Título: Taxonomía y función de organismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales

Autora: Licenciada Patricia Bovio Winkler

Datos institucionales: Facultad de Ciencias, Universidad de la República

Montevideo

2021

Ministerio
de Educación
y Cultura

AGENCIA NACIONAL
DE INVESTIGACIÓN
E INNOVACIÓN

Título: Taxonomía y función de organismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales

Autora: Licenciada Patricia Bovio Winkler

Tesis presentada con el objetivo de obtener el título de Doctor en Biotecnología otorgado por Facultad de Ciencias, Universidad de la República

Directoras de tesis:

- Prof. Agregado, Dra. Claudia Etchebehere

Laboratorio de Ecología Microbiana, Instituto de Investigaciones Biológicas Clemente Estable, Uruguay.

- Prof. Asociado, Dra. Angela Cabezas

Laboratorio de Microbiología ambiental y biotecnología, Instituto Tecnológico Regional Centro Sur, Universidad Tecnológica, Durazno, Uruguay.

Tribunal:

- Prof. adjunto en investigación, Dr. Javier Menes

Laboratorio de Ecología Microbiana Medioambiental, Facultad de Química, Universidad de la República, Uruguay

- Prof. Asociado en investigación, Dr. Andrés Iriarte

Departamento de Biotecnología, Instituto de Higiene, Facultad de Medicina, Universidad de la República, Uruguay

- Prof. Dra. Valeria Maia Merzel

Centro Multidisciplinario de Investigaciones Químicas, Biológicas y Agrícolas (CPQBA), Universidad Estatal de Campiñas (UNICAMP), Brasil

Lugar y fecha: Instituto de Investigaciones Biológicas Clemente Estable, Montevideo, Abril 2021

Agradecimientos

A mis tutoras de tesis Claudia y Angela, por el apoyo, las oportunidades, las discusiones, la guía constante, por la calidad humana y por transmitirme la pasión por la ciencia.

A Leonardo y Leandro, por abrirme las puertas y facilitarme el acceso al mundo de la metagenómica.

Al Instituto de Investigaciones Biológicas Clemente estable, donde además aprendí que: “Con ciencia grande no hay país pequeño” .

A las agencias financiadoras de mis becas de posgrado y proyectos en los cuales participé: ANII, SNI y CAP.

A la Facultad de Ciencias y al posgrado en Biotecnología.

A mis compañeras y amigas del Laboratorio de Ecología Microbiana por su apoyo y compañía, y por los hermosos after office.

A los compañeros de BIOGEM, por los gratos momentos compartidos.

A los colegas de BioPROA, en especial a Cecilia por su apoyo y buenos consejos.

A mi familia y los tremendos amigos que me dio la vida.

Y a mi vieja, por su apoyo incondicional.

Sin todos ellos, nada hubiese sido posible.

Índice de figuras y tablas

Figuras

Capítulo 1

Figura 1.1. Esquema típico de un sistema de tratamiento aerobio de lodos activados (compuesto por el tanque de aireación y el clarificador secundario) (Sancho et al. 2019).

Figura 1.2. Micrografías de contraste de fases de la estructura de flóculos: a) “Pocas”, b) “Raras” (Deepnarain et al. 2020)

Figura 1.3. Esquema del proceso de digestión anaerobia (Cabezas et al. 2015a).

Figura 1.4. Esquema de tres configuraciones distintas de reactores anaerobio: A) Reactor UASB (Chong et al. 2012), B) Reactor EGSB (Chan et al. 2009), C) Reactor IC, modificado de (Su et al. 2020)

Figure 1.5. Micrografía electrónica de barrido de gránulos enteros de reactores UASB (A, B, C) y de la superficie de los gránulos (E, D, F) (Yamada et al. 2005a)

Figura 1.6. Esquema global del ciclo del nitrógeno (Cardenas et al. 2018)

Figura 1.7. Gránulos pertenecientes a un reactor anammox con nitritacion parcial (PNA) (Speth et al. 2016)

Figura 1.8. Árbol filogenético del filo Chloroflexi a partir del gen de ARN ribosomal de 16S de representantes cultivados y *Candidatus* (Yang et al. 2020)

Figura 1.9. Micrografías de contraste de fases de la estructura de flóculos: a) Presencia de bacterias filamentosas “excesiva”, b) “Dominantes” (Deepnarain et al. 2020)

Figura 1.10. Gránulos de un reactor UASB: A) Con estructura compacta, B) Con sobrecrecimiento de bacterias filamentosas del filo Chloroflexi, C) Gránulo con sobrecrecimiento de bacterias filamentosas (Sekiguchi et al. 2001a)

Figure 1.11. Fotografías de los gránulos anammox: a) y b) flotación de gránulos, d) gránulos con aumento de volumen (esponjamiento), y e) gránulos compactos (Song et al. 2017).

Figura 1.12. Esquema de distintas preguntas que pueden ser formuladas en un trabajo de investigación y las técnicas moleculares que pueden ser utilizadas para su abordaje (Cabezas et al. 2015b).

Figura 1.13. Esquema del procedimiento de la técnica de FISH (Moter and Gobel 2000).

Figura 1.14. Químicas de detección más usadas en q-PCR: (A) SYBR Green I y (B) sondas TaqMan (Kim et al. 2013)

Figura 1.15. Esquema de la técnica de T-RFLP (Walker et al. 2017)

Figure 1.16. Esquema del proceso de Pirosecuenciación (Heather and Chain 2016)

Figura 1.17. Esquema del proceso Illumina (Heather and Chain 2016)

Figura 1.18. Esquema simplificado de los principales pasos en el análisis de metagenomas. Los rectángulos indican resultados, y los óvalos representan pasos del procesamiento (Jünemann et al. 2017). WGS= whole generation sequence

Capítulo 2

Figure 2.1 Quantification of members of phylum Chloroflexi in samples taken over time (days) for each reactor by specific q-PCR. Each point represents the mean value from duplicate reactions. The error bars represent the standard deviations. The data were presented from each reactor a AL, b COA, c COB, d SR, e MO

Figure 2.2 Abundance of Chloroflexi phylum in the five studied reactors. BoxPlots of Chloroflexi 16S rRNA gene copy numbers (log 10 copy number per ng of DNA) from each reactor are presented. The bottom and the top of each box present the first and third quartiles, the band inside is the median and the whiskers (short horizontal lines) represent the minimum and maximum value of each data set. Values outside the inner fences are shown as circles; values further than 3 times the box height from the box (the "outer fences") are shown as stars

Figure 2.3 a Specific Chloroflexi T-RFLP analysis of reactor samples taken along operation time. Relative abundances of T-RFs are indicated in different colors. Only T-RFs with relative abundance higher than 3 % for each sample were considered. (■) 52 (■) 58 (■) 70 (■) 79 (■) 87 (■) 140 (■) 148 (■) 161 (■) 167 (■) 182 (■) 200 (■) 272 (■) 369 (■) 447 (■) 457 (■) 470 (■) 501 (■) 511 (■) 520 (■) 625 (■) 633 b NMDS ordination using T-RFLP data (●) AL (●) COA (●) COB (●) SR (●) MO. Community similarity is based on the Bray–Curtis index which includes the presence/absence as well as the relative abundances of T-RFs. The stress value of 0.16 indicates good quality of the NMDS result.

