uncover causative agents of micropollutant biotransformation, and to thus provide a roadmap for its wider adoption.

Methods

Full details on reactor operation, analysis of chemicals, determination of rate constants and metatranscriptome analysis are provided elsewhere.[30, 31] In the following subsections, a summary of the applied methods is provided.

Cultivation of activated sludge at SRTs between 1 and 15 days

Six fully-automated 12-L sequencing batch reactors were operated in parallel at SRTs of 1, 3, 5, 7, 10 and 15 days and a hydraulic retention time of 12 hours. The reactors were inoculated with activated sludge from a municipal wastewater treatment plant (ARA Niederglatt, Switzerland) and operated with local wastewater. At two time points during reactor operation (48 days (Exp1) and 187 days (Exp2) after start-up of the reactors), biotransformation batch experiments were conducted with a broad selection of micropollutants. For the batch experiments, micropollutants were spiked into the activated sludge communities (to yield an initial batch concentration of 6 µg/L for each compound), and, for each batch, 11 (Exp1) or 9 (Exp2) samples for chemical analysis were collected over three days.[30] Activated sludge samples for metatranscriptome analysis were collected 5 hours after the start (i.e., addition of micropollutants) of both experiments and immediately frozen using liquid nitrogen.[31] The sampling time point was chosen to allow sufficient time for transcriptional responses due to the addition of micropollutants, yet to capture the functions that were active during the experimental period during which we observed biotransformation for all biotransformed compounds, including the comparably fast degrading ones. Parallel to both Exp1 and Exp2, chemical control experiments were conducted with autoclaved activated sludge and autoclaved activated sludge filtrate to estimate the degree of adsorption (to sludge solids) and abiotic transformation, respectively.

Micropollutant biotransformation rate constants

Samples for chemical analysis were measured using liquid chromatography coupled to high-resolution mass spectrometry (LC-HRMS). From the obtained concentration-time series, first-order rate constants (k_{obs}) were determined that were then normalized by total suspended solids concentrations (TSS) as proxies for biomass concentration in each reactor to obtain second-order biotransformation

rate constants (k_{bio}). Biomass-normalized rate constants were calculated for all micropollutants for which abiotic degradation was minor (<20%), no strong sorption to activated sludge solids was observed (<20%) and first-order degradation was observed.[30] In total, k_{bio}-values were obtained for 33 (Exp1) and 42 (Exp2) micropollutants and 5 (Exp1) and 6 (Exp2) different SRTs (because of the low biomass concentration in Exp1 for the reactor at 1 day SRT, this reactor was not considered when calculating k_{bio} values). Additionally, a suspect transformation product screening was conducted to identify the major initial biotransformation reactions these chemicals were undergoing (details on reaction type characterization are provided elsewhere).[30]

Metatranscriptomic analysis

The samples collected for metatranscriptomic analysis were stored at -80 °C until processing. For RNA extraction, a phenol-chloroform extraction method was used, followed by DNA removal (TURBO DNase Kit, Invitrogen) and purification (MoBIO RNA Pro Clean-Up Kit, MoBio).[31] Prior to sequencing, a ribosomal RNA (rRNA) removal step was conducted to enrich the messenger RNA (mRNA) fraction of the samples (RiboZero Epidemiology Kit, Illumina). Sequencing was performed on the Illumina NextSeq platform (Illumina TruSeq Single-End-Read 150 bp) and the raw data are freely available at EMBL-EBI (https://www.ebi.ac.uk) under the study number ERP024418. The raw reads (41.2-54.3 million reads per sample; because of similar sampling depths, the data was not rarefied) were trimmed and filtered and additional rRNA was removed using the software SortMeRNA.[32] Using the software DIAMOND[33] (minimum bitscore cutoff of 50), the sequences were annotated with the descriptors provided in the Enzyme Commission (EC) Number Uniprot database, resulting in 5.1-9.8 million reads being annotated per sample. The full Uniprot-TrEMBL database was downloaded on March 6th, 2018. The read counts were aggregated per EC number. The read counts for all EC numbers were normalized by the total number of reads identified to encode a protein to account for variability in annotation efficiency between samples. Because variability in annotation efficiency was found to be higher than variability in sampling depth, we preferred this normalization procedure to raw read rarefaction to minimize information loss. Descriptions for EC categories were obtained from the BRENDA database (http://www.brenda-enzymes.org/), downloaded on May 1st, 2018.

Additionally, the metatranscriptomic analysis assigned a putative taxonomic read origin based on the identity of the best matching sequence, allowing to estimate the fraction of reads originating from eukaryotic and bacterial organisms or, more specifically, from different taxonomic genera.

Correlation analysis between EC numbers and biotransformation rate constants

Pearson correlation coefficients (denoted r) were calculated using the software R (version: 3.3.0). Reported p-values of the respective correlations correspond to two-tailed tests. To account for multiple hypothesis testing (here, testing gene transcript read abundances of *n* different ECs against a given micropollutant rate constant), the Benjamini-Hochberg method was applied to control the false discovery rate and to obtain adjusted p-values, denoted P_{BH}.[34] Because of the explorative character of many analyses presented in this work, providing new hypotheses about linkages between genes and micropollutant biotransformation reactions that we propose to further assess in future work, we are not only concerned about false positive detection (type I error) but we also want to minimize false negative results (type II error). Therefore, the Benjamini-Hochberg procedure was preferred over more conservative methods. To statistically assess whether higher correlation coefficients (here defined as $r > 0.5$) were overrepresented in certain subsets of the correlation tables, bootstrapping was used. Therefore, sampling with replacement (n = 1000) was performed, and we tested whether the fraction of r values > 0.5 of a given distribution lies within the empirical 95% confidence interval of the sampled entity (corresponding two-tailed p-values are denoted P_B). Shifts in median correlation coefficients were assessed in the same way.

In the EC number classification scheme, enzymatic reaction types are typically defined at the 3rd level (sub-subclasses) of the four-digit EC numbers, whereas the 4th digit characterizes substrate specificity.[35, 36] We performed correlation analysis at the level of individual ECs (4th level EC numbers) because gene transcript abundances between individual ECs within each sub-subclass level class were not always strongly correlated. Correlation coefficients between gene transcript abundances and biotransformation rate constants were calculated for Exp1 (across 5 SRTs, n = 5) and Exp2 (across 6 SRTs, n = 6) separately. For chemicals that were only included in Exp2 (iprovalicarb, amisulprid,

irgarol, isoproturon, metoxuron, MMclB, BEclB and terbutryn), only one correlation coefficient was obtained. To construct a correlation heatmap, correlation coefficients were averaged if available from both experiments, and hierarchical clustering was performed using Euclidean distances and complete linkages (using the R package 'pheatmap', v1.0.10).

