A review paper re-submitted to *Biotechnology Advances*

Insights into microbial diversity in wastewater treatment systems: How far have we come?

Isabel Ferrera^a and Olga Sánchez^{b*}

^a Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar, ICM-CSIC, Barcelona, Spain; iferrera@icm.csic.es, tel. (+34) 93 230 9500

^b Departament de Genètica i Microbiologia, Facultat de Biociències, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain; olga.sanchez@uab.es, tel. (+34) 93 586 8022

*Correspondence: Olga Sánchez: olga.sanchez@uab.es, tel. (+34) 93 586 8022, FAX (+34) 93 581 2387

ABSTRACT

Biological wastewater treatment processes are based on the exploitation of the concerted activity of microorganisms. Knowledge on the microbial community structure and the links to the changing environmental conditions is therefore crucial for the development and optimization of biological systems by engineers. The advent of molecular techniques occurred in the last decades quickly showed the inadequacy of culture-dependent methodologies to unveil the great level of diversity present in sludge samples. Initially, culture-independent technologies and more recently the application of *-omics* in wastewater microbiology, has drawn a new view of microbial diversity and function of wastewater treatment systems. This article reviews the current knowledge on the topic placing emphasis on crucial microbial processes carried out in biological wastewater treatment systems driven by specific groups of microbes, such as nitrogen and phosphorus removal bacteria, filamentous and electrogenic microorganisms, as well as Archaea. Despite the recent *-omics* has offered substantial insights into the diversity and ecophysiology of these bacteria never envisioned before by providing millions of sequence reads at an unprecedented scale, studies based on high-throughput sequencing are still scarce. In order to obtain significant gains in the analysis of structure-function relationships, a greater sequencing investment is needed, particularly to uncover gene expression patterns of functionally relevant genes.

Keywords: molecular techniques, wastewater microbiology, *-omics*, nitrogen removal, phosphorus removal, electrogenic bacteria, archaea, bulking

1. Introduction

Through the history of mankind, water availability has played an essential role in human development. In this sense, global demand for water has been continuously rising due to population growth and increasing socioeconomic activities; in a hundred years the world population has tripled while water consumption has increased six-fold. This circumstance is particularly critical in areas where deficiency in hydric resources prevails, a situation that has prompted the need for wastewater treatment and subsequent reuse. The quantity and quality of waste generated and discharged into natural water bodies is a topic of serious concern in our world today, and the wealthiest nations invest a considerable effort for mitigating the pollution of dense populations (Oh et al., 2010). In this context, knowledge of the wastewater source is fundamental to determine its physical, chemical and biological characteristics and to establish an appropriate treatment strategy. Composition of wastewaters depends on their origin but in general, major contaminants found include organic compounds, xenobiotics, metals, suspended soils, nutrients (mainly nitrogen and phosphorus) and pathogenic microorganisms (Bitton, 2005). Depending on the final use of the water (drinking, recreation, irrigation, etc), several technologies can apply to remove these compounds. Among them, biological wastewater treatment systems rely on the use of microorganisms; the unique abilities of microbes to degrade organic matter, remove nutrients and transform toxic compounds into harmless products makes them the essential players in waste removal.

For a century, biological wastewater treatment plants (BWWTPs) based in activated sludge have been in charge of treating both domestic and industrial wastes and have constituted an essential instrument in environmental protection. Initially, by the end of the nineteenth century, removal of organic matter and suspended soils was the main requirement, but nowadays BWWTPs have additional objectives, such as elimination of toxic metals, odors, nutrients (N, P) and pathogens (Bitton, 2005). However, a drawback of this technology is the high performance costs, given the need for aeration as well as moving biomass between treatment tanks (Oh et al., 2010; Sheik et al., 2014). Additionally, powerful greenhouse gases such as CH₄ and N₂O are generated in BWWTPs (Sheik et al., 2014). Giving an answer to this problem, Sheik et al. (2014) proposed a bottom-up design approach for BWWTPs based in the concept of “wastewater biorefinery column”, which relies on the engineering of distinct ecological niches to guarantee the targeted enrichment of specific functional groups for sustainable production of bioenergy, bioplastic and fertilizers.