Figure 2.4 Taxonomic distribution according to the 16S rRNA gene amplicon sequencing analysis. One sample from each reactor were analyzed. The results were presented as: a at phylum level: (■) Acidobacteria (■) Aminicenantes (■) Bacteroidetes (■) Caldiserica (■) Chlamydiae (■) Chloroflexi (■) Fibrobacteres (■) Firmicutes (■) Hydrogenedentes (■) Lentisphaerae (■) Omnitrophica (■) Parcubacteria (■) Planctomycetes (■) Proteobacteria (■) Spirochaetae (■) Synergistetes (■) Tenericutes (■) Thermotogae (■) Verrucomicrobia (■) Others lower than 1%. b at class level within the phylum Chloroflexi: (■) Anaerolineae (■) Caldilineae (■) Chloroflexia (■) Dehalococcoidia (■) Gitt-GS-136 (■) JG30-KF-CM66 (■) JG37-AG-4 (■) KD4-96 (■) Ktedonobacteria (■) SBR2076 (■) SHA-26 (■) SJA-15 (■) SJA-68 (■) Thermoflexia (■) Thermomicrobia (■) TK10 (■) unc. Bacteria. c at OTU level

within the phylum Chloroflexi: (■) OTU271 (■) OTU299 (■) OTU304 (■) OTU371 (■) OTU677 (■) OTU733 (■) OTU928 (■) OTU1228 (■) OTU1353 (■) OTU1407 (■) OTU1632 (■) OTU1649 (■) OTU1650 (■) OTU1668 (■) OTU1821 (■) OTU2086 (■) OTU2360 (■) OTU2545 (■) OTU2612 (■) OTU2616 (■) OTU2620 (■) OTU2814 (■) OTU3200 (■) OTU3208 (■) OTU3509 (■) OTU4169 (■) OTU4170 (■) Others lower than 2%. The sequences showing an abundance below 1.0% for phyla a and 2% for OTUs c in all samples were grouped into 'Others'

Figure 2.5 Phylogenetic distance tree of the bacterial phylum Chloroflexi constructed based on 16S rRNA gene sequence, including environmental clones, cultured representatives belonging to the phylum Chloroflexi and OTU sequences classified within the phylum Chloroflexi (●) AL (●) COA (●) COB (●) SR (●) MO. Bootstrap values (greater than 50%) obtained with the maximum likelihood method based on 1000 replications are shown at branch nodes. *Thermus aquaticus* YT-1 was used as outgroup. Bar indicates 0.05 substitutions per site. The relative abundance of the OTUs in the samples is show in brackets. Each reactor was identified with a different label. Only OTUs higher than 5% was used in the analysis

Figure 2.6 Micrographs taken from samples from COB and MO by confocal laser scanning microscopy (digitally produced images with pseudocolors). The samples were simultaneously hybridized with probes Cy3-labeled directed to the phylum Chloroflexi GNSB941 and CFX1223R (green), and DAPI staining (blue) (bar, 10 μm) a COB and b MO. Arrows indicate different types of filaments detected in each reactor

Capítulo 3

Figure 3.1. Boxplot of the total community for STARs and WTARs showing the phyla with a relative abundance higher than 2% in at least one reactor.

Figure 3.2. PCoA analysis and PERMANOVA test of the total community. The different colors indicate the different wastes or wastewaters used to fed the reactors. The PERMANOVA test results are shown at the top of the figure (see Table 3.3).

Figure 3.3. Boxplot of the Chloroflexi phylum for STARs and WTARs showing the relative abundance at: (A) class level (showing all the classes present in the systems); (B) genus level (showing the 25 most abundant genus which belong to the Anaerolinea class).

Figure 3.4. PCoA analysis and PERMANOVA test of the phylum Chloroflexi populations, the different colors indicate the different hypervariable regions. The PERMANOVA test results are shown at the top of the figure (see Table 3.3).
Figure 3.5. Boxplot of the Chloroflexi phylum for STARs and WTARs showing the relative abundance at species level (25 most abundant).
Figure 3.6. Phylogenetic tree of the phylum Chloroflexi inferred from the 16S rRNA gene sequences. The tree was reconstructed using the ML method and the GTR model. ML bootstrap values greater than or equal to 70% are shown at each node. Bar, 0.01 substitutions per site. The reference OTUs with a relative abundance greater than 5% in at least one reactor are represented by their GenBank accession numbers. Only OTUs from the core microbiome are indicated in color: from STARs are in blue color, from WTARs are in red color. Thermotoga spp. sequence served as an outgroup for rooting tree.

Capítulo 4

Figure 4.1. Taxonomic composition of the microbial communities in the samples taken from the three systems according to the 16S rRNA gene amplicon sequence. A) Relative abundance of the different phyla B) Relative abundance of the Anaerolineae class within the phylum Chloroflexi (only reads classified to the Anaerolineae class were observed in the Chloroflexi phylum).
Figure 4.2. Heatmaps showing the relative abundance of members of the phylum Chloroflexi classified at A) genus level, B) ASV level according to the 16S rRNA genus analysis
Figure 4.3. Fluorescence *in situ* hybridization (FISH) micrographs: (A) MO1, MO2, and MO3, (B) RH4, RH5 and RH6, (C) AMX. 100X Scale bar = 10 µm.
Figure 4.4. Phylogenetic analysis of Chloroflexi based on the metagenomics results and the 16S rRNA gene sequence analysis: A) Phylogenomic tree constructed with the 17 genomes from this work and reference genomes retrieved from the NCBI, B) Phylogenetic tree constructed with the most abundant ASVs retrieved from the 16s rRNA gene amplicon sequences (with abundance higher than 5% in at least one sample) and the sequences of the 16S rRNA gene from 10 MAGs from this work, and reference sequences of the phylum Chloroflexi retrieved from NCBI. The tree was reconstructed using the ML method and the GTR model. ML bootstrap values greater than or equal to 70% are shown at each node. Bar, 0.01 substitutions per site. The ASV are colored by reactor: from RH are in blue color, from MO are in red color, from AMX are in green color. Four species from four genus of phylum Three sequences from Thermotoga phylum were used outgroup for rooting trees.

Figure 4.5. Heatmap showing the presence or absence of each metabolic pathway in each genome. Pathway absences are colored with gray. The presence of the different pathways is colored with different colors.
Figure 4.6. Composite metabolic overview of the 17 Chloroflexi genomes from RH, MO and AMX based on identified genes and pathways.

Tablas

Capítulo 1

Tabla 1.1. Compuestos contaminantes según el sector agroindustrial. Fuente: https://www.dinama.gub.uy/
Tabla 1.2. Tipos de tratamiento de aguas residuales. Fuente: https://opus.lib.uts.edu.au/
Tabla 1.3. Resumen de cepas aisladas y genomas ensamblados de Chloroflexi. En violeta genomas ensamblados. Amarillo genomas obtenidos de reactores de lodos activados. Verde genomas y aislamientos en reactores anaerobios y digestores.

Capítulo 2

Table 2.1 Characteristics of isolated strains from species belonging to the classes *Anaerolineae*, *Caldilineae* and *Dehalococcoidia* within the phylum *Chloroflexi*
Table 2.2 Operational parameters of the five UASB reactors
Table 2.3 Primers and probes used in this study for q-PCR, T-RFLP, 454-pyrosequencing, Ion Torrent sequencing and FISH
Table 2.4 Pearson Linear correlation parameters between abundance of the phylum *Chloroflexi* and reactor's operational data. The data analysis was performed using PAST Software

Capítulo 3

Table 3.1. Studies included in this meta-analysis and operational parameters
Table 3.2. *In silico* primer coverage prediction for the 10 primers sets used in the studies restricted to the phylum Chloroflexi in SILVA Test Prime analysis (SSU Ref 138 NR). Color red indicates higher primer coverage, while green color indicates lower primer coverage. Numbers indicate the primer coverage as a percentage of the total sequences of the database detected by the primer set.

Table 3.3. PERMANOVA (adonis) and multivariate homogeneity of groups dispersions analysis (betadisper) results. Groups that satisfy the homogeneity group condition to perform PERMANOVA test are highlighted in grey. P values were corrected for multiple tests using the Benjamini-Hochberg correction factor. Substrate refers to the feeding according to Table 1

Table 3.4. Minimum Core microbiome obtained for STARs and WTARs. The accessing number of the sequences and the classification at the level of family, genera and species is shown. NA: not assigned

Capítulo 4

Table 1. Characteristics of the studied reactors

Table 2. Statistics overview of the Chloroflexi MAGs. Genomes with complete 16S rRNA sequences are colored in grey.