Unless stated otherwise, a relative abundance threshold of 10^{-6} was applied, and ECs were only considered when their relative abundance exceeded this threshold in at least one sample of both experiments. This threshold corresponds to a minimum of 12 to 22 reads annotated to the respective EC in individual samples. For oxidative biotransformation reactions, the results were tested for robustness against higher (10^{-5}) and lower (10^{-7}) thresholds (see results and Figure S6 in the Supporting Information (SI)).

Nitrile hydratase experiments

A mix of nine different nitrile hydratases was purchased as selectAZyme™ enzyme screening kit (Almac). Batch experiments were performed in 1.5-mL HPLC vials filled with the nitrile hydratase solutions (200 µL, 15 mg/mL of total enzyme concentration in 0.05 M phosphate buffer, blend of all nine nitrile hydratases or individual nitrile hydratases), micropollutant containing solution (50 µL, 1.5 mg/L of each contaminant in water) and phosphate buffer (750 µL, 0.1 M solution provided with the enzyme screening kit). The batch reactors (initial concentration of each micropollutant: 75 µg/L) were placed on a shaker table (220 rpm) in a temperature-controlled environment (30 °C). The experiment was performed in triplicates and for a runtime of 72 hours. Sample workup for chemical analysis using liquid chromatography coupled to mass spectrometry (LC-MS) was performed according to Polson *et al.*[37] In short, 200 µL of the reaction mixture were added to 200 µL trichloroacetic acid (20% in water), mixed by vortexing, incubated at 4 °C for 25 minutes and centrifuged (10 min, 1700 × g, 4 °C). In total, 200 µL of supernatant were transferred to 800 µL of water, internal standard was added and the samples were measured within 7 days as detailed elsewhere.[30]

Results and Discussion

Evaluation of data treatment procedure and validation of correlation analysis for nitrification

In total, gene transcripts assigned to 4165 different ECs were detected, of which 2760 ECs were found in at least one sample from both Exp1 and Exp2 with a relative abundance of at least 10^{-6}. Prior to the analysis of correlations between gene transcript read abundances of individual EC categories and contaminant biotransformation rate constants, we assessed the validity of (i) our metatranscriptome normalization procedure by comparison with ECs representing previously used reference genes, and (ii) the correlation analysis by testing it for a well-characterized metabolic function, i.e., nitrification.

Normalized gene transcript abundance. To normalize gene expression data obtained in other methods such as RT-qPCR, reference genes are frequently used.[38] When a reference gene represents a function that is present in all microorganisms contained in our activated sludge samples and that does not show much variability in expression levels amongst different microorganisms and conditions, a constant relative abundance can be expected in all samples if our normalization procedure is valid. Therefore, we calculated the relative standard deviation (RSD), defined as the standard deviation of the relative abundance over the mean relative abundance, for the EC categories representing specific, frequently used reference genes, namely DNA-directed RNA polymerase (RNAP, EC 2.7.7.6), DNA topoisomerase (5.99.1.2), DNA gyrase (5.99.1.3) and glyceraldehyde-3-phosphate dehydrogenase (GAPDH, 1.2.1.12), across all samples (n = 12) (Table S1).[38] RSDs thus calculated were found to range between 11−39% initially. Because the first three reference genes have primarily been used for bacteria, the respective fractions of eukaryotes and bacteria were additionally estimated for all samples (Figure S1). Based on the total number of reads assigned to EC categories, we observed an increase in gene transcripts that were predicted to be of eukaryotic origin towards higher SRTs (Exp1: from 11 to 40%, Exp2: from 19 to 38%), which is in accordance with previous reports of longer SRTs tending to promote the growth of higher forms of life and increased abundances of different protozoa species at higher SRTs.[39-42] Because fractions of the read abundances of the ECs 2.7.7.6, 5.99.1.2 and 5.99.1.3

233 originated from eukaryotes, we recalculated their RSDs for the bacterial fractions only (after
234 calculating the relative abundances of reads originating from bacteria), leading to smaller RSD values
235 in the range of 6 to 10% (Table S1). This low variability in abundance for the tested reference genes
236 provides empirical support for the here applied normalization procedure. For EC 1.2.1.12, the RSD
237 was also considerably smaller when the bacterial and eukaryotic fractions were considered separately
238 (RSD bacteria: 6%, RSD eukaryotes: 32%), but remained relatively high for eukaryotes. The latter
239 might be because of fluctuating expression levels in eukaryotic organisms under different conditions
240 including stress levels,[43] or it might result from the overall low relative abundance of eukaryotes at
241 low SRTs, which might lead to increased uncertainties when calculating RSD values for eukaryotes.

242 ***Correlation of amo gene transcripts with nitrifying activity.*** We selected nitrifying activity to test the
243 validity of the correlation analysis approach because (1) we had a strong hypothesis regarding the
244 main enzymes involved, (2) we measured nitrifying activity and observed a strong trend with SRT
245 across the samples, and, (3), as for the biotransformation rate constants, we could calculate a biomass-
246 normalized nitrifying activity. In wastewater treatment, a certain minimal SRT is known to be required
247 to achieve nitrification because ammonia oxidizing bacteria are slow-growing and are washed out at
248 low SRTs.[44] In the transformation of ammonium to nitrate, the initial step (the oxidation of
249 ammonium) is typically rate-limiting and performed by the enzyme ammonia monooxygenase (*amo*,
250 EC 1.14.99.39).[45] In agreement with our expectations, higher nitrifying activity was observed at higher
251 SRTs in our experiments[30] and the abundance of *amo* gene transcripts increased with increasing SRT.
252 Correlating the *amo* gene transcripts with the measured nitrifying activity resulted in correlation
253 coefficients of 0.97 and 0.96 (Pearson's r in Exp1 and Exp2, P <0.05; Figure 1). Furthermore, as
254 earlier observed for a microbial function added at different relative levels into a community,[29] the
255 intercepts obtained from linear regression analysis are not significantly different from zero (Figure 1),
256 supporting not only linearity but proportionality of the relationship. These results demonstrate that
257 metatranscriptomic information can indeed quantitatively predict the native relative activity levels of a
258 microbial community, at least with respect to a well-known metabolic function.