Currently, besides BWWTPs, other ecologically engineered systems are becoming alternative solutions to conventional processes due to their efficiency, low establishment costs and low operation and management requirements. Under the denomination of non-conventional technologies, different strategies can be distinguished. For instance, green filters use land planted with trees flooded with the wastewater to be treated; purification is the result of physical, chemical and biological factors (Fahd et al., 2007). Also, constructed wetlands are passive treatment systems constituted by lagoons or shallow ponds or channels planted with characteristic wetland vegetation (aquatic macrophytes), which mimic natural wetlands under a more controlled environment. Different categories of constructed wetlands exist according mainly to hydrology (open water-surface flow and subsurface flow), type of macrophytic growth (emerged, submerged, free-floating) and flow path (horizontal and vertical) (Vymazal, 2011). Other constructed systems simulate the natural processes of purification that occur in rivers and lakes (Middelbrooks et al., 1982) or are based on wastewater filtration through beds using peat (Couillard, 1994) or sand (Lian et al., 2014). An additional sustainable option that has gained thrust in the last decade includes the use of microbial fuel cells (MFCs), i.e. bioelectrochemical systems where bacteria oxidize organic compounds and the electrons generated in this process are

converted into electrical energy (Kelly and He, 2014; Mathuriya et al., 2014). The energy produced from wastewater treatment could recover the cost of the process, a promising perspective for this alternative.

A variety of lab-scale systems, such as particulate biofilm reactors or membrane bioreactors (MBR) among others, have also been used as pilot systems for wastewater treatment (Ballesteros Martín et al., 2011; Nicolella et al., 2000). They have the advantage of being compact and well-controlled systems with small area requirements which minimize the excess of sludge production; they can also maintain a high biomass age (of several weeks), while the high biomass concentration and mass transfer area result in high conversion capacities to specifically degrade key pollutants. Besides, they have the advantage that changes in the design can be made more economically. However, biofilm formation on carriers can lead to long start-up times, and scaling-up from laboratory scale reactors to full-scale applications exhibit some unknown aspects, since it is difficult to predict the behaviour of a complex process (Nicolella et al., 2000).

In all these technologies, performance of the wastewater treatment process is based on the exploitation of the concerted activity of the microorganisms involved. Knowledge on the microbial community structure and the links to the changing environmental conditions is therefore crucial for the development and optimization of biological systems by engineers. Before the development of cultivation-independent techniques, the key drivers of pollutant removal processes were hardly known. This scenario overturned in the last decade of the 20th century with the development of a suite of molecular methods that allowed the study of these organisms circumventing cultivation. Further, the *-omics* revolution occurred in the last years has boosted our understanding on the biology of these biotechnological processes. In this review, we provide an outline on the current scene of the diversity and genomics of prokaryotes in various wastewater treatment systems highlighting the findings of some crucial microbial processes in this field and discussing some unresolved questions.

2. Approaches to study wastewater microbiology

The advent of molecular methods such as Polymerase Chain Reaction (PCR) techniques revolutionized the study of microbial diversity previously constrained by our inability to grow the majority of microorganisms in culture (Rappé and Giovannoni, 2003). PCR amplification and sequencing (Fig. 1) have been used now for decades to study the presence, diversity and expression of ribosomal and protein-coding genes in these constructed environments, allowing the identification of the key microorganisms involved in processes such as carbon-degrading and ammonia oxidation (Gilbride et al., 2006; Wagner and Loy, 2002). These PCR approaches shed light on the diversity of main players in BWWTPs and how genetic diversity is structured in relation to certain environmental parameters (Wang et al., 2012; Wells et al., 2011). Yet, there are inherent PCR amplification- and primer-related biases that limit the quantitative information that can be extracted from these approaches. Perhaps the most critical step from these is primer design, which may be remarkably challenging for the highly diverse functional genes. However, fluorescent in situ hybridization (FISH) using group-specific rRNA-targeted oligonucleotide probes circumvents the PCR bias and has been largely applied to wastewater microbiology providing quantitative information on the dominant groups involved in the different processes of waste removal (Joss et al., 2011; Nielsen et al., 2009b; Wagner et al., 1996).

