Genomics update

Genomics of biological wastewater treatment

Roland J. Siezen1,2,3* and Marco Galardini3

1Kluyver Centre for Genomics of Industrial Fermentation, TI Food and Nutrition, 6700AN Wageningen, the Netherlands.
2NIZO food research, 6710BA Ede, the Netherlands.
3Center for Molecular and Biomolecular Informatics, Radboud University Nijmegen Medical Centre, 6500HB Nijmegen, the Netherlands.

Cleaning up wastewater from domestic sewage and industrial processes is essential for safeguarding public health and the environment (for an overview of wastewater treatment processes see http://en.wikipedia.org/wiki/Sewage_treatment). Biological wastewater treatment plants (WWTPs) employing sludge represent one of the most widely used biotechnological processes, with more than 120 billion litres treated daily in the USA alone (Fig. 1). Removal of organic carbon and other nutrients, mainly nitrogen (N) and phosphorus (P), by sludge microbes is essential in order to avoid eutrophication and deterioration of recipient surface waters (Seviour et al., 2003). The complexity of wastewater microbial communities, based on the analysis of 16S rRNA sequences, is known to be enormous (Chouari et al., 2005a,b) (Fig. 2). As we have very limited knowledge of composition, dynamics and stability of microbial communities, several processes in wastewater treatment are generally regarded as being a ‘black box’. Their improvement is often based on a trial-and-error approach, which leads to inconsistent results and unexpected behaviours (Daims et al., 2006a). In recent years, with the development of several new high-throughput sequencing platforms (Hall, 2007; Marsh, 2007), metagenome sequencing strategies (Schloss and Handelsman, 2005; Tringe and Rubin, 2005; Schmeisser et al., 2007) and bioinformatics toolboxes (Chen and Pachter, 2005; Tringe et al., 2005; Foerstner et al., 2006; Raes et al., 2007), the genome analysis of complex communities has become much more accessible, albeit by no means easier. Here we give a brief update of the current status of (meta)genome sequencing and mining of wastewater-related microorganisms and communities.

Trends in wastewater genomics

For decades, the putative key players in wastewater treatment processes were those that were culturable and well studied, while the really important species remained non-culturable or unknown (Crocetti et al., 2000; Chouari et al., 2005a). A variety of genomes of cultured microbes have been sequenced in the past 5 years, but now the focus is shifting towards culture-independent metagenome sequencing. Table 1, extracted from the GOLD

---

Fig. 1. Wastewater treatment plant near Arnhem, the Netherlands (photo M. Galardini).

*For correspondence. E-mail r.siezen@cmbH.ru.nl; Tel. (31) 2436 19559; Fax (31) 2436 19395.

© 2008 The Authors
Journal compilation © 2008 Society for Applied Microbiology and Blackwell Publishing Ltd
database (Liolios et al., 2008), gives an overview of completed and ongoing sequencing projects, divided into relevant steps in wastewater treatment. Clearly the focus is still on bacteria, with virtually no information on archaea and eukarya, even though they are known to play a role in wastewater processes (Liu et al., 2008; Rehman et al., 2008). In many cases, wastewater microbes have first been enriched in laboratory bioreactors prior to metagenome sequencing (Strous et al., 2006; Maixner et al., 2008; Pelletier et al., 2008). In addition, metagenome sequences have provided a first insight into metabolic requirements of specific microbes, allowing further optimization of enrichment conditions (Garcia Martin et al., 2006). 16S rRNA sequences from metagenome libraries have been used to identify new bacterial divisions and species (Guermazi et al., 2008; Pelletier et al., 2008) and to visualize them using molecular techniques such as FISH (Roh et al., 2008; Yoon et al., 2008). Some recent breakthroughs are highlighted in more detail below.