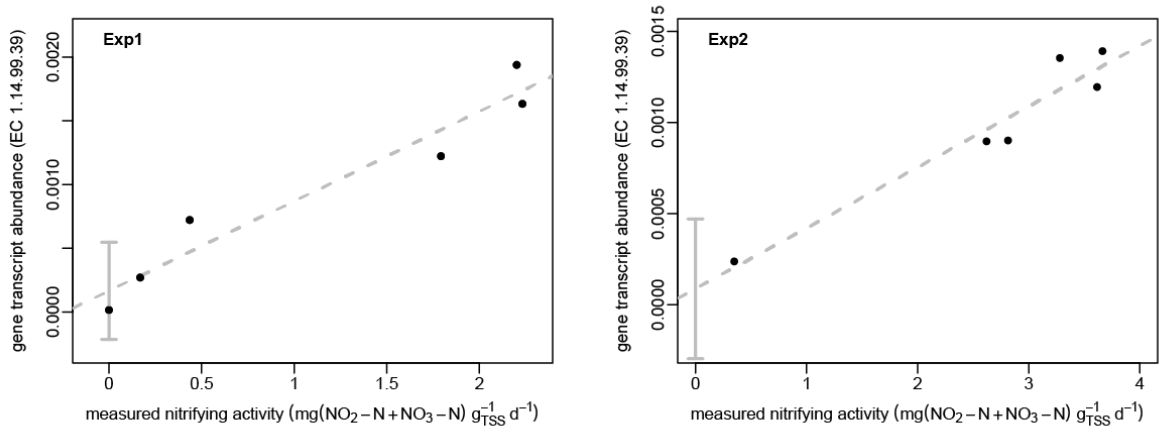


Figure 1. Relative gene transcript abundance of *amo* (EC 1.14.99.39) against the measured nitrifying activity in Exp1 and Exp2 (the applied methods for measuring nitrifying activity are described elsewhere[30]). Confidence intervals (confidence level: 95%) for the intercept (gray bar) were calculated after linear regression analysis (indicated by dashed line).

Validation of correlation analysis for micropollutant biotransformation

The validity of the correlation analysis approach for micropollutant biotransformation (a supposedly minor function of activated sludge communities) was assessed for three selected substance classes. They were selected because (i) we either had prior knowledge or, based on transformation product analysis, could generate a strong hypothesis on involved enzymes, and (ii) they had shown highly consistent within-substance-class patterns across the SRT gradient, suggesting that biotransformation for substances in these classes was catalyzed by a number of shared enzymes or enzyme systems that were similarly regulated[30].

Correlation analysis for the biotransformation of nitrile-containing compounds and confirmation of results by enzyme assays. In our experiments, bromoxynil and acetamiprid were both shown to react at the nitrile functional group, and, for both micropollutants, the corresponding primary amide transformation products were detected.[30] Nitrile hydratase (EC 4.2.1.84) has been previously described to catalyze the nucleophilic attack of nitriles by water to form the corresponding amide. Reassuringly, strong and significant correlations were obtained between the biotransformation rate constants of the two nitrile-containing compounds and abundances of transcripts annotated as nitrile hydratase, i.e., r=0.95, P<0.05 (Exp1), r=0.78, P>0.05 (Exp2) for bromoxynil, and r=0.91, P<0.05 (Exp1), r=0.95,

P<0.05 (Exp2) for acetamiprid. The lower correlation observed for bromoxynil in Exp2 (r = 0.78) likely originates from a larger uncertainty in the corresponding k_{bio} value (for an error estimation, see Figure S2). Furthermore, as for nitrification, the correlations with nitrile hydratase are in agreement with the proportionality assumption as shown in Figure S3. The proportional relationships between nitrile hydratase transcript abundances and biotransformation of two nitrile-containing compounds demonstrates that quantitative relationships can be uncovered not only for highly abundant community functions such as nitrification but also for biotransformation of low concentration chemicals (<6 µg/L) and for low relative gene transcript abundances (<0.001%) in a genuine complex community. To provide evidence for the causality of the thus demonstrated relationship, independent experiments were performed using commercially available nitrile hydratases. Incubation with a mix of nine nitrile hydratase enzymes for 72 h resulted in complete depletion for bromoxynil and a clear reduction in concentration by 47% (mean concentrations of triplicates) for acetamiprid (Figure S4). In parallel, formation of the corresponding amide transformation products could be observed. Although the transformation products could not be quantified due to a lack of analytical standards, increased peak areas at the expected *m/z* values were detected relative to control conditions when the nitriles were incubated either with the same mix of enzymes (Figure S4) or increasing concentrations of selected individual nitrile hydratases (Figure S5).

Correlation analysis for acetanilide biotransformation. Acetanilide biotransformations have repeatedly been associated with glutathione-S-transferase (*GST*, EC 2.5.1.18).[46, 47] Correlation analysis revealed positive, and for two compounds, significant, but, in comparison to the nitriles, weaker correlation coefficients (r: 0.3 – 0.95, P: 0.02−0.51; Figure S7). We hypothesize that mainly two reasons impede the direct correlation between gene transcript abundances and rate constants here. First, although the oxanilic acid (OXA) and ethanesulfonic acid (ESA) transformation products that were reported to form after an initial substitution by *GST* were detected, additional transformation products related to reductive dehalogenation were observed.[30] If multiple initial biotransformation reactions occur in parallel, separate rate constants for each pathway would be required to obtain meaningful linear correlations since the relative importance of parallel reactions might change across conditions. To properly quantify these, separate spike experiments with transformation product

standards would have been needed, which were outside the scope of this study. The second reason for the insignificant correlations in some of the cases is the relatively small variation in k_{bio} for the acetanilides ($k_{bio,max}/k_{bio,min}$ <3, except for flufenacet which actually showed the highest correlation coefficient) and also in *GST* gene transcript abundances (GST_{max}/GST_{min} <3) across all samples. Whereas the consistency of low variations in both biotransformation rate constants and gene transcript abundances actually supports the hypothesized proportionality, it likely impedes the detection of a significant correlation given that various experimental and analytical uncertainties introduce scatter into the data.

Correlation analysis for sulfonamide biotransformation. Relatively low variations in biotransformation rate constants were also observed for the five investigated sulfonamide antibiotics (mean $k_{bio,max}/k_{bio,min}$ of 2.6) and (except for sulfathiazole) negative associations with SRT were observed.[30] In a recent study, the biotransformation of sulfonamide antibiotics has been associated with folic acid synthesis and, more specifically, dihydropteroate synthase (*DHPS*, EC 2.5.1.15), which catalyzes one of the essential steps in folic acid synthesis.[48] However, correlations between sulfonamide biotransformation rate constants with transcripts of *DHPS* did not indicate a strong relationship (r ranging from -0.52 to 0.31). Again, this may be caused by a lack of sufficient variation between the different SRT conditions with respect to the function in question. Indeed, $DHPS_{max}/DHPS_{min}$ was 1.4 across all samples, consistent with the fact that any bacteria growing need to express *DHPS* to a certain degree.