Lista de abreviaciones

DBO: Demanda Biológica de Oxígeno

DQO/COD: Demanda Química de Oxígeno/Chemical Oxygen Demand

TRH/HRT: Tiempo de Retención Hidráulico/Hydraulic Retention Time

UASB: Up-flow Anaerobic Sludge Blanket Reactor

IC: Internal Circulation

EGSB: Expanded Granular Sludge Bed Reactor

CSTR: Continuously Stirred Tank Reactors

AGV/VFA: Ácidos Grasos Volátiles/Volatile Fatty Acid

ARNr/rRNA: ARN Ribosomal/Ribosomal RNA

FISH: Fluorescent *In situ* Hybridization

OLR: Organic Loading Rate/Carga Orgánica

OTU: Operational Taxonomic Unit

Q-PCR: Quantitative Polymerase Chain Reaction

RDP: Ribosomal Database Project

Resumen de la tesis

El filo Chloroflexi se encuentra frecuentemente en alta abundancia en distintos tipos de sistemas de tratamiento de aguas residuales tanto aerobios como anaerobios, de remoción de carbono y nitrógeno. A pesar de su alta abundancia no se conoce aún la importancia de estos organismos en estos sistemas. Debido a la gran dificultad de aislarlos se los considera que forman parte de la "dark matter", la fracción de seres vivos que aún no se conocen. Se ha propuesto que podrían tener un rol importante en la formación de gránulos y flóculos en los distintos tipos de sistemas de tratamiento ya que tienen morfología filamentosa, generando la matriz inicial a la cual se van adhiriendo otros microorganismos. También se ha propuesto que podrían tener un rol como "scavengers" degradando restos celulares. Por otro lado, su sobrecrecimiento, principalmente en sistemas de lodos activados, ha sido reportado como agente causante de episodios de bulking. En dichos episodios, el sobrecrecimiento de bacterias filamentosas causa la interconección de los flóculos, disminuyendo su sedimentabilidad y causando graves problemas operacionales. Debido a la dificultad en aislarlos en cultivo puro, este grupo de microorganismos se ha estudiado mediante distintas técnicas de biología molecular para intentar profundizar en el conocimiento sobre su metabolismo, y su rol en estos ecosistemas.

Una de las líneas de investigación del Laboratorio de Ecología Microbiana del Departamento BIOGEM del IIBCE, es el estudio de las comunidades microbianas de sistemas de tratamiento de aguas residuales y residuos sólidos. Dentro de estos estudios se ha detectado una gran abundancia de los organismos del filo Chloroflexi en prácticamente todos los sistemas estudiados (reactores anaerobios, aerobios y de remoción de nitrógeno). Hasta el momento, se conoce muy poco sobre este grupo de microorganismos por lo cual el objetivo de esta tesis fue profundizar sobre la ecofisiología de estos organismos desconocidos.

En estas tesis se planteó como primer objetivo estudiar la diversidad, abundancia y morfología de organismos del filo Chloroflexi en cinco reactores metanogénicos escala real instalados en Uruguay. Tres de estos reactores tenían operación estable hace más de 5 años, y dos de los reactores fueron estudiados desde la etapa de arranque. Para esto se utilizaron distintas técnicas de biología molecular: T-RFLP, q-PCR, FISH y secuenciación masiva del gen de ARN ribosomal de 16S. Los resultados mostraron que el filo Chloroflexi presentaba morfología filamentosa y una alta abundancia en la mayoría de los reactores. En todos ellos predominó la clase Anaerolineae con una gran diversidad, pero sin representantes cultivados. También se observó que los reactores con una operación estable durante años y con biomasa

granular, tenían una mayor abundancia del filo Chloroflexi, que aquellos estudiados desde el periodo de arranque, los cuales presentaban biomasa flocular.

Como segundo paso se decidió ampliar el estudio a reactores metanogénicos a escala real distribuidos globalmente utilizando datos de secuencias depositados en los bancos de datos. El segundo objetivo de la tesis fue entonces estudiar la abundancia y diversidad del filo Chloroflexi en reactores metanogénicos a escala real distribuidos globalmente. Este trabajo consistió en la recopilación de sets de datos de secuenciación masiva del gen de ARNr 16S publicados en 17 trabajos científicos, sumando un total de 62 reactores. Los datos crudos fueron procesados mediante herramientas bioinformáticas, abordando la complejidad de los datos como son la utilización de distintas plataformas de secuenciación, distintos sets de primers y distintas profundidades de secuenciación. Los resultados mostraron que la clase Anaerolineae fue la predominante en todos los reactores metanogénicos, presentando mayor abundancia en los reactores de tratamiento de aguas residuales con biomasa granular, y menor abundancia en reactores de residuos sólidos los cuales presentan en general biomasa dispersa. En ambos tipos de reactor, predominaron 4 géneros de la clase Anaerolineae, pero presentaron distintas especies, en todos los casos sin representantes cultivados. Los resultados de los primeros dos objetivos subrayan la importancia de obtener aislamientos en cultivo puro o genomas ensamblados a partir de metagenomas, para llenar el vacío filogenético y funcional de este grupo de microorganismos sumamente abundante y diverso en sistemas de tratamiento de aguas residuales.

Se decidió entonces profundizar en la función de estos microorganismos. El tercer objetivo de la tesis fue estudiar la diversidad taxonómica y el potencial metabólico de estos microorganismos en tres sistemas diferentes de tratamiento de aguas residuales: sistemas aerobios de lodos activados, reactores metanogénicos y reactores anammox. Se pretendía determinar si existían diferencias en los genomas de los organismos de los distintos sistemas de tratamiento de aguas residuales. Para esto se secuenciaron metagenomas de cada uno de los sistemas, se ensamblaron genomas de Chloroflexi y posteriormente se realizaron análisis filogenómicos y metabólicos. A partir de estos análisis se lograron obtener 17 nuevas especies del filo Chloroflexi, donde más del 90% de ellas pertenecieron a la clase Anaerolineae. Estas especies se distribuyeron en varios géneros, mostrando la gran diversidad presente en los sistemas de tratamiento. A pesar de su gran diversidad, pudimos determinar que eran funcionalmente redundantes, ya que todas las especies presentaron el potencial de ser anaerobias estrictas e hidrolizar y fermentar compuestos simples y complejos, lo que incluye el potencial de degradar restos celulares. Por otro lado, se observó que los genomas

ensamblados a partir de muestras de los reactores anammox presentaban más genes relacionados a la reducción y oxidación de compuestos nitrogenados, que los genomas de los otros reactores. Por lo tanto, en los sistemas anammox el filo Chloroflexi podría estar colaborando en la remoción de nitrógeno. En cuanto al rol del filo Chloroflexi en la formación de flóculos y gránulos, pudimos determinar que la mayoría de los genomas presentaba propiedades de adhesividad. Dicha adhesividad en conjunto con la morfología filamentosa podría tener un papel crucial en la formación de gránulos y flóculos.
En esta tesis se profundizó en dos aspectos muy importantes del filo Chloroflexi, por un lado, se generó conocimiento sobre su diversidad y abundancia en una amplia variedad de reactores metanogénicos. Por otro lado, se generó conocimiento en el potencial metabólico y rol estructural de nuevas especies en distintos sistemas de tratamiento de aguas residuales. Conocer la taxonomía y metabolismo de organismos no cultivados en sistemas de tratamiento de aguas residuales ayudará a mejorar los procesos de remoción de materia orgánica y nitrógeno, y a mantener las condiciones de operación alejadas de aquellas que podrían ocasionar el sobrecrecimiento de bacterias filamentosas. Si bien mediante las nuevas herramientas de biología molecular es posible inferir los potenciales roles metabólicos de las comunidades, son necesarios más estudios que confirmen estas funciones como puede ser la metatranscriptómica. Por otro lado, siguen siendo imprescindibles los aislamientos en cultivo puro, y queda mucho por investigar sobre la interacción de los microorganismos en las complejas comunidades presentes en estos sistemas. En este trabajo se generó conocimiento básico de Microbiología aplicado a sistemas biológicos de sistemas de tratamiento de aguas residuales y residuos sólidos.

Palabras claves: Chloroflexi, tratamiento de aguas residuales, metanogénesis, lodos activados, reactores anammox, comunidades microbianas, metagenómica, genomas, secuenciación masiva