### Anaerobic digestion: Candidatus Cloacamonas acidaminovorans

Metagenomics projects involve not only species that are predominant in the wastewater treatment, but also minor members of the population. For instance, a new bacterial division (WWE1) belonging to the *Spirochaetes*, with no cultured representative, has been characterized (Chouari et al., 2005b). Members of this division are present in anaerobic digesters, generally as a subdominant group (< 10% of the population), where they perform the decomposition of organic matter into methane and CO₂. Starting from a metagenomic analysis of sludge microbes from such a digester, Pelletier and colleagues (2008) succeeded in reconstructing the complete genome of a WWE1 bacterium, represented by only 2% of the fosmid library reads, using a novel sequencing and assembly strategy specifically designed during this project. The reconstruction started with seed sequences (fully sequenced fosmids on which WWE1 16S rRNA sequences were found), to which > 99% identical fosmid paired-end sequences from the metagenome library were added in an iterative fashion (Fig. 3). A circular genome of 2.2 Mb was then obtained by standard finishing techniques; its reconstruction was probably made possible by the extremely low sequence polymorphism of the dominant phylotype (Pelletier et al., 2008). This is the first time that such an iterative procedure had been used, and the first time ever a ‘complete’ genome from a previously unknown, uncultured and subdominant bacterium was reconstructed out of such
<table>
<thead>
<tr>
<th>Phylum</th>
<th>Organism(s)</th>
<th>Type</th>
<th>Sequencing</th>
<th>Habitat/isolation</th>
<th>Oxygen</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaerobic digestion</td>
<td>Unclassified Candidatus Cloacamonas acidaminovorans</td>
<td>Metagenome</td>
<td>Complete</td>
<td>Sludge, wastewater</td>
<td>Anaerobe</td>
<td>Pelletier et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Rubrivivax gelatinosus IL144</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic, sludge</td>
<td>Facultative</td>
<td>NITE</td>
</tr>
<tr>
<td></td>
<td>Firmicutes Aminomonas paucivorans GLU-3</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic, wastewater</td>
<td>Anaerobe</td>
<td>DSMZ, JGI</td>
</tr>
<tr>
<td></td>
<td>Firmicutes Aminobacterium colombiense ALA-1</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic, wastewater</td>
<td>Anaerobe</td>
<td>DSMZ, JGI</td>
</tr>
<tr>
<td>Anaerobic digestion and hydrogen production</td>
<td>Firmicutes Ethanoligenens harbinense YUAN-3T</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Sludge</td>
<td>Anaerobe</td>
<td>DSMZ, JGI</td>
</tr>
<tr>
<td>Phosphorus removal</td>
<td>Proteobacteria Candidatus Accumulibacter phosphatis</td>
<td>Metagenome</td>
<td>Complete</td>
<td>Sludge</td>
<td>Facultative</td>
<td>Garcia Martin et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Gemmatimonadetes Gemmatimonas aurantiaca T-27T</td>
<td>Genome</td>
<td>Complete</td>
<td>Wastewater</td>
<td>Aerobe</td>
<td>NITE</td>
</tr>
<tr>
<td></td>
<td>Unclassified Wastewater EBPR community</td>
<td>Metagenome</td>
<td>Complete</td>
<td>Sludge</td>
<td>Facultative</td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Firmicutes Microbacterium phosphovorus NM-1</td>
<td>Genome</td>
<td>Complete</td>
<td>Sludge</td>
<td>Anaerobe</td>
<td>DSMZ, JGI</td>