After the initial application of these approaches, the rise of the “*-omics* era” has resulted in a turning point in studying phylogenetic and functional diversity of wastewater treatment systems (Fig. 1). The development of the so-called High Throughput Sequencing (HTS) techniques, i.e., 454-pyrosequencing and Illumina

among others, has allowed producing millions of sequence reads at an unimaginable low cost. Whole genome sequencing of isolated microorganisms was the first step taken to elucidate the functional potential of wastewater relevant microorganisms (Table 1). Beyond the initial genome sequencing of cultures, metagenomics (the massive DNA sequencing of all genes present in a microbial community) and metatranscriptomics (sequencing of all genes being transcribed, and therefore, likely expressed in a microbial community) have been applied to unveil the genomic potential and transcripts of whole microbial communities circumventing isolation and PCR (Venter et al., 2004; Yu and Zhang, 2012). However, these approaches have also the limitation that most of the information is derived from the most abundant members of the community. Despite treatment systems present in general less diversity than natural ecosystems, significant gains in the analysis of gene expression patterns require a greater sequencing investment and the combination of both metagenomics and metatranscriptomics, but these studies are still rare (Yu and Zhang, 2012). More effort is needed to understand the functional capabilities of microbial communities in wastewater treatments and what genes are expressed under which circumstances to better understand the relationship between structure and function in these microbially-mediated processes.

In addition to nucleic acid based methods, the study of metaproteomics (Fig. 1) can lead to the identification of proteins involved in known processes and also to the discovery of unrecognized metabolic processes and functional roles, since there are still many protein families whose function is unknown. Gene sequence similarity does not necessarily reflect functional conservation and simple modifications can lead to major modifications on the activity and the specificity of the corresponding enzyme (Ufarté et al., 2015). A recent metaproteomics pilot study allowed the identification of most proteins involved in glycolysis, citrate cycle and nitrogen removal in a wastewater treatment plant (Püttker et al., 2015). Metabolomics (the analysis of all cellular metabolites) is perhaps currently the least used field of the *-omics* approaches primarily due to analytical complexity but it is clear that its emergence will be critical to understand microbe–microbe and microbe–molecule interactions.

Despite the potential of *-omics* for studying functional capabilities within a microbial community, the power of culturing individual species cannot be neglected. Cultures have provided the basis for our understanding of the microbial world and are necessary for physiological studies in a simplified and controlled environment. These studies are crucial to interpret *-omics* data and to ascertain the relevance of the newly discovered functions of proteins through *-omics*. Until recently, culturing was also necessary to generate reference genomes, but access to the genome of uncultured organisms is nowadays granted thanks to the development of single-cell sorting and manipulation and whole-genome amplification (SAGs) (Fig. 1) (Blainey, 2013; Dichosa et al., 2015). However, despite of the enormous advances gained on knowledge of the identity and function of wastewater microorganisms, we are still far from reaching a full understanding of these largely complex microbial communities. The combination of all these approaches with systematic measurements and experimental validation (ecosystems biology) will be the path to achieve a holistic characterization of microbial consortia responsible for waste treatment (Narayanasamy et al., 2015).