Tabla de contenido

Agradecimientos 4
Índice de figuras y tablas 5
Figuras 5
Tablas 9
Resumen de la tesis 11
CAPÍTULO 1 18
INTRODUCCIÓN GENERAL 18
1.1 Generación de aguas residuales en Uruguay 19
1.2 Sistemas biológicos de tratamiento de aguas residuales 20
1.2.1 Reactores aerobios 21
1.2.1.1 Floculación de la biomasa en sistemas de lodos activados 23
1.2.2 Sistemas anaerobios 23
1.2.2.1 Microbiología de la digestión anaerobia 24
1.2.2.2 Reactores anaerobios de tratamiento de aguas residuales 26
1.2.2.3 Granulación de la biomasa 29
1.2.3 Sistemas de remoción de Nitrógeno 30
1.2.3.1 Reactores anammox 31
1.3 El filo Chloroflexi 32
1.3.1 Relevancia de la clase Anaerolineae en sistemas de tratamiento de aguas residuales 34
1.3.2 Rol de organismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales 36
1.3.3 Problemas de bulking ocasionados por sobrecrecimiento de bacterias filamentosas del filo Chloroflexi 37
1.3.3.1 Bulking en reactores de lodos activados 38
1.3.3.2 Bulking en reactores anaerobios tipo UASB 40
1.3.3.3 Bulking en reactores anammox 41
1.4 Métodos moleculares aplicados al estudio de comunidades microbianas en sistemas de tratamiento 42
1.4.1 Hibridación *in situ* Fluorescente (FISH) 44
1.4.2 Reacción en cadena de la polimerasa cuantitativa (q-PCR) 45
1.4.3 Terminal restriction fragment length polymorphism analysis (T-RFLP) 46
1.4.4 Técnicas de secuenciación masiva de segunda generación 47
1.5.5 Procesamiento de datos obtenidos mediante secuenciación masiva de amplicones 49
1.5.6 Aplicaciones de la secuenciación masiva de amplicones del gen de ARNr 16S para el estudio de comunidades microbianas 50
1.5.7 Procesamiento de datos obtenidos mediante secuenciación de metagenomas 51
1.5.8 Aplicaciones de la Metagenómica al estudio de comunidades microbianas 52

1.5 Marco del trabajo y objetivos 53
1.5.1 Marco del trabajo: persistencia de microorganismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales 53
1.5.2 Objetivo general 55
1.5.3 Objetivos específicos 55
1.5.4 Estrategia de investigación 55
1.6 Referencias introducción general 57
CAPÍTULO 2 72
Diversidad y abundancia del filo Chloroflexi en sistemas de tratamiento metanogénicos de Uruguay 72
Abstract 73
2.1 Introduction 74
2.2 Materials and methods 76
2.2.1 UASB methanogenic reactors and sample collection 76
2.2.2 DNA extraction 77
2.2.3 Terminal restriction fragment length polymorphism (T-RFLP) analysis 78
2.2.4 Quantitative PCR 79
2.2.5 Fluorescence *in situ* hybridization (FISH) 80
2.2.6 Bacterial 16S rRNA gene analysis by next generation sequencing 80
2.2.7 Bioinformatic processing 81
2.2.8 Phylogenetic analysis of selected reads 82
2.3 Results 82
2.3.1 Chloroflexi were abundant and persistent in UASB reactors 82
2.3.2 High dominance and stability of a particular Chloroflexi T-RF detected by specific T-RFLP 85
2.3.3 Uncultured Chloroflexi dominate the communities 87
2.3.4 FISH revealed a filamentous morphology of Chloroflexi cells 90
2.4 Discussion 90
2.4.1 Abundance of Chloroflexi in UASB reactors 90
2.4.2 The biomass used as inoculum influences the Chloroflexi community structure and composition 92
2.4.3 Uncultured Chloroflexi were predominant in the reactors 93
2.5 References 95
2.6 Material suplementario 101
CAPÍTULO 3 105
Meta-análisis: determinación de la ecología del filo Chloroflexi en reactores metanogénicos escala real 105
Abstract 106
3.1 Introduction 107

3.2 Materials and methods 109
3.2.1 Data collection and composition 109
3.2.2 Data processing, clustering and taxonomic assignments 110
3.2.3 Analysis strategy of the sequencing data 110
3.2.4 Data analysis and visualization 111
3.2.5 Phylogenetic tree 112
3.2.6 Primer coverage rate 112
3.3 Results 112
3.3.1 In silico analysis of the coverage rate of Chloroflexi primer pairs 113
3.3.2 Microbial community structure in STARs and WTARs 114
3.3.3 Chloroflexi community composition 117
3.3.4 Phylogenetic analyses of the sequences affiliated to the phylum Chloroflexi 123
3.4 Discussion 125
3.4.1 Anaerolineae class dominate the Chloroflexi microbial communities in the anaerobic reactors studied 125
3.4.2 Four genera within Anaerolineae predominate in the full scale anaerobic reactors studied 126
3.4.3 Challenges and limitations of comparing sequences retrieved from databases generated with different primer sets and sequencing platforms 128
3.4.4 Importance of standardized metadata repositories for full scale methanogenic reactors 129
3.5 Conclusions and future directions 130
3.6 References 131
3.5 Material suplementario 137
CAPÍTULO 4 142
Genomas ensamblados a partir de metagenomas muestra la redundancia funcional de organismos no cultivables del filo Chloroflexi en reactores de lodos activados, anammox y metanogénicos 142
4.1 Introduction 144
4.2 Material and methods 146
4.2.1 Sampling and DNA extraction 146
4.2.2 Community profiling using 16S rRNA gene amplicon sequencing 147
4.2.3 Phylogenetic analysis 148
4.2.4 Fluorescence in situ hybridization (FISH) 148
4.2.5 Metagenome sequencing 149
4.2.6 Genome assembly, binning and genome annotation 149
4.2.7 Metabolic Analysis 150
4.2.7 Phylogenomic analysis 150
4.3 Results and discussion 150

4.3.1 Abundance, diversity and taxonomy of Chloroflexi members in the different wastewater systems 151
4.3.2 Morphology determined by FISH 153
4.3.3 General genomic features of Chloroflexi metagenomic assembled genomes 154
4.3.4 Phylogenomic analysis of Chloroflexi MAG 155
4.3.5 General metabolic pathways and genes according to the MAGs 158
4.3.5.1 The potential for anaerobic respiration 158
4.3.5.2 Metabolic pathways involved in hydrolysis of polymers 162
4.3.5.3 Central carbon metabolism 162
4.3.5.4 Ability to generate energy by fermentation 164
4.3.5.5 Wood–Ljungdahl pathway 165
4.3.5.6 Potential for storage of polymers and lipids 165
4.3.5.7 Nitrogen metabolism 166
4.3.5.8 Filamentous morphology, adhesion capability and exopolysaccharides production 168
4.3.6 Ecological significance and concluding remarks 169
4.4 References 172
4.5 Material suplementario 180
CAPÍTULO 5 182
DISCUSIÓN GENERAL, CONCLUSIONES Y PERSPECTIVAS 182
5.1 ¿Por qué es importante estudiar organismos del filo Chloroflexi sistemas de tratamiento de aguas residuales? 183
5.2 Los organismos del filo Chloroflexi son diferentes pero tienen la misma función en los diferentes tipos de reactores, ¿que puede implicar esto? 184
5.3 La gran mayoría de los organismos del filo Chloroflexi descripto en sistemas de tratamiento de aguas residuales tiene morfología filamentosa, ¿Esto tiene una importancia especial en estos sistemas? 185
5.4 ¿Por qué es tan difícil aislar organismos del filo Chloroflexi? 186
5.5 ¿Cuál es la limitación de cada técnica y cuál de las técnicas de estudio utilizadas sería la que más aporta? 188
5.6 Conclusiones 189
5.7 Perspectivas 190
5.8 Referencias 191

CAPÍTULO 1

INTRODUCCIÓN GENERAL

1.1 Generación de aguas residuales en Uruguay

Uruguay es un país con una extensión territorial de 176.215 km², caracterizado por un uso intensivo de la tierra para la agricultura y la ganadería, y una importante actividad industrial (López 2016). El vertido de efluentes contaminados provenientes de dichas actividades, genera un impacto negativo en el medio acuático y terrestre afectando a todo el ecosistema. La composición de las aguas residuales depende de su origen, pero en general, los principales contaminantes encontrados incluyen compuestos orgánicos, xenobióticos, metales, sólidos en suspensión, nutrientes (principalmente nitrógeno y fósforo) y patógenos (Tabla 1.1).