</tr>
<tr>
<td>Anaerobic ammonia oxidation (anammox)</td>
<td>Planctomycetes Candidatus Kuenenia stuttgartensis</td>
<td>Metagenome</td>
<td>Incomplete</td>
<td>Wastewater</td>
<td>Anaerobe</td>
<td>Strous et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Planctomycetes Brocadia anamnnoxidans</td>
<td>Genome</td>
<td>Complete</td>
<td>Aquatic</td>
<td>Anaerobe</td>
<td>University of Nijmegen, NL</td>
</tr>
<tr>
<td></td>
<td>Planctomycetes Gemmata obscuriglobus UQM 2246</td>
<td>Genome</td>
<td>Draft</td>
<td>Aquatic</td>
<td>Aerobe</td>
<td>JCVI</td>
</tr>
<tr>
<td></td>
<td>Unclassified Anaerobic ammonium-oxidizing community</td>
<td>Metagenome</td>
<td>Incomplete</td>
<td>Aquatic</td>
<td>Anaerobe</td>
<td>JGI</td>
</tr>
<tr>
<td>Ammonia oxidation</td>
<td>Proteobacteria Nitrosomonas europaea IFO</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil, aquatic</td>
<td>Aerobe</td>
<td>Chain et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Nitrosomonas eutropha C91 (C71)</td>
<td>Genome</td>
<td>Complete</td>
<td>Aquatic</td>
<td>Aerobe</td>
<td>Stein et al. (2007)</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Nitrosomonas oligotropha Nm45</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil</td>
<td>Aerobe</td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Nitrosomonas sp. 17</td>
<td>Genome</td>
<td>Complete</td>
<td>Marine</td>
<td>Aerobe</td>
<td>JCVI</td>
</tr>
<tr>
<td>Nitrile oxidation</td>
<td>Proteobacteria Nitrobacter hamburgensis X14</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil</td>
<td>Aerobe</td>
<td>Starkenburg et al. (2008)</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Nitrobacter winogradskyi Nb-255</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil</td>
<td>Facultative</td>
<td>Starkenburg et al. (2006)</td>
</tr>
<tr>
<td></td>
<td>Nitrospirae Nitrospira marina Nb-295</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic</td>
<td>Facultative</td>
<td>JCVI</td>
</tr>
<tr>
<td></td>
<td>Nitrospirae Candidatus Nitrospira defluvi</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic</td>
<td>Facultative</td>
<td>Genoscope, France</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Nitrospira gracilis Nb-211</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Aquatic</td>
<td>Facultative</td>
<td>JCVI</td>
</tr>
<tr>
<td>Denitrification</td>
<td>Proteobacteria Paracoccus denitrificans PD1222</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil</td>
<td>Aerobe</td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Deftia acidivorans SPH-1</td>
<td>Genome</td>
<td>Complete</td>
<td>Aquatic, sludge</td>
<td>Aerobe</td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria Thiobacillus denitrificans ATCC 25259</td>
<td>Genome</td>
<td>Complete</td>
<td>Soil</td>
<td>Facultative</td>
<td>Beller et al. (2006)</td>
</tr>
<tr>
<td>Biofilm formation and flocculation</td>
<td>Proteobacteria Acidovorax temperans</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Sludge</td>
<td>Aerobe</td>
<td>University of Auckland, NZ</td>
</tr>
<tr>
<td>Bulking</td>
<td>Chloroflexi Herpetosiphon aurantius ATCC 23779</td>
<td>Genome</td>
<td>Complete</td>
<td>Aquatic</td>
<td>Aerobe</td>
<td>JGI</td>
</tr>
<tr>
<td></td>
<td>Chloroflexi Anaerolinea thermophila UNI-1</td>
<td>Genome</td>
<td>Incomplete</td>
<td>Sludge</td>
<td>Anaerobe</td>
<td>NITE</td>
</tr>
</tbody>
</table>