Taken together, results for the three selected substance classes support the notion that quantitative relationships indeed exist between gene transcript abundances derived from metatranscriptomic information and specific micropollutant biotransformation functions in complex microbial communities. However, these only manifest themselves as strong and significant correlations if the function of interest shows sufficient variation across the microbial communities investigated. This was the case for nitrile hydrolysis (i.e., $k_{bio,max}/k_{bio,min}$ > 3, except for bromoxynil in Exp2), a function with low abundance in our metatranscriptomic data (i.e., relative gene transcripts ranging between $1–8\times10^{-6}$), and which so far has been reported to be associated with only a low number of bacterial species

(e.g., of the only 2937 annotated sequences available in the UNIPROT database, 1161 originate from the single bacterial class of *Rhizobiales*). Therefore, the capacity for nitrile hydrolysis can be expected to strongly depend on community composition. In contrast, functions such as acetanilide and sulfonamide biotransformation are most likely catalyzed by enzymes intimately linked to cellular function (i.e., stress response[49] and cellular metabolism and growth[50], in the case of *GST* and *DHPS*, respectively) and are therefore widely distributed across bacterial species. This seems to be confirmed by the orders of magnitude higher relative gene transcript abundances in our data compared to nitrile hydratase genes (i.e., gene transcripts range between $3–6\times10^{-4}$ and between $0.7–2\times10^{-3}$, for *GST* and *DHPS*, respectively) and explains why biomass normalization was sufficient to explain most of the variation in acetanilide and sulfonamide biotransformation rate constants. As a consequence, causal linkages to such more widely distributed functions do not lend themselves to be uncovered through correlation analysis.

Application of correlation analysis for association mining

Given the results discussed so far, we concluded that association mining can plausibly be employed to identify enzyme candidates that catalyze observed reactions in the much more abundant case where such knowledge is lacking, if the biotransformation of interest shows sufficient variation across the microbial communities studied. However, as pointed out previously, unless metatranscriptomic and biotransformation kinetics data are available for large numbers of different microbial communities, such an analysis runs the risk of generating many false positive associations.[23] Since such data are still very costly to generate, and, accordingly, our number of experimentally characterized communities, i.e., the six communities grown along a SRT gradient, was too small to prevent false positive associations, we used our data set to test whether the consideration of additional experimental information on the general enzymatic reaction type could increase the probability of detecting plausible enzyme candidates. In the following, we will first demonstrate this for the case of nitrile-containing compounds, and then apply the approach to identify specific fourth-level enzyme classes potentially involved in oxidative transformations.

In the case of nitrile-containing compounds, not only nitrile hydratase, but also nitrilases (sub-subclass EC 3.5.5), of which we detected transcripts for ECs 3.5.5.1, 3.5.5.4, 3.5.5.7 and 3.5.5.8, have previously been described to hydrolyze nitriles, yet to yield carboxylic acids as final products.[51] When considering all 5 ECs and accounting for multiple hypothesis testing (at the 95% confidence level, Benjamini-Hochberg method[34]), only one significant result was obtained (EC 4.2.1.84 in Exp2 for acetamiprid, Figure 2). In contrast, when only EC 4.2.1.84 was considered as justified using the additional evidence from transformation product analysis and literature, additional significant correlations in Exp1 emerged for both chemicals. This example illustrates how additional experimental evidence on the products of the enzymatic transformation can reduce the number of considered hypotheses and lead to more statistical power when searching for meaningful associations with gene transcript abundances.

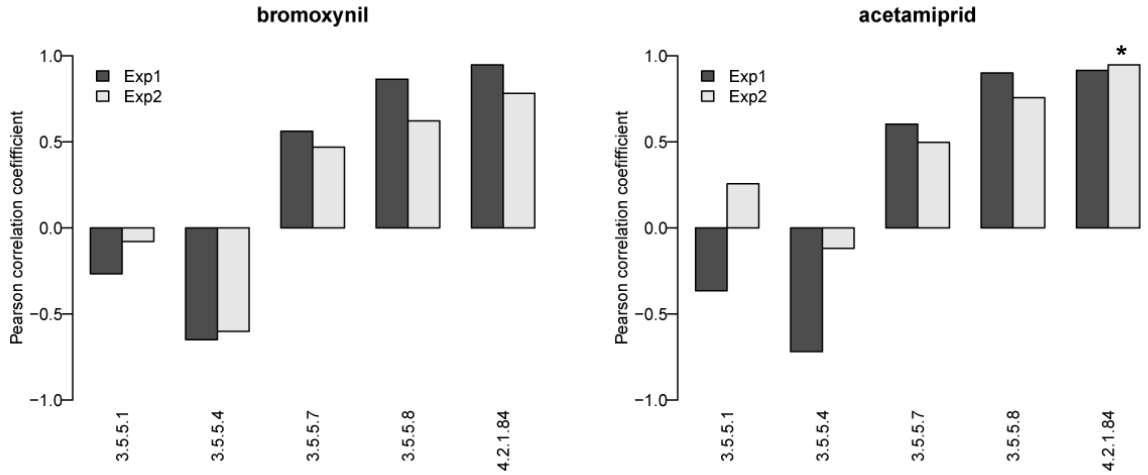


Figure 2. Pearson correlation coefficients for bromoxynil and acetamiprid with ECs describing nucleophilic reactions of water with nitriles. Asterisks indicate a significant correlation after correcting for multiple hypothesis testing (P_{BH} <0.05, n=5).

Association mining for compounds undergoing oxidative transformation reactions. In our experiments, for 19 out of 42 analyzed micropollutants, an initial oxidation reaction was confirmed by chemical analysis as detailed previously.[30] Most of these oxidative transformations displayed clear trends of increasing degradation with SRT and in more than two third of cases considerable variation (i.e., $k_{bio,max}/k_{bio,min} > 3$). [30] Consistently, the majority of the observed oxidation reactions, i.e.,

dealkylation, S-/N-oxidation and hydroxylation, are typically catalyzed by monooxygenases,[52] which are known to be rather rare (i.e., not widespread among different bacterial species) and highly differentially expressed.[53] For these reasons, we chose to use the case of oxidative micropollutant biotransformation as a case study to illustrate the potential, but also limitations of using association mining to detect enzyme candidates plausibly involved in catalyzing these oxidative transformations.

Differently from nitrile hydrolysis, a large number of ECs designated as monooxygenases may catalyze the different types of observed oxidation reactions. Also, certain dioxygenases[54, 55] and peroxygenases[56, 57] have been shown to catalyze the types of monooxygenation reactions observed. For three micropollutants, dihydroxylated transformation products were observed alongside dealkylated products, additionally suggesting a potential direct relevance of dioxygenases in the investigated experimental system.[30] We therefore searched for EC sub-subclasses associated with monooxygenases and dioxygenases and selected all 144 detected 4th level ECs contained therein (Table S2). No peroxygenase ECs (sub-subclass EC 1.11.2) were detected above the minimum relative abundance threshold.