Tabla 1.1. Compuestos contaminantes según el sector agroindustrial. Fuente: https://www.dinama.gub.uy/

Sector industrial	Principales sustancias contaminantes
Alimenticia	Contaminantes orgánicos, grasas y aceites
Cárnica	Contaminantes orgánicos e inorgánicos
Construcción	Sólidos en suspensión, metales
Energía	Calor, hidrocarburos y químicos
Fertilizantes	Nitritos y fosfatos
Lácteas	Suero de leche y grasas
Lavaderos de lana	Grasas y aceites. Detergentes, pesticidas
Metalúrgica	Aceites, metales disueltos, emulsiones, álcalis y ácidos
Minería	Sólidos en suspensión, metales pesados, materia orgánica, pH, cianuros
Oleaginosa	Grasas y aceites
Pasta y papel	Sólidos en suspensión. Compuestos organoclorados y azufrados. Metales pesados
Pinturas, barnices y tintas	Zn, Cr, Se, Mo, Ti, Sn, Ba, Co, etc
Plaguicidas	Organohalogenados, organofosforados, biocidas
Química inorgánica	Hg, P, fluoruros, cianuos, amoniaco, nitritos, ac. sulfhídrico, F, Mn, Mo, Pb, etc
Química orgánica	Organohalogenados, organosilícicos
Textil y cuero	Cr, taninos, tensoactivos, sulfuros, grasas, sólidos en suspensión

La cantidad y calidad de las aguas residuales generadas por los procesos agroindustriales y su descarga en los cuerpos de agua naturales tales como ríos, arroyos y lagunas, son un tema de gran preocupación a nivel nacional. En este sentido, el proyecto Biovalor (https://biovalor.gub.uy/) llevado adelante por el Ministerio de Industria, Energía y Minería, Ministerio de Vivienda Ordenamiento Territorial y Medio Ambiente, y Ministerio de Ganadería Agricultura y Pesca, asesora desde 2014 a productores agropecuarios sobre la transformación de residuos orgánicos contaminantes en energía y/o subproductos aprovechables. Este proyecto promueve la generación de biogás y otros activos provenientes de residuos orgánicos en distintas partes de Uruguay, que redunda en ahorros para el sector y suman al cuidado ambiental. El punto de partida de este proyecto fue la problemática que

afronta la industria, en general, y los productores, en particular, para disponer los residuos que generan sus respectivas actividades.
Una solución ambiental a ese problema es su tratamiento y disposición en el terreno o su vertido a cursos de agua, en el caso de efluentes no contaminantes, en cumplimiento de la normativa nacional (Decreto N° 253/979). Los residuos orgánicos totales registrados en el año 2014 provenientes de la producción lechera, porcina y de engorde en corral muestran que Montevideo y Canelones generan entre 40.000 y 90.000 toneladas por año, mientras que Paysandú, Río Negro, Soriano, Colonia, San José y Florida generan entre 15.000 y 40.000 toneladas por año. Los departamentos con menor generación de residuos son Rivera, Tacuarembó, Cerro Largo, Treinta y Tres, y el departamento de Maldonado, los cuales tienen un promedio de entre 1 y 5.000 toneladas de residuos por año. Por lo tanto, resulta imprescindible una buena gestión y tratamiento de las aguas residuales y residuos sólidos para reducir al mínimo el impacto ambiental. El mejoramiento del proceso de eliminación de la materia orgánica y otros contaminantes presentes en aguas residuales es objetivo de investigación y desarrollo en el mundo.

1.2 Sistemas biológicos de tratamiento de aguas residuales

Dependiendo de la carasterística química de los contaminantes del agua residual, se pueden aplicar combinaciones de procesos físicos, químicos y/o biológicos para su tratamiento a través del tratamiento primario, secundario y/o terciario (Sonune and Ghate 2004) (Tabla 1.2).

Tabla 1.2. Tipos de tratamiento de aguas residuales. Fuente: https://opus.lib.uts.edu.au/

Tratamiento	Objetivo	Proceso	Resultado
Primario (mecánico)	Purificación preliminar para eliminar residuos suspendidos y flotantes	Cribado para atrapar compuestos sólidos	Reducción de la demanda biológica de oxígeno (DBO entre un 20% y 30%
		Sedimentación por gravedad para eliminar sólidos en suspensión.	Reducción de compuesto sólidos en suspensión entre un 50% y 60%
Secundario (biológico)	Oxidación de la materia orgánica y contaminantes a través de comunidades microbianas	Utilización de microorganismos que degradan la materia orgánica transformándola en CO_2, agua y energía para su propio crecimiento	Elimina aproximadamente el 85% de la DBO y sustancias sólidas suspendidas
	Eliminación de la materia orgánica disuelta que escapa al tratamiento primario.	El agua circula en un reactor que contiene una alta concentración de microorganismos Luego es seguido de tanques de sedimentación adicionales	
Terciario	Proceso adicional para convertir el agua residual en agua de buena calidad para diferentes usos	Métodos químicos o biológicos avanzados para eliminar nitrógeno y fósforo.	Puede eliminar el 99% de todas las impurezas
		La tecnología relacionada puede ser muy cara, muy técnica y requiere energía	Produciendo un efluente de calidad de agua potable

DBO: Demanda biológica de oxígeno

El tratamiento primario tiene como objetivo la remoción por medios físicos o mecánicos de una parte sustancial del material sedimentable o flotante. Entre los principales procesos y operaciones de tratamiento primario se destacan la sedimentación (donde los sólidos se separan por gravedad), la coagulación y la floculación, que facilitan el retiro de los sólidos suspendidos y de las partículas coloidales (Sonune and Ghate 2004).

El tratamiento secundario se basa en la eliminación biológica de la contaminación de las aguas residuales. En este sentido, actualmente se usan dos tipos de tratamiento biológico, basados en la acción de comunidades microbianas: los sistemas de tratamiento aerobios y anaerobios. Las capacidades únicas de las comunidades microbianas para degradar la materia orgánica, eliminar nutrientes y transformar compuestos tóxicos en productos inocuos, las convierten en actores esenciales en el tratamiento de aguas residuales.

1.2.1 Reactores aerobios

Los sistemas aerobios son ampliamente utilizados para el tratamiento biológico de aguas residuales. Los sistemas más convencionales son los llamados sistemas de lodos activados. Estos sistemas constan de dos etapas, una etapa biológica que se lleva a cabo en el tanque de

aireación, donde se elimina la materia orgánica por parte de las comunidades microbianas, y una etapa física que se lleva cabo en el clarificador secundario, donde la biomasa sedimenta y se separa del agua tratada (Figura 1.1).

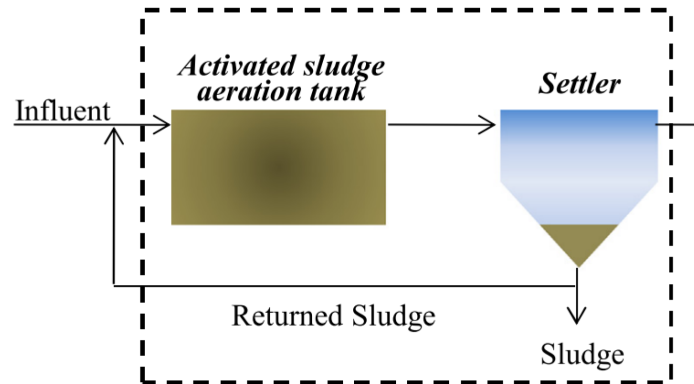

Figura 1.1. Esquema típico de un sistema de tratamiento aerobio de lodos activados (compuesto por el tanque de aireación y el clarificador secundario) (Sancho et al. 2019).

La biomasa de estos sistemas contiene una gran diversidad de virus, bacterias, protozoos, hongos, metazoos y algas (Jenkins et al. 2003a). En este complejo ecosistema, las bacterias representan alrededor del 95% de la población total de microorganismos y juegan el rol clave en la degradación de materia orgánica y nutrientes (Jenkins et al. 2003a). La ventaja de los sistemas de lodos activados es que presentan una alta eficiencia de remoción de materia orgánica (Chong et al. 2012). Una de las desventajas de estos sistemas es que requieren una gran cantidad de energía debido a que es necesaria una aireación constante. Se estima que la aireación representa entre el 30% y el 60% del consumo total de energía de una planta de lodos activados (McCarty et al. 2011). Sumado a esto, es necesario considerar el tratamiento y disposición final de los lodos, ya que se genera gran cantidad de biomasa, suponiendo entre un 30% y un 60% del costo operativo (Saby et al. 2003). Aunque hay distintas alternativas para el tratamiento del exceso de lodos, la digestión anaerobia juega un papel importante por su capacidad para transformar la materia orgánica en biogás (60-70% en volumen de metano). De este modo se reduce la cantidad de sólidos finales, la mayoría de los patógenos presentes y se eliminan los problemas de olor asociados con el lodo residual (Appels et al. 2008).

1.2.1.1 Floculación de la biomasa en sistemas de lodos activados

La buena sedimentación y compactación de la biomasa en el clarificador secundario es una condición necesaria para garantizar una buena calidad del efluente final. En los sistemas de lodos activados, la biomasa está dispuesta en forma de flóculos los cuales constituyen la unidad fundamental en estos sistemas. Los flóculos están compuestos por diversos grupos de bacterias, entre ellas, existen las bacterias formadoras de flóculos (Zooglea) y las bacterias filamentosas representadas por varios grupos filogenéticos, que junto con los polímeros extracelulares colaboran en la formación de los flóculos (Figura 1.2 (A,B)) (Örmeci and Vesilind 2000).