202 **Table 1. Some examples of wastewater relevant microorganisms for which sequencing projects**
 203 **have been completed**

Process	Organism	Reference and/or Genbank Accession Number
Carbon and toxic compounds removal	<i>Aminomonas paucivorans</i>	Pitluck et al., 2010; NZ_AEIV000000000
	<i>Anaerobaculum mobile</i>	Mavromatis et al., 2013; NC_018024
	<i>Arthrobacter nitroguajacolicus</i>	Niewerth et al., 2012; BBFS000000000
	<i>Cloacibacillus evryensis</i>	NZ_JFBR000000000
	<i>Comamonas testosteroni</i>	Fukuda et al., 2014; CP006704
	<i>Cand. Competibacter denitrificans</i>	NZ_CBTJ020000020
	<i>Can. Contendobacter odensis</i>	NZ_CBTk000000000
	<i>Exiguobacterium alkaliphilum</i>	NZ_JJMW01000019
	<i>Methanocorpusculum labreanum</i>	Anderson et al., 2009; CP000559
	<i>Methanofollis liminatans</i>	NZ_AGCO000000000
	<i>Methanolinea tarda</i>	Yamamoto et al., 2014; AGIY020000000
	<i>Pseudomonas moraviensis</i>	Hunter et al., 2014; AYMZ000000000
	<i>Pseudomonas otitidis</i>	NZ_JGYF01000080
	<i>Pseudomonas stutzeri</i>	Busquets et al., 2013; AMVM010000000
	<i>Rhodococcus ruber</i>	Shumkova et al., 2015; LDUF010000000
	<i>Sediminibacterium</i> sp.	Ayarza et al., 2014; AXUM010000000
	<i>Sphingomonas</i> sp.	Chen et al., 2014; JMUB000000000
	<i>Thauera</i> sp.	Dichosa et al., 2015; JTDM010000000
Nitrogen removal		
Ammonia oxidation	<i>Nitrosomonas europaea</i>	Chain et al., 2013; AL954747
	<i>Nitrosomonas eutropha</i>	Stein et al., 2007; CP000450
Nitrite oxidation	<i>Nitrobacter hamburgensis</i>	Starkenburger et al., 2008; CP000319
	<i>Nitrobacter winogradskyi</i>	Starkenburger et al., 2006; CP000115
	<i>Nitrospina gracilis</i>	Lücker et al., 2013; NZ_CAQJ010000000
	' <i>Cand. Nitrospira defluvii</i> '	Lücker et al., 2010; GCA_000196815.1
Denitrification	<i>Paracoccus denitrificans</i>	Siddavattam et al., 2011; CP002897
	<i>Thiobacillus denitrificans</i>	Beller et al., 2006; CP000116
Anammox	' <i>Cand. Kuenenia stuttgartiensis</i> '	Strous et al., 2006; CT030148
Phosphorus removal	<i>Gemmatimonas aurantiaca</i>	AP009153
	' <i>Cand. Accumulibacter phosphatis</i> '	Mao et al., 2014; NC_013194.1
	<i>Tetrasphaera jenkinsii</i>	Kristiansen et al., 2013; NZ_HF571038

204

205 **3. Microbial diversity in wastewater treatment systems**

206 Initial description of the microbial structure in a variety of treatment systems was
 207 conducted in the last decades through the application of culture-independent
 208 techniques such as DGGE (Adrados et al., 2014; Ballesteros Martín et al., 2011; Boon
 209 et al., 2002), T-RFLP (Eschenhagen et al., 2003; Hiraishi et al., 2000; Wang et al.,
 210 2011), cloning (Sánchez et al., 2011; Snaird et al., 1997; Vainio et al., 1997), and
 211 particularly FISH (Kämpfer et al., 1996; Manz et al., 1994; Sánchez et al., 2011;
 212 Wagner et al., 1994; Wallner et al., 1995). Concerning activated sludge, perhaps the
 213 most commonly used technology for treating sewage, these studies highlighted the
 214 dominance of the phylum Proteobacteria, followed by other groups (like Bacteroidetes,
 215 Chloroflexi, Actinobacteria, Planctomycetes, Firmicutes, etc) in different proportions