In the following, we compared the results of association mining for the 19 micropollutants undergoing oxidation reactions, if their rate constants were correlated with gene transcript abundances across all 2760 ECs or only across those 144 4th level oxygenase ECs plausibly associated with the observed oxidative transformation reactions. The distribution of correlation coefficients resulting from correlation analysis against the selected oxygenases showed, on average, significantly higher values (median r = 0.47, fraction of r > 0.5: 40%, Figure S6b) than the distribution of correlation coefficients resulting from correlation analysis against all ECs (median r = 0.12, fraction of r > 0.5: 32%, Figure S6e) ($P_B < 0.05$). The same was also true when correlating gene transcript abundances of individual selections of sub-subclasses containing either only mono- or only dioxygenases against rate constants of oxidative transformation reactions (monooxygenases: median r = 0.45; dioxygenases: median r = 0.43) (Table S3, $P_B < 0.05$ relative to the all ECs case). This statistically significant overrepresentation of ECs showing strong correlations with oxidation reactions (see Table S3) within the group of oxygenase ECs, which was observed independent of the applied minimum abundance threshold

(Figure S6), supports the validity of restricting the EC search space for association mining based on the observed transformation reactions to increase the probability of detecting plausible enzyme candidates.

Finally, to allow for a more detailed inspection of individual correlations between ECs and oxidation reactions, we constructed a heatmap (Figure 3). Because all oxidation reactions (except for the micropollutant MMclB) showed increasing trends with increasing SRT, the variability across the micropollutants is relatively small compared to the differences observed among the analyzed ECs. We therefore applied hierarchical clustering and detected three clusters showing strong (cluster A, 98 ECs), moderate (B, 20 ECs) or mostly anti-correlations (C, 26 ECs). For several micropollutants, statistically significant correlations were observed with ECs of cluster A as marked in Figure 3. The selection of ECs in Figure 3 provides an opportunity for comparison with oxygenase-related ECs that have previously been associated with micropollutant biotransformation. Reassuringly, both *Amo* (EC 1.14.99.39)[20, 58-61] and vanillate monooxygenase (EC 1.14.13.82)[62], which both have both been linked to micropollutant biotransformation earlier, are found in cluster A. In the human liver, the most important enzyme systems responsible for oxidative biotransformation of xenobiotics are the cytochrome P450 (CYP) family and flavin-containing monooxgenases (*FMOs*).[63] Members of both families are also represented in cluster A, for instance EC 1.14.13.8 (*FMO*) or EC 1.14.14.1 (unspecific monooxygenase), which are both reported to be rather unspecific and catalyze a broad range of substrates and reactions.[63, 64] Whereas the presence of literature-reported unspecific monooxygenases in cluster A lends some support to the structure of the heatmap, likely, only a subset of the oxygenases in that cluster is responsible for the observed oxidation reactions. Whether a majority of micropollutants is biotransformed by a small number of different ECs or whether a larger number of ECs is more equally involved in the observed biotransformation reactions therefore remains unclear at present. Notably, if multiple ECs are involved in a specific biotransformation reaction, then the best correlation would not be expected for an individual EC-micropollutant pair but for a linear combination of involved ECs. However, testing the significance of different linear combinations is beyond what can be achieved given the sample size of the data set at hand.

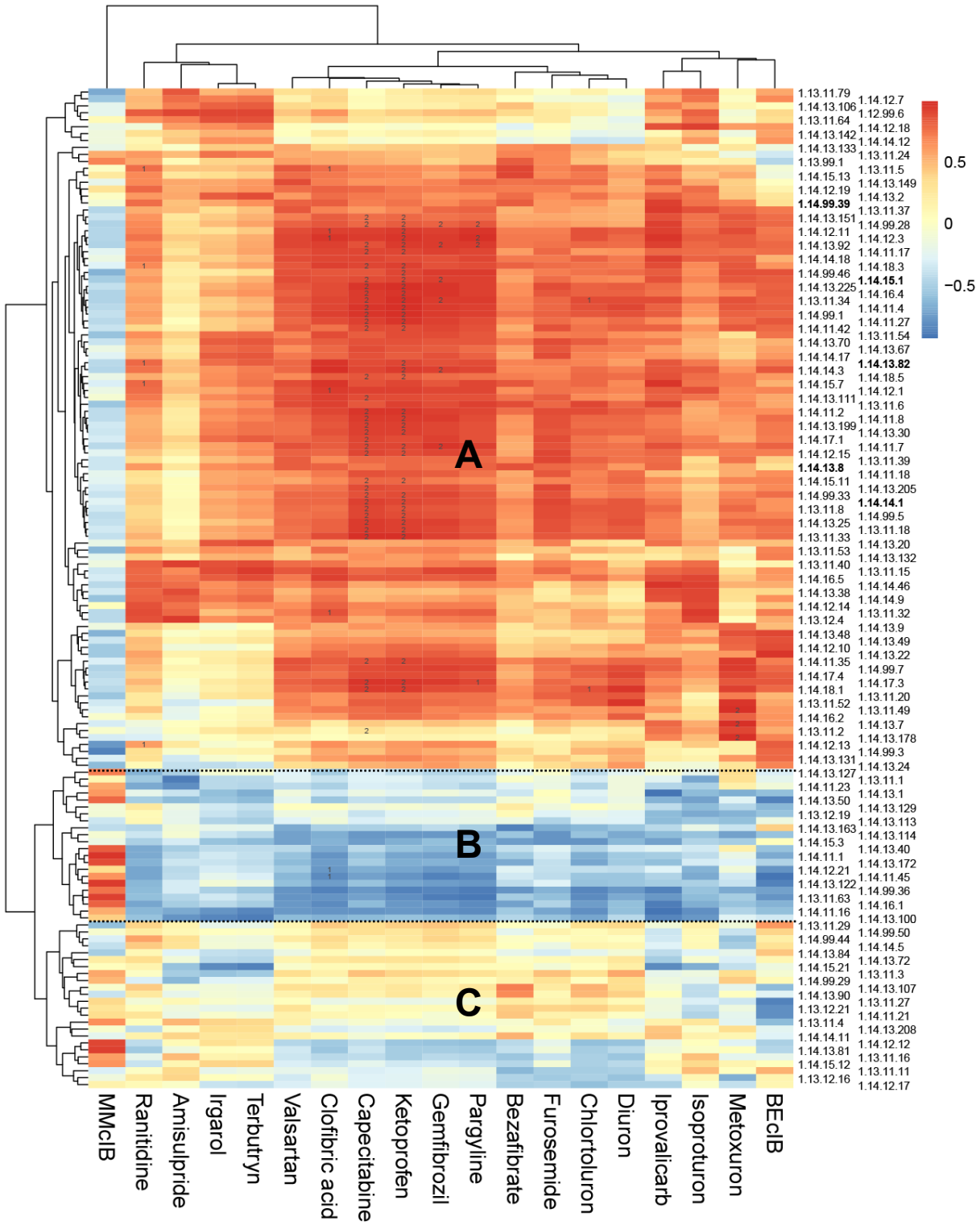


Figure 3. Heatmap showing correlations between gene transcripts of 144 ECs in sub-subclasses associated with mono- or dioxygenases and rate constants of 19 micropollutants transformed by oxidation reactions. For micropollutants analyzed in Exp1 and Exp2, the mean r was calculated. The three indicated clusters represent groups of ECs showing mainly strong positive correlations (A), moderate correlations (B) or anti-correlation (C). The numbers in the heatmap represent significant correlations in Exp1 (1) or Exp2 (2) when accounting for multiple hypothesis testing ($P < 0.05$, n = 144 for each micropollutant and experiment). ECs that are further discussed in the text are printed in bold. Descriptions of the EC numbers are provided in the SI (Table S4).