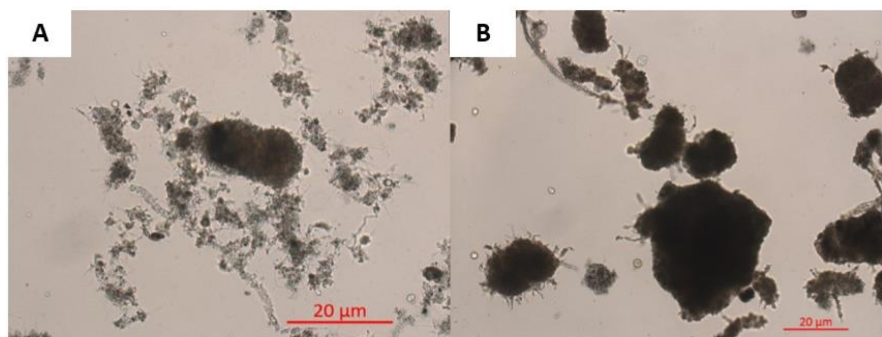

Figura 1.2. Micrografías de contraste de fases de la estructura de flóculos: A) Pocas bacterias filamentosas, B) Mínima presencia de bacterias filamentosas (Deepnarain et al. 2020).

Uno de los grupos que cumplen un rol fundamental en la formación de los flóculos son bacterias filamentosas del filo Chloroflexi, y se ha propuesto que proporcionan una estructura estabilizadora en los flóculos (Björnsson et al. 2002). La localización intraflocular de Chloroflexi y la alta abundancia relativa reportada por Kragelund et al. (2007) apoyó esta hipótesis que probablemente explica un papel importante de estos organismos en sistemas de lodos activados.

1.2.2 Sistemas anaerobios

Los reactores anaerobios se basan en el proceso de digestión anaerobia, en el cual la materia orgánica es degradada por una compleja red de microorganismos en ausencia de oxígeno, generando dióxido de carbono y metano como productos finales del proceso (Verstraete et al. 2005; Wheeldon et al. 2007). En algunos sistemas anaerobios también es posible obtener como productos finales biofertilizantes como el biosol (sólido) y el biól (líquido). El metano generado es un biogás que puede ser utilizado como fuente de energía renovable (Demirel et

al. 2005). Por lo tanto, a través de este tipo de sistema es posible generar una economía circular en los procesos productivos, convirtiendo las plantas de tratamiento en productores netos de energía (Weiland 2010). La digestión anaerobia involucra relaciones sintróficas estrechamente acopladas entre microorganismos de la comunidad (Nealson 1997; Demirel and Scherer 2008). Por lo tanto, la función general, la estabilidad y la eficiencia del proceso dependen del buen funcionamiento de estas comunidades (Schink 1997; Weiland 2010).

En Uruguay, debido a la relativa disponibilidad de tierra, comúnmente se utilizaban lagunas anaerobias para el tratamiento de aguas residuales (López 2016). Sin embargo, actualmente varias empresas tienen reactores anaerobios con distintas configuraciones (de residuos líquidos o sólidos). Algunas de ellas son: una industria procesadora de malta (Montevideo), dos industrias lácteas, la fábrica principal de lavado de lana, una granja lechera a gran escala con producción de leche en polvo, fábricas nacionales de cerveza (Montevideo), y dos plantas de tratamiento de OSE (Pando y Costa de Oro). Las restricciones medioambientales, las mejoras energéticas y nuevos conceptos como producción limpia, incentivan a las empresas a buscar tecnologías de tratamiento más eficaces y optimizadas.

1.2.2.1 Microbiología de la digestión anaerobia

El proceso de digestión anaerobia puede dividirse conceptualmente en cuatro fases definidas por las reacciones catabólicas primarias llevadas a cabo por las comunidades microbianas presentes en cada fase, éstas son: hidrólisis de polímeros complejos (I, hidrólisis), fermentación de los productos finales de hidrólisis a ácidos grasos de cadena corta (II, acidogénesis), la conversión de ácidos grasos de cadena corta de tres o más carbonos en acetato (III, acetogénesis sintrófica) y finalmente, la producción de metano (IV, metanogénesis) (Angenent et al. 2004; Yu et al. 2010b) (Figura 1.3).

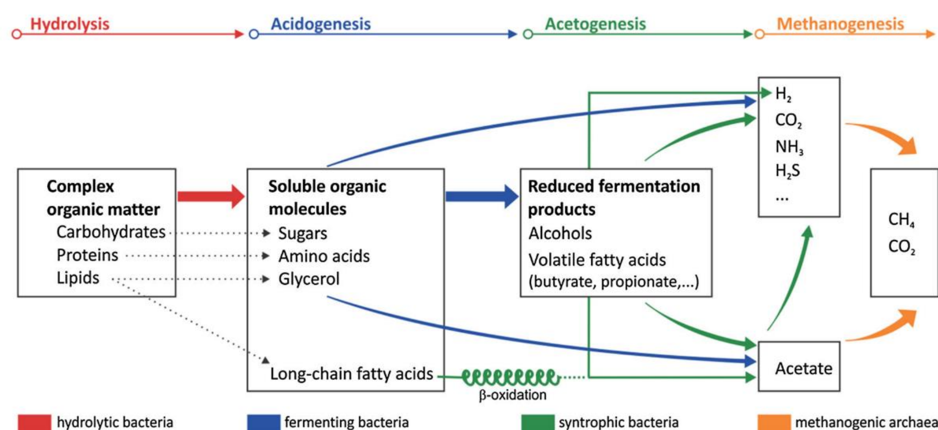

Figura 1.3. Esquema del proceso de digestión anaerobia (Cabezas et al. 2015a).

Las primeras tres etapas del proceso de digestión anaerobia las llevan a cabo diferentes grupos de bacterias, mientras que la etapa final la realizan exclusivamente Archaeas metanogénicas. Los reactores anaerobios en general están dominados por los mismos grupos filogenéticos pertenecientes al dominio bacteria (Chloroflexi, Proteobacteria, Firmicutes y Sinergistetes) (Bovio et al. 2020), y al dominio de las Archaeas (Euryarchaeotas), variando en sus abundancias relativas (Qiao et al. 2013).

En la primera etapa, la hidrólisis, los compuestos orgánicos complejos como polisacáridos, proteínas y grasas, que son en su mayoría insolubles, son degradados a monómeros solubles más simples como monosacáridos, aminoácidos y ácidos grasos. Dichos monómeros pueden penetrar la membrana celular de bacterias fermentadoras. En la etapa de hidrólisis, participan enzimas extracelulares del grupo de hidrolasas (amilasas, proteasas y lipasas) producidas por las bacterias hidrolíticas (Conrad 1999; Amani et al. 2010).

En la segunda etapa, la fermentación o etapa acidogénica, las bacterias fermentativas convierten los compuestos solubles (incluidos los productos de la hidrólisis previa), en ácidos grasos volátiles (por ejemplo: fórmico, acético, propiónico, butírico y pentanoico), alcoholes (metanol y etanol), aldehídos, dióxido de carbono e hidrógeno. Esta etapa suele ser la más rápida en todo el proceso y la llevan a cabo un grupo muy diverso de bacterias fermentadoras (Van Lier 2008)

En la tercera etapa, la acetogénesis, las bacterias acetogénicas convierten los ácidos grasos volátiles generados en la fase de acidogénesis en acetato, dióxido de carbono e hidrógeno, que serán utilizados por las Archaeas metanogénicas en la etapa final del proceso (Schink 1997). La formación de hidrógeno, tanto en la etapa de fermentación como en la de acetogénesis, es desfavorable termodinámicamente a 25 °C y 1 atm (variación de energía

libre de Gibbs positiva). Por lo tanto, la presión parcial de hidrógeno debe mantenerse en bajos niveles mediante su consumo por Archaeas metanogénicas hidrogenotróficas. Esta relación entre bacterias acetogénicas y Archaeas metanogénicas se denomina sintrofismo (Schink 1997; de Bok et al. 2005). De esta forma la concentración de hidrógeno se mantiene en niveles adecuados cuando las Archaeas metanogénicas crecen activamente (Hattori 2008). La concentración de hidrógeno en el medio y la transferencia de hidrógeno interespecies es de vital importancia para regular el proceso global en la digestión anaerobia.