depending on the conditions (Boon et al., 2002; Snaidr et al., 1997; Wagner et al., 1994; Wang et al., 2011). These observations have been confirmed when further high throughput techniques (HTS) were applied (Ju et al., 2014; Kwon et al., 2010; Sanapareddy et al., 2009; Sánchez et al., 2013). Furthermore, HTS allowed the identification of groups that had remained undetected with the former molecular methods, deepening our information on the diversity of activated sludge. Additionally, metagenomic studies pointed out to the dominance of functional categories involved in carbohydrates, protein metabolism and amino acids derivatives (Ju et al., 2014) or the metabolism of aromatic compounds (Sanapareddy et al., 2009). On the other hand, these new methodologies confirmed that members of the Domain Bacteria are largely responsible for the majority of carbon removal (Sánchez et al., 2011; Yu and Zhang, 2012) while Archaea seem to be less relevant.

Proteobacteria also showed to be dominant in other wastewater treatment processes such as constructed wetlands, although other groups as Bacteroidetes, Actinobacteria, Firmicutes, Chloroflexi or Acidobacteria were retrieved in some samples with fingerprinting methods (Adrados et al., 2014; Sidrach-Cardona et al., 2015). In these engineered systems, microbial communities appeared to be mainly affected by hydraulic configuration and plant presence, as well as organic load in a lesser extend. Pyrosequencing analyses of different types of constructed wetlands showed a wider range of phyla involved in wastewater treatment besides those mentioned above, for instance, Verrucomicrobia, Planctomycetes, Nitrospirae, Cyanobacteria, TM7 or Gemmatimonadetes among others (Ansola et al., 2014; Arroyo et al., 2015; Zhao et al., 2015a; Zhong et al., 2015). These results were confirmed with recent metagenomic analyses (Bai et al., 2014). In any case, all these studies have reasserted the predominance of Proteobacteria.

4. Microorganisms involved in nitrogen removal

Nitrification and denitrification constitute key processes in wastewater treatment because they are considered the major mechanisms of nitrogen (N) removal. Ammonia and nitrite are toxic to aquatic life and can also contribute to eutrophication of receiving waters. It is known that ammonia oxidation can be performed by aerobic and anaerobic oxidizers, which include the proteobacterial ammonia oxidizers (AOB) and the anaerobic ammonia oxidation (anammox) bacteria (Fig. 2). The proteobacterial ammonia oxidizers, which can carry out aerobic or anaerobic ammonia oxidation depending on oxygen concentration (Schmidt et al., 2003), include two monophyletic groups, the Betaproteobacteria ammonia oxidizers *Nitrosomonas* and *Nitrosospira*, and the Gammaproteobacteria *Nitrosococcus* (with the exception of *Nitrosococcus mobilis*, a beta-proteobacterium lineage); under oxic conditions, the main product from their metabolism is nitrite, while dinitrogen, nitrite and nitric oxide are produced under anoxic conditions. On the other hand, anammox, a novel and promising biotechnological process for nitrogen removal from ammonia to nitrogen gas, is mediated by a monophyletic group of bacteria within the phylum Planctomycetes; it includes five genera identified by culture-independent techniques (*Candidatus* 'Brocadia', *Candidatus* 'Kuenenia', *Candidatus* 'Scalindua', *Candidatus* 'Anammoxoglobus' and *Candidatus* 'Jettenia') (Fig. 2) (Ali et al., 2013; Hu et al., 2011).

Additionally, other heterotrophic bacteria (Kim et al., 2005) and some Archaea (AOA) have been reported to oxidize ammonia (You et al., 2009). Although little attention has been paid to the role of Archaea in wastewater treatment processes, the results obtained comparing the abundance of AOB and AOA (now classified within the phylum Thaumarchaeota) (Brochier-Armanet et al., 2008; Spang et al., 2010) by means of molecular methods are controversial, since some studies report either the complete absence of AOA (Mußmann et al., 2011; Yu and Zhang, 2012), a minimal or

equal contribution (Jin et al., 2010; Sonthiphand et al., 2011; Wells et al., 2009; Yapsakli et al., 2011), or a dominance of AOA over AOB (Bai et al., 2012; Kayee et al., 2011; Sauder et al., 2012). Thus, no clear conclusion can be drawn from these studies about the role of Archaea in ammonia oxidation and therefore, more studies are needed to ascertain under which circumstances they play a significant role.