NITE, National Institute of Technology and Evaluation, Japan; DSMZ, Deutsche Sammlung von Mikroorganismen und Zellkulturen, Germany; JGI, Joint Genome Institute, USA; NL, the Netherlands; JCVI, J. Craig Venter Institute, USA; NZ, New Zealand.
a complex environment. Their procedure clearly shows promise for application to other wastewater-sequencing projects as well as to other environmental samples.

Analysis of the in silico proteome revealed that the primary energy source for this bacterium is derived from the fermentation of amino acids and thus the provisional name ‘Candidatus Cloacamonas acidaminovorans’ was coined. Survival under minimal oxygen concentrations appears to be due to the presence of various enzymes that give protection against oxidative stress. Moreover, predictions about the consortium composition was made knowing that C. acidaminovorans can generate H₂ and CO₂ from the oxidation of propionate, because this pathway is possible only when H₂-scavenging bacteria are present. An attempt to isolate this bacterium was made starting from the predicted nutrient needs and adding the putative syntrophic partners: unfortunately the attempt failed, probably because the real syntrophic bacteria are also non-culturable. Although it appears that this bacterial division is widespread in many anaerobic digesters, the importance of this species for the function of the digester is not clear yet.

Enhanced biological phosphorus removal (EBPR)

Phosphorus removal is one of the key steps in wastewater treatment. Low inorganic phosphate (Pᵢ) levels in treated water do not allow algal growth and thereby prevent eutrophication of water systems. The possibility to use so-called PAOs (polyphosphate-accumulating organisms) instead of more expensive chemical treatments is a strong incentive to improve our understanding of their physiology. WWTPs for enhanced biological phosphorus removal (EBPR) have an anaerobic treatment phase that precedes an aerobic phase, and selects for bacteria which accumulate large amounts of polyphosphate (Blackall et al., 2002). The consensus model of EBPR assumes that in the aerobic stage Pᵢ is removed from wastewater by uptake and storage in PAOs as polyphosphate. In the anaerobic stage, the stored polyphosphate is consumed to provide energy for uptake and transformation of volatile fatty acids (VHAs, mostly acetate and propionate) into polyhydroxyalkanoates (PHAs). From a macroscopic point of view, the objective is to let the PAOs accumulate the phosphate inside their cells and then remove the bacteria by settling in a clarifier.

For a long time, culture-dependent analyses have pointed at members of the genus Acinetobacter as the key organism for the EBPR process (Deinema et al., 1985; Streichan et al., 1990). In recent years, new molecular techniques had revealed that in fact the most important member of the PAO community, called ‘Candidatus Accumulibacter phosphatis’, in EBPR reactors belongs to the order Rhodocyclales which has no cultured representatives (Hesselmann et al., 1999; Crocetti et al., 2000). Now, this organism has been highly enriched (60–80% of biomass) in two independent lab-scale reactors seeded with sludge from EBPR plants in the USA and Australia (Garcia Martin et al., 2006). Metagenome analysis of these two samples allowed assembly of a nearly complete composite genome of A. phosphatis, and indicated that the dominant strains in both reactors are closely related (> 95% identity at nucleotide level) (Garcia Martin et al., 2006). A comprehensive metabolic reconstruction was performed, and the EBPR-relevant anaerobic metabolism inferred from the genome is shown in Fig. 4. Moreover, gene sets for N fixation, CO₂ fixation and high-affinity Pᵢ transporters were identified, suggesting that A. phosphatis is also well adapted to nutrient-limited habitats such as freshwater. Cobalt dependence was inferred from the presence of cobalamin biosynthesis genes, allowing an enrichment of this species by growing in an N-free, cobalt-rich medium. Finally, a gene-centric analysis of all metagenome sequences obtained from the wastewater plant sample revealed an over-representation of genes related to the EBPR process (e.g. phosphate transporters, cobalt transporters and VFA-handling genes).

Metaproteomics studies of sludge microbes from a similar EBPR reactor in the UK confirmed the activities of these EBPR processes, but also highlighted the importance of fatty acid cycling, denitrification and the glyoxylate bypass in EBPR (Wilmes et al., 2008a,b). Many proteins were identified with the help of the template composite A. phosphatis genome (Garcia Martin et al., 2006), but numerous protein variants were also found, which presumably came from related species and strains, highlighting the likely importance of genetic variation for overall community homeostasis. These insights should provide leads to optimize consortia and maintain stable performance of EBPR systems.

Nitrogen removal

Nitrogen elimination from wastewater involves three main processes: (i) nitrification, the oxidation of ammonia (NH₃) via nitrite (NO₂⁻) to nitrate (NO₃⁻), (ii) denitrification, the reduction of nitrate to nitrogen gas (N₂), and (iii) anaerobic ammonium oxidation (anammox), the direct combination of ammonia and nitrite into N₂. Nitrogen gas is released to the air and thus removed from the wastewater.

Metagenome sequencing was used to decipher the 4.2 Mb genome of the uncultured anammox bacterium Kuenenia stuttgartiensis; this bacterium dominated the complex community in a laboratory bioreactor 1 year after inoculation with wastewater sludge (Strous et al., 2006). More than 200 genes relevant to anammox catabolism and respiration were detected in the genome, and con-
firmed earlier hypotheses that hydrazine (N₂H₄) and hydroxylamine (NH₂OH) are intermediates in anammox catabolism. Now the question is how to culture this bacterium more rapidly, as controllable anammox is a highly desirable process in N removal. Analysis of the genome for nutrient requirements may provide the answers for overcoming their slow grow rate.