La cuarta y última fase, la metanogénesis, consiste en la producción de metano exclusivamente por Archaeas metanogénicas a través de tres vías principales. Dichas vías se diferencian en el compuesto dador de electrones: vía acetotrófica/acetoclástica (utilizan el acetato), vía hidrogenotrófica (utilizan el H_2 y CO_2), y vía metilotrófica (utilizan compuestos metilados). Se ha postulado que la vía acetoclástica contribuye hasta en un 72% de la generación total de metano a partir del acetato (Gujer and Zehnder 1983), aunque este hecho tiene algunas excepciones. Por ejemplo, se ha reportado que ante una concentración elevada de amoníaco en el reactor, las Archaeas metanogénicas acetoclásticas podrían verse inhibidas por ser más susceptibles. Por lo tanto, el acetato se podría degradar a través de la vía de oxidación sintrófica del acetato (SAO), al mismo tiempo que la metanogénesis se daría por la vía hidrogenotrófica (Schnürer and Nordberg 2008; Fotidis et al. 2014). Las Archaeas metanogénicas se caracterizan por ser anaerobias estrictas y ser muy sensibles a variaciones de pH (Van Lier 2008). Es en esta etapa de la digestión anaerobia en la que la DQO (materia orgánica medida como demanda química de oxígeno) del agua residual o del residuo sólido se convierte a una forma gaseosa que abandona el reactor.

1.2.2.2 Reactores anaerobios de tratamiento de aguas residuales

Los reactores anaerobios de aguas residuales pueden dividirse en dos grandes grupos con base en el tipo de crecimiento microbiano: 1. de lecho fijo, formando biopelículas (los microorganismos se unen a un soporte inerte fijo), y 2. de crecimiento libre o suspendido (formación de gránulos a flóculos debido a las propiedades de auto-inmovilización de los microorganismos). Dentro de los reactores de biomasa no unida a un soporte, se encuentran por ejemplo los reactores CSTR (Continuously Stirred Tank Reactors), IC (Internal Circulation), EGSB (Expanded Granular Sludge Bed Reactor) y los UASB (Up-flow Anaerobic Sludge Blanket Reactor).

El reactor tipo UASB fue desarrollado en la década del 70-80 por un grupo de Holanda dirigido por el Profesor Lettinga (Lettinga et al. 1980). Esta tecnología anaerobia, tiene la ventaja de soportar un alto contenido de materia orgánica en el agua residual. El éxito de los reactores UASB puede atribuirse a su capacidad de retener altas concentraciones de biomasa, debido a que es capaz de separarla eficientemente del agua residual tratada y del biogás generado. El reactor UASB consiste en un tanque rectangular o circular, en el cual el agua residual fluye en dirección de flujo ascendente tomando contacto con la biomasa granular (altamente sedimentable), la cual ocupa alrededor de la mitad del volumen del reactor (Figura 1.4 (A)).

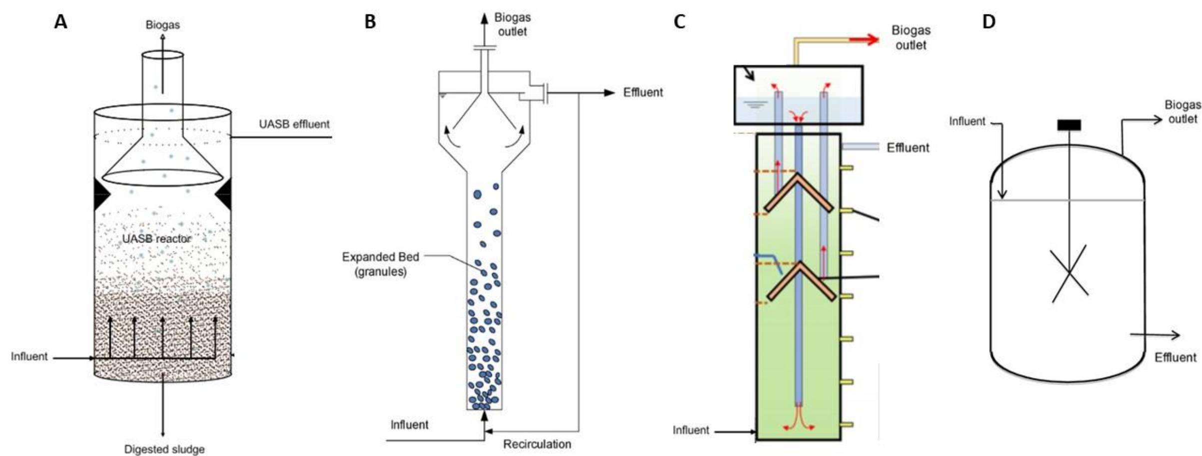

Figura 1.4. Esquema de cuatro configuraciones distintas de reactores anaerobios: A) Reactor tipo UASB (Chong et al. 2012), B) Reactor tipo EGSB (Chan et al. 2009), C) Reactor tipo IC, modificado de (Su et al. 2020), D) Reactor tipo CSTR (Merlin and Hervé 2013).

De esta forma, la materia orgánica presente en el agua residual es degradada por la comunidad microbiana, generando gas metano, el cual a su vez ayuda a la mezcla dentro del reactor, facilitando el contacto del agua residual con la biomasa (Lettinga and Hulshoff Pol 1991). La parte superior del reactor UASB contiene un separador trifásico (gas-sólido-líquido), el cual separa el agua residual tratada, del biogás generado y permite que la biomasa granular permanezca dentro del reactor (Lettinga and Hulshoff Pol 1991). La ventaja de este tipo de reactor es que permite desacoplar el tiempo de residencia hidráulica del tiempo de residencia celular, logrando una mayor concentración de biomasa, altas cargas y bajos tiempos de residencia hidráulica.

A partir del reactor UASB se han realizado modificaciones en su configuración, generando por ejemplo los reactores EGSB y los reactores IC.

Los reactores EGSB han sido desarrollados principalmente para mejorar el contacto sustrato-biomasa dentro del sistema de tratamiento, mediante la expansión del lecho de lodo con una mayor velocidad de flujo ascendente del líquido (Figura 1.4 (B)) (Chan et al. 2009). El aumento de la velocidad se logra mediante un mayor flujo de recirculación de lodos y una mayor relación entre la altura y el ancho del reactor (Von Sperling and de Lemos Chernicharo 2015). Esto mejora el rendimiento y la estabilidad del reactor.

Los reactores IC también están diseñados para mejorar la transferencia de masa. Por otro lado, es más económico ya que contiene un elevador de aire interno para lograr la circulación de líquido mixto en lugar de la circulación externa, que depende de una bomba (Figure 1.4 (C)). Por lo tanto, el reactor IC podría mantener una mayor concentración de biomasa en un tiempo de retención hidráulica menor manteniendo una alta eficiencia operativa y estabilidad, incluso para una alta tasa de carga orgánica (Luo et al. 2016).

En general los reactores UASB, EGSB e IC tienen una alta eficiencia de remoción de materia orgánica tanto a cargas altas como a temperaturas bajas; bajo consumo de energía, ya que no se necesita calentar el influente ni airear el sistema, y además se produce energía en forma de biogás (metano). Estos sistemas generan menor cantidad de biomasa comparado con sistemas de tratamiento aerobios debido a que los microorganismos anaerobios tienen tasas de crecimiento más bajas (Chong et al. 2012). Los sistemas anaerobios presentan la desventaja de tener baja remoción de patógenos y nutrientes (como nitrógeno y fósforo), por lo que en general se hace necesario un post-tratamiento (Chernicharo 2006; Khan et al. 2011). A pesar de ser sistemas de tratamiento robustos, los microorganismos involucrados en cada fase del proceso de digestión anaerobia son interdependientes a través de la alimentación cruzada y/o del mantenimiento de gradientes quimio-termodinámicos. Es por esto que el proceso puede alterarse cuando una de las cuatro fases de la digestión anaerobia está en desequilibrio, como podría suceder ante una acumulación de ácidos grasos de cadena corta, lo que generaría una acidificación de todo el sistema (Chen and Dong 2005).

Además de los reactores anaerobios antes descriptos que tratan aguas residuales, los reactores de tipo CSTR son comúnmente utilizados para el tratamiento de residuos sólidos, como puede ser el lodo excedente de plantas de tratamiento de lodos activados. La mezcla en este tipo de sistema es mediante un agitador mecánico (Figura 1.4 (D)), lo que da como resultado un sistema completamente mezclado.