Nitrification implies also a second oxidative step from nitrite to nitrate (nitrite oxidation), which is performed by the aerobic nitrite bacteria (NOB) and traditionally comprise members of the genera *Nitrobacter* (Alphaproteobacteria), *Nitrococcus* (Gammaproteobacteria) and *Nitrospira* (Nitrospirae) (Wagner et al., 2002). However, little is known about their diversity and ecological niche differentiation in complex communities. In a recent study combining different molecular techniques, Gruber-Dorninger and colleagues (2015) revealed the presence of diverse *Nitrospira* clusters with functional differences concerning the preferred nitrite concentration, the use of formate and the spatial co-aggregation with other AOB; they also reported the presence of as many as 121 species-level *nrxB* (encoding the beta subunit of nitrite oxidoreductase) operational taxonomic units of *Nitrospira*, showing an unusual high diversity not documented before.

In the last years, molecular techniques such as DGGE, cloning, tag sequencing, metagenomics or FISH, or a combination of some of them, have shed light in the identification of the microorganisms involved in nitrogen removal through nitrification in wastewater treatment systems. Table 2 offers a description of the main nitrifiers (AOB and NOB) identified in activated sludge samples, as well as in non-conventional engineered systems. Some of these systems are dominated by a single ammonia or nitrite oxidizer species, while others harbor different populations. *Nitrosomonas* sp. appears to be virtually present in all the systems described, as well as the NOB *Nitrospira*. Overall, the results confirm the presence of the already described AOB and NOB in previous works, but also, other not so common genera, such as *Nitrosovibrio* or *Nitrospina* and *Nitrotoga* were found.

Denitrification, the transformation of nitrate or nitrite into gaseous forms (N_2 and N_2O), is an ability dispersed in diverse phylogenetic lineages. Different chemoorganotrophic, lithoautotrophic, and phototrophic bacteria, as well as some Archaea and fungi, can perform this process (Zumft, 1997). It is essentially a facultative trait, triggered by environmental parameters such as oxygen reduced conditions and availability of a nitrogen oxide. Similarly to other groups, molecular techniques have proved to be valuable for the study of the structure and function of wastewater denitrification communities (Lu et al., 2014). The genera *Alcaligenes*, *Pseudomonas*, *Methylobacterium*, *Bacillus*, *Paracoccus*, *Hyphomicrobium*, as well as members of the order Rhodocyclales among others are able to carry out denitrification (Hosselhoe et al., 2009; Wagner et al. 2002; Zumft, 1997). Recently, genomic analyses have revealed novel processes concerning non-conventional types of denitrification (Isobe and Ohte, 2014). For example, some microorganisms, such as *Candidatus Methyloirabilis oxyfera*, could couple the formation of N_2 with the anaerobic oxidation of CH_4 (Ettwig et al., 2010), and non-denitrifying species, like *Anaeromixobacter dehalogenans*, could catalyze the reduction of N_2O to N_2 by means of an atypical nitrous oxide reductase (Sanford et al., 2012).

Furthermore, recent metagenomic studies in a laboratory-scale sequencing batch reactor (SBR) treating artificial wastewater have offered new insights into the simultaneous removal of nitrogen and organic carbon with the coupling of anammox, denitrification and dissimilatory nitrate reduction to ammonium (Shu et al., 2015). Metagenomics was also useful to explore microbial community composition and potential function in a constructed wetland (Bai et al., 2014), where nitrification and denitrification were the main transformation processes in the N cycle, highlighting the major role of microbes on the removal of nutrients and contaminants from wastewater.

Figure