Implications

For nitrification and the biotransformation of nitriles we observed significant proportional relationships between relative gene transcript abundances annotated to 4th level enzyme classes containing enzymes known or demonstrated to catalyze the respective transformation reactions and their chemical biotransformation rate constants in complex activated sludge communities. Whereas *amo* transcripts reach relative abundances of 0.2% of all annotated reads, the relative abundance of gene transcripts annotated as nitrile hydratases remained below 0.001% (10^{-5}) in all samples. These results demonstrate that metatranscriptomic information can indeed quantitatively predict community functions such as micropollutant biotransformation in a genuine complex microbial community, even if the function is of low abundance. While these results are promising, one potential shortcoming of our data set is the lack of replicates for metatranscriptomic analysis. Reproducibility and precision of metatranscriptomic-based gene transcript abundances should therefore be further explored in subsequent work.

Still, the application of metatranscriptomics-based association mining in micropollutant biotransformation research is potentially hampered by a lack of statistical power to detect meaningful associations. In this study, we therefore used the available data to illustrate, for the cases of nitrile biotransformation and oxidation reactions, how additional information from transformation product analysis, which allows characterizing general reaction types and hence selection of EC numbers that potentially catalyze the respective reactions, can increase the statistical power of association mining. Specifically, we found a number of ECs that are known to be associated with unspecific enzymes, e.g., EC 1.14.14.1 (unspecific monooxygenase), EC 1.14.13.8 (*FMO*) and EC 4.2.1.84 (nitrile hydratase), to strongly correlate with the rate constants of compounds undergoing the respective types of transformations. Furthermore, for oxidative reactions, the distributions of correlation coefficients were shifted towards higher correlation coefficients when selecting plausible ECs using prior knowledge. In analogy to terminology used in high-resolution mass spectrometry where a distinction between suspect and non-target analysis is made when searching for either somewhat expected (e.g., predicted transformation products) or completely unknown compounds,[65] the approach presented here could be

considered a suspect association mining rather than a completely untargeted approach. In high-resolution mass spectrometry-based transformation product analysis, suspect screening has been shown to capture >80% of the products discovered when using both suspect and non-target analysis jointly, while being much less data greedy and time consuming than non-target analysis.[66] Based on our results, we are confident that suspect association mining based on additional experimental evidence from transformation product analysis (or also based on reaction-associated selections of ECs as derived from specialized databases, e.g., the Eawag-BBD[22]) is similarly promising in that it reduces data needs without losing too many true associations.

Although a statistically significant correlation never implies causality, the established linkages provide a basis for further investigating the causality for individual enzyme-micropollutant relationships using orthogonal information, e.g., from studies with selective inhibitors or genetically modified cells that overexpress specific enzymes. Also, linkages may be of predictive value even when the causality remains unclear. To this end, we propose to further test the robustness of the here obtained correlations in future studies with activated sludge and microbial communities from different natural or engineered systems. With respect to oxidative biotransformation reactions, such orthogonal information will reveal whether a sub-selection of the identified 98 ECs correlating with oxidation reactions could serve as indicators of oxidative biotransformation capacity more generally, i.e., in different environments. Such an increased general understanding of the relationship between different biotransformation reactions and metatranscriptomic information may eventually allow predicting environment-specific biotransformation rates and pathways from metatranscriptomic data, which would strongly support environmental risk assessment of chemical contaminants in the future.

Acknowledgements

We thank Dr. Jean-Claude Walser (Genomic Diversity Center, Zürich), Dr. Thomas B. Hofstetter and Andreas Scheidegger (both Eawag) for helpful discussions and Bernadette Vogler (Eawag) for help with chemical analyses. We acknowledge financial support from the European Research Council

under the European Union's Seventh Framework Programme (ERC grant agreement no. 614768, PROduCTS).

Supporting Information

Contains additional information on data normalization, nitrile hydration reactions, oxidation reactions and substitution reactions.

References

1. K. Fenner; S. Canonica; L. P. Wackett; M. Elsner, Evaluating pesticide degradation in the environment: blind spots and emerging opportunities. *Science* **2013,** *341*, 752-8.
2. R. P. Schwarzenbach; P. M. Gschwend; D. M. Imboden, *Environmental Organic Chemistry*. John Wiley & Sons: 2005.
3. B. Khan; L. S. Lee, Soil temperature and moisture effects on the persistence of synthetic androgen 17α-trenbolone, 17β-trenbolone and trendione. *Chemosphere* **2010,** *79*, 873-9.
4. T. Borch; R. Kretzschmar; A. Kappler; P. Van Cappellen; M. Ginder-Vogel; A. Voegelin; K. Campbell, Biogeochemical Redox Processes and their Impact on Contaminant Dynamics. *Environ. Sci. Technol.* **2010,** *44*, 15-23.
5. B. C. Okeke; J. E. Smith; A. Paterson; I. A. Watson-Craik, Influence of environmental parameters on pentachlorophenol biotransformation in soil by *Lentinula edodes* and *Phanerochaete chrysosporium*. *Appl. Microbiol. Biotechnol.* **1996,** *45*, 263-6.
6. Y. Luo; W. Guo; H. H. Ngo; L. D. Nghiem; F. I. Hai; J. Zhang; S. Liang; X. C. Wang, A review on the occurrence of micropollutants in the aquatic environment and their fate and removal during wastewater treatment. *Sci. Total Environ.* **2014,** *473-474*, 619-41.
7. M. S. McLachlan; H. Zou; T. Gouin, Using Benchmarking To Strengthen the Assessment of Persistence. *Environ. Sci. Technol.* **2017,** *51*, 4-11.
8. V. de Lorenzo, Systems biology approaches to bioremediation. *Curr. Opin. Biotech.* **2008,** *19*, 579-589.
9. T. Gu; C. Y. Zhou; S. R. Sorensen; J. Zhang; J. He; P. W. Yu; X. Yan; S. P. Li, The Novel Bacterial *N*-Demethylase PdmAB Is Responsible for the Initial Step of *N*,*N*-Dimethyl-Substituted Phenylurea Herbicide Degradation. *Appl. Environ. Microbiol.* **2013,** *79*, 7846-7856.
10. X. Maymo-Gatell; Y. T. Chien; J. M. Gossett; S. H. Zinder, Isolation of a bacterium that reductively dechlorinates tetrachloroethene to ethene. *Science* **1997,** *276*, 1568-1571.
11. J. Dunbar; S. White; L. Forney, Genetic diversity through the looking glass: Effect of enrichment bias. *Appl. Environ. Microbiol.* **1997,** *63*, 1326-1331.
12. S. Larcher; V. Yargeau, Biodegradation of sulfamethoxazole by individual and mixed bacteria. *Appl. Microbiol. Biotechnol.* **2011,** *91*, 211-218.
13. B. Jiang; N. Jin; Y. Xing; Y. Su; D. Zhang, Unraveling uncultivable pesticide degraders via stable isotope probing (SIP). *Crit. Rev. Biotech.* **2018,** *38***,** 1025-1048.
14. M. G. Dumont; J. C. Murrell, Stable isotope probing - linking microbial identity to function. *Nat. Rev. Microbiol.* **2005,** *3*, 499-504.
15. I. B. Lolas; X. J. Chen; K. Bester; J. L. Nielsen, Identification of triclosan-degrading bacteria using stable isotope probing, fluorescence in situ hybridization and microautoradiography. *Microbiology* **2012,** *158*, 2796-2804.