Nitrification is a two-step process, catalysed sequentially by ammonia-oxidizing bacteria (AOB) and nitrite-oxidizing bacteria (NOB). Nitrifying bacterial communities in flocs and biofilms in WWTPs are extremely complex and elucidation of their composition and dynamics offers unique opportunities for studying interactions, competition and niche differentiation (Juretschko et al., 1998; Daims et al., 2006a,b; Maixner et al., 2006; Moussa et al., 2006) (Fig. 5). A few genome sequences of some culturable members are available now (Table 1). The ammonia-oxidizing *Nitrosomonas eutropha* was found to have many genes for heavy metal and drug resistance, which could explain its presence in a polluted environment such as a WWTP (Stein et al., 2007). Members of the genus *Nitrosomonas* and *Nitrobacter* were long believed to be the main nitrifiers in wastewater treatment, but culture-independent techniques have now revealed a large diversity of yet uncultured and poorly characterized AOB and NOB, and suggest that the more important members may in fact be *Nitrosococcus* and *Nitrospira* (Juretschko et al., 1998; Moussa et al., 2006). An environmental metagenomics approach was used to sequence, identify and assemble a 137 kb genome fragment of ‘*Candidatus Nitrospira defluvii*’, which had been enriched from activated sludge from a domestic WWTP (Maixner et al., 2008). This contig, harbouring a 16S rRNA gene of *N. defluvii*, was surprisingly found to encode a chlorite dismutase, which converts chlorite (ClO₂⁻) to chloride (Cl⁻) and O₂. Enzymatic tests then
confirmed that this highly active chlorite dismutase is expressed in situ in Nitrospira, indicating that this nitrite oxidizer may be involved in bioremediation of perchlorate ($\text{ClO}_4^-$) and chlorite. Survival of this species in a nitrifying WWTP may depend on this somewhat unique activity, as (per)chlorate-containing fertilizers, disinfectants and bleaching agents are all discharged into the environment. Furthermore, hypochlorite ($\text{ClO}^-$) is often added to activated sludge to reduce the outgrowth of filamentous bacteria (Seka et al., 2003).

Understanding biofilm formation and flocculation

Good settling properties of activated sludge are important for the separation of sludge and treated wastewater. Setting problems, such as bulking and foaming, are often caused by excessive growth of filamentous bacteria (Martins et al., 2004). When the formation of so-called ‘flocs’, generally regarded as the centre of the desired metabolic reactions, is not controlled, this can result in excessive foaming and problems in oxygenation and settling. To better understand cell–cell interactions in wastewater treatment, the 4.2 Mb genome has been sequenced of the low-abundant bacterium Acidovorax temperans strain CB2, a model organism for biofilm and floc formation (Turner et al., 2008). Acidovorax temperans is present in the wastewater environment in microcolonies or as single cells. It exhibits two distinct morphological colony types with different properties in terms of biofilm and floc formation; for instance, it would appear that biofilm formation leads to poor settling ability. Data mining of the genome suggests that cell–cell interactions in biofilm formation are mediated by numerous factors, including nutritional signals, extracellular DNA and cell structures such as pili and exopolysaccharides.

Future focus

Clean water is one of the most important aspects of life. Its availability is sometimes taken for granted, but when rivers and lakes turn eutrophic action is required. In our ‘developing world’ we need to focus on avoiding these problems. Opening the black box of biological wastewater treatment is not the daunting experience that is once was. The viable but non-culturable is no longer the non-sequencable, non-characterizable enigma that existed at one time. Metagenomics is leading the way for more specific studies in related fields. New microbial species can be named and classified, while associations, dependencies and hierarchies can be described. New intriguing enzymes and metabolic pathways will be discovered, and interactions between consortium members will be unravelled. Many bacterial and archaeal divisions will be studied in greater detail, and it is expected that many of them will become culturable, allowing a better use in the construction of bioreactors specific for different types of wastewater treatment. This new knowledge may someday lead to starter cultures for wastewater treatment in the same manner as yogurt and cheese starter cultures for the dairy industry. And finally, genomic studies of wastewater treatment microbes, apart from their biotechnological applications, are also an excellent test field for a variety of other burning ecological and environmental questions.
Acknowledgements

We thank Eric Pelletier, Denis Le Paslier, Holger Daims and Abdelghani Sghir for permission and contributing their figures for re-use, and Greer Wilson for reading and correcting the manuscript. This project was carried out within the research programme of the Kluiver Centre for Genomics of Industrial Fermentation which is part of the Netherlands Genomics Initiative/Netherlands Organization for Scientific Research.

References