1.2.2.3 Granulación de la biomasa

La clave para un proceso biotecnológico moderno de alta carga, sin importar el sistema que sea considerado, es la inmovilización adecuada de los microorganismos. Los reactores UASB y sus variantes (como EGSB e IC) se caracterizan por tener la biomasa inmovilizada en forma de gránulos compactos (Figure 1.5 (A, B, C)).

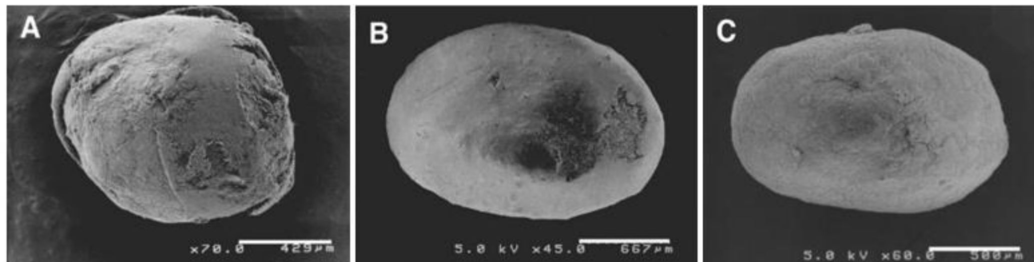

Figure 1.5. Micrografía electrónica de barrido de gránulos enteros de distintos reactores UASB (Yamada et al. 2005a).

El proceso de granulación ocurre gracias a las propiedades de auto-inmovilización de los microorganismos. La parte central del gránulo consiste en una matriz de compuestos inertes orgánicos e inorgánicos, a la cual se van adhiriendo los microorganismos. Existen varios factores microbiológicos que afectan la granulación de la biomasa, como la hidrofobicidad superficial de las células (Grotenhuis et al. 1992; Daffonchio et al. 1995; Schmidt and Ahring 1996), la presencia de polímeros extracelulares (Schmidt and Ahring 1996; Veiga et al. 1997) y la composición de la comunidad microbiana (Wiegant and de Man 1986; Pol et al. 1988; Schmidt and Ahring 1996). Por ejemplo, los microorganismos del género *Methanosaeta* (Archaea) son filamentosos y se consideran importantes en la generación de los núcleos de los gránulos mediante la construcción de estructuras en forma de red (Wiegant and de Man 1986). La etapa de granulación ocurre al inicio de la etapa de arranque de los reactores y puede alcanzar una duración de entre 2 y 8 meses (Adhikari and Lohani 2019). Las partículas más pequeñas y dispersas se eliminarán del sistema saliendo junto con el agua residual ya tratada, mientras que los gránulos más compactos y de mayor densidad quedarán retenidos en el reactor (Hulshoff Pol et al. 2004a). Normalmente, los gránulos tienen un diámetro que varía de 0,14 mm a 5 mm. La distribución espacial de Archaeas y bacterias en los gránulos ha sido estudiada mediante Hibridación *in situ* Fluorescente (FISH) con sondas de oligonucleótidos específicas para ambos grupos (Sekiguchi et al. 1999; Yamada et al. 2005b; Satoh et al. 2007). Estos estudios han determinado que los gránulos tienen estructuras multicapa de diferentes grupos filogenéticos, donde las bacterias hidrolíticas se encuentran en

la parte más externa del gránulo, seguidas por bacterias acetogénicas, y en la parte más interna se encuentran las Archaeas metanogénicas (Sekiguchi et al. 1999).

1.2.3 Sistemas de remoción de Nitrógeno

Sistemas de remoción de nitrógeno convencionales

La remoción biológica de nitrógeno convencional implica los procesos de nitrificación y de desnitrificación, las cuales se llevan a cabo en dos reactores separados. El proceso de nitrificación, implica la oxidación de amonio a nitrito o nitrato, disminuyendo la demanda de oxígeno del efluente, pero sin eliminar el nitrógeno del sistema (Wei et al. 2014). Esta etapa del proceso está caracterizada por ocurrir en condiciones aerobias estrictas, llevándose a cabo en dos etapas de oxidación. Inicialmente el amonio (NH_4^+) es oxidado a nitrito (NO_2^-), proceso conocido como nitritación, y posteriormente se da la oxidación de nitrito a nitrato (NO_3^-), proceso denominado nitratación (Figura 1.6).

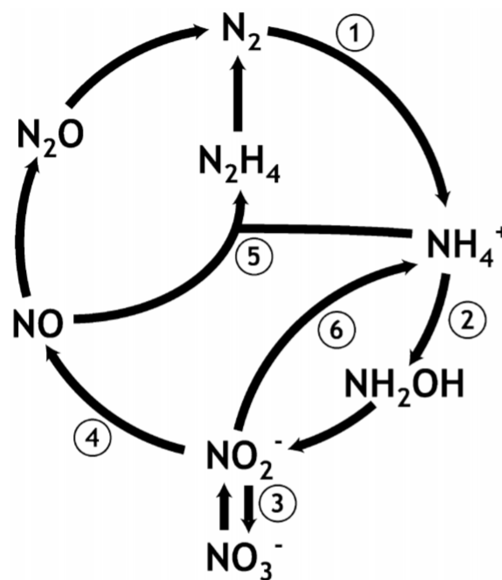

Figura 1.6. Ciclo del nitrógeno microbiano: (1) Fijación de nitrógeno; (2) oxidación aerobia de amonio; (3) oxidación aerobia de nitrito; (4) desnitrificación; (5) oxidación anaerobia de amonio; y (6) reducción disimilatoria de nitrato y nitrito a amonio (Jetten, 2008).

Ambos procesos son realizados por bacterias nitrificantes autótrofas las cuales requieren de la presencia de oxígeno, por lo tanto, si predominan condiciones reductoras se dificulta la formación del nitrato (Locey and Lennon, 2016). Cada etapa de oxidación es llevada a cabo por diferentes géneros de bacterias, como las bacterias oxidantes de amonio y las bacterias

oxidantes de nitrito, que usan amonio o nitrito como fuente de energía respectivamente (Ahn 2006). El género más común de bacterias autótrofas que lleva a cabo la oxidación del amonio es *Nitrosomonas*, en el caso de la oxidación de nitrito es *Nitrobacter* (Kim et al. 2008). *Nitrosococcus*, *Nitrosospira*, *Nitrosovibrio* y *Nitrosolobus* también son capaces de oxidar amonio a nitrito (Yu et al. 2010a).

En un segundo reactor, se lleva a cabo la desnitrificación, en donde el nitrato es convertido en óxido nitroso (N_2O, desnitrificación parcial) o nitrógeno gaseoso (N_2, desnitrificación completa), los cuales finalmente son eliminados a la atmósfera (Figura 1.6) (Zhang et al. 2012). El proceso de desnitrificación se lleva a cabo por enzimas específicas que catalizan el proceso de reducción en etapas: nitrato reductasa (Nar), nitrito reductasa (Nir), óxido nítrico reductasa (Nor) y óxido nitroso reductasa (Nos) catalizan el proceso de reducción en etapas (Zumft 1997; Ji et al. 2015). La desnitrificación ocurregeneralmente en condiciones anóxicas por bacterias heterótrofas facultativas que reducen el nitrato hasta nitrógeno molecular. Esta reacción ocurre en etapas consecutivas apareciendo como productos intermedios el nitrito (NO_2^-), óxido nítrico (NO) y óxido nitroso (N_2O); dado que estos compuestos nitrogenados son gases insolubles, no pueden ser incorporados al material celular, sino que escapan a la atmósfera (Gerardi 2002). Las bacterias desnitrificantes utilizan el carbono orgánico para la síntesis celular y fuente de energía y el nitrato como aceptor final de electrones constituyendo un mecanismo respiratorio alterno.

1.2.3.1 Reactores anammox

La oxidación anaerobia de amonio (anammox) convierte directamente el amonio (NH_4^+) en gas nitrógeno (N_2) utilizando el nitrito (NO_2^-) como aceptor de electrones en condiciones anóxicas (Strous et al. 1998). Desde su descubrimiento a mediados de la década de 1990, el proceso anammox es considerado un método eficiente y económico en la eliminación de altas concentraciones de amonio en aguas residuales con baja cantidad de