16. P. Falås; K. S. Jewell; N. Hermes; A. Wick; T. A. Ternes; A. Joss; J. L. Nielsen, Transformation, CO2 formation and uptake of four organic micropollutants by carrier-attached microorganisms. *Water Res.* **2018,** *141*, 405-416.
17. K. Kundu; S. Marozava; B. Ehrl; J. Merl-Pham; C. Griebler; M. Elsner, Defining lower limits of biodegradation: atrazine degradation regulated by mass transfer and maintenance demand in Arthrobacter aurescens TC1. *ISME J.* **2019**, *13*, 2236-2251
18. M. Alexander, Biodegradation of Organic-Chemicals. *Environ. Sci. Technol.* **1985,** *19*, 106-111.
19. K. Fischer; M. Majewsky, Cometabolic degradation of organic wastewater micropollutants by activated sludge and sludge-inherent microorganisms. *Appl. Microbiol. Biotechnol.* **2014,** *98*, 6583-97.
20. D. E. Helbling; D. R. Johnson; M. Honti; K. Fenner, Micropollutant Biotransformation Kinetics Associate with WWTP Process Parameters and Microbial Community Characteristics. *Environ. Sci. Technol.* **2012,** *46*, 10579-10588.
21. C. Monard; F. Martin-Laurent; O. Lima; M. Devers-Lamrani; F. Binet, Estimating the biodegradation of pesticide in soils by monitoring pesticide-degrading gene expression. *Biodegradation* **2013,** *24*, 203-13.
22. Eawag-BBD; http://eawag-bbd.ethz.ch/index.html. (14.08.2019),
23. D. R. Johnson; D. E. Helbling; Y. Men; K. Fenner, Can meta-omics help to establish causality between contaminant biotransformations and genes or gene products? *Environ. Sci.-Wat. Res.* **2015,** *1*, 272-278.
24. B. Futcher; G. I. Latter; P. Monardo; C. S. McLaughlin; J. I. Garrels, A sampling of the yeast proteome. *Mol. Cell. Biol.* **1999,** *19*, 7357-68.
25. R. de Sousa Abreu; L. O. Penalva; E. M. Marcotte; C. Vogel, Global signatures of protein and mRNA expression levels. *Mol. BioSyst.* **2009,** *5*, 1512-26.
26. T. Maier; M. Guell; L. Serrano, Correlation of mRNA and protein in complex biological samples. *FEBS Letters* **2009,** *583*, 3966-73.
27. P. Wilmes; A. Heintz-Buschart; P. L. Bond, A decade of metaproteomics: Where we stand and what the future holds. *Proteomics* **2015,** *15*, 3409-3417.
28. L. B. Stadler; J. Delgado Vela; S. Jain; G. J. Dick; N. G. Love, Elucidating the impact of microbial community biodiversity on pharmaceutical biotransformation during wastewater treatment. *Microb. Biotechnol.* **2017**, *11*, 995-1007.
29. D. E. Helbling; M. Ackermann; K. Fenner; H. P. Kohler; D. R. Johnson, The activity level of a microbial community function can be predicted from its metatranscriptome. *ISME J.* **2012,** *6*, 902-4.
30. S. Achermann; P. Falås; A. Joss; C. B. Mansfeldt; Y. Men; B. Vogler; K. Fenner, Trends in micropollutant biotransformation along a solids retention time gradient. *Environ. Sci. Technol.* **2018,** *52*, 11601−11611.

578 31. C. Mansfeldt; S. Achermann; Y. Men; J.-C. Walser; K. Villez; A. Joss; D. R. Johnson; K.
579 Fenner, Microbial residence time is a controlling parameter of the taxonomic composition and
580 functional profile of microbial communities. *ISME J.* **2019**, *13*, 1589-1601.
581 32. E. Kopylova; L. Noe; H. Touzet, SortMeRNA: fast and accurate filtering of ribosomal RNAs
582 in metatranscriptomic data. *Bioinformatics* **2012,** *28*, 3211-7.
583 33. B. Buchfink; C. Xie; D. H. Huson, Fast and sensitive protein alignment using DIAMOND.
584 *Nat. Methods* **2015,** *12*, 59-60.
585 34. Y. Benjamini; Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful
586 Approach to Multiple Testing. *J. Roy. Stat. Soc. B Met.* **1995,** *57*, 289-300.
587 35. K. Tipton; S. Boyce, History of the enzyme nomenclature system. *Bioinformatics* **2000,** *16*,
588 34-40.
589 36. M. Kotera; Y. Okuno; M. Hattori; S. Goto; M. Kanehisa, Computational assignment of the EC
590 numbers for genomic-scale analysis of enzymatic reactions. *J. Am. Chem. Soc.* **2004,** *126*, 16487-98.
591 37. C. Polson; P. Sarkar; B. Incledon; V. Raguvaran; R. Grant, Optimization of protein
592 precipitation based upon effectiveness of protein removal and ionization effect in liquid
593 chromatography-tandem mass spectrometry. *J. Chromatogr. B* **2003,** *785*, 263-275.
594 38. D. J. Rocha; C. S. Santos; L. G. C. Pacheco, Bacterial reference genes for gene expression
595 studies by RT-qPCR: survey and analysis. *Anton. Leeuw. Int. J. G.* **2015,** *108*, 685-93.
596 39. K. Kaewpipat; C. P. Grady, Jr., Microbial population dynamics in laboratory-scale activated
597 sludge reactors. *Water Sci. Technol.* **2002,** *46*, 19-27.
598 40. J. H. Sherrard; E. D. Schroeder, Relationship between Observed Cell Yield Coefficient and
599 Mean Cell Residence Time in Completely Mixed Activated-Sludge Process. *Water Res.* **1972,** *6*, 1039.
600 41. H. Salvado, Effect of Mean Cellular Retention Time on Ciliated Protozoan Populations in
601 Urban Waste-Water Treatment Plants Based on a Proposed Model. *Water Res.* **1994,** *28*, 1315-1321.
602 42. J. S. Cech; P. Hartman; M. Macek, Bacteria and Protozoa Population-Dynamics in Biological
603 Phosphate Removal Systems. *Water Sci. Technol.* **1994,** *29*, 109-117.
604 43. T. Suzuki; P. J. Higgins; D. R. Crawford, Control selection for RNA quantitation.
605 *BioTechniques* **2000,** *29*, 332-7.
606 44. P. Kos, Short SRT (solids retention time) nitrification process/flowsheet. *Water Sci. Technol.*
607 **1998,** *38*, 23-29.
608 45. J. M. Caffrey; N. Bano; K. Kalanetra; J. T. Hollibaugh, Ammonia oxidation and ammonia-
609 oxidizing bacteria and archaea from estuaries with differing histories of hypoxia. *ISME J* **2007,** *1*, 660-
610 2.
611 46. J. A. Field; E. M. Thurman, Glutathione conjugation and contaminant transformation.
612 *Environ. Sci. Technol.* **1996,** *30*, 1413-1418.
613 47. B. Singh; K. Singh, Microbial degradation of herbicides. *Crit. Rev. Microbiol.* **2016,** *42*, 245-
614 261.

615 48. S. Achermann; V. Bianco; C. B. Mansfeldt; B. Vogler; B. A. Kolvenbach; P. F. X. Corvini; K.
616 Fenner, Biotransformation of Sulfonamide Antibiotics in Activated Sludge: The Formation of Pterin-
617 Conjugates Leads to Sustained Risk. *Environ. Sci. Technol.* **2018,** *52*, 6265-6274.
618 49. N. Allocati; L. Federici; M. Masulli; C. Di Ilio, Glutathione transferases in bacteria. *Febs J.*
619 **2009,** *276*, 58-75.
620 50. D. Voet; J. G. Voet, *Biochemistry*. 3rd ed.; John Wiley & Sons, Inc.: 2004.
621 51. M. Kobayashi; S. Shimizu, Versatile Nitrilases - Nitrile-Hydrolyzing Enzymes. *Fems*
622 *Microbiol. Lett.* **1994,** *120*, 217-223.
623 52. F. S. Sariaslani, Microbial Enzymes for Oxidation of Organic-Molecules. *Crit. Rev.*
624 *Biotechnol.* **1989,** *9*, 171-257.
625 53. D. R. Johnson; D. E. Helbling; T. K. Lee; J. Park; K. Fenner; H. P. E. Kohler; M. Ackermann,
626 Association of Biodiversity with the Rates of Micropollutant Biotransformations among Full-Scale
627 Wastewater Treatment Plant Communities. *Appl. Environ. Microbiol.* **2015,** *81*, 666-675.
628 54. S. M. Resnick; K. Lee; D. T. Gibson, Diverse reactions catalyzed by naphthalene dioxygenase
629 from *Pseudomonas* sp strain NCIB 9816. *J. Ind. Microbiol. Biotechnol.* **1996,** *17*, 438-457.
630 55. D. T. Gibson; R. E. Parales, Aromatic hydrocarbon dioxygenases in environmental
631 biotechnology. *Curr. Opin. Biotech.* **2000,** *11*, 236-243.
632 56. M. R. Anari; S. Khan; Z. C. Liu; P. J. Obrien, Cytochrome-P450 Peroxidase Peroxygenase
633 Mediated Xenobiotic Metabolic-Activation and Cytotoxicity in Isolated Hepatocytes. *Chem. Res.*
634 *Toxicol.* **1995,** *8*, 997-1004.
635 57. M. J. Pecyna; R. Ullrich; B. Bittner; A. Clemens; K. Scheibner; R. Schubert; M. Hofrichter,
636 Molecular characterization of aromatic peroxygenase from Agrocybe aegerita. *Appl. Microbiol.*
637 *Biotechnol.* **2009,** *84*, 885-97.
638 58. E. Fernandez-Fontaina; F. Omil; J. M. Lema; M. Carballa, Influence of nitrifying conditions
639 on the biodegradation and sorption of emerging micropollutants. *Water Res.* **2012,** *46*, 5434-5444.
640 59. N. H. Tran; T. Urase; O. Kusakabe, The characteristics of enriched nitrifier culture in the
641 degradation of selected pharmaceutically active compounds. *J. Haz. Mat.* **2009,** *171*, 1051-1057.
642 60. T. Yi; W. F. Harper, The link between nitrification and biotransformation of 17α-
643 ethinylestradiol. *Environ. Sci. Technol.* **2007,** *41*, 4311-4316.
644 61. H. Roh; N. Subramanya; F. Zhao; C.-P. Yu; J. Sandt; K.-H. Chu, Biodegradation potential of
645 wastewater micropollutants by ammonia-oxidizing bacteria. *Chemosphere* **2009,** *77*, 1084-1089.
646 62. D. Li; M. Alidina; J. E. Drewes, Role of primary substrate composition on microbial
647 community structure and function and trace organic chemical attenuation in managed aquifer recharge
648 systems. *Appl. Microbiol. Biotechnol.* **2014,** *98*, 5747-56.
649 63. S. K. Krueger; D. E. Williams, Mammalian flavin-containing monooxygenases:
650 structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacol. Therapeut.* **2005,**
651 *106*, 357-87.

652 64. R. A. Gottlieb, Cytochrome P450: major player in reperfusion injury. *Arch. Biochem. Biophys.*
653 **2003,** *420*, 262-7.
654 65. M. Krauss; H. Singer; J. Hollender, Anal. Bioanal. Chem. *Anal. Bioanal. Chem.* **2010,** *397*,
655 943-951.
656 66. D. E. Helbling; J. Hollender; H.-P. E. Kohler; K. Fenner, High-throughput identification of
657 microbial transformation products of organic micropollutants. *Environ. Sci. Technol.* **2010,** *44*, 6621-
658 6627.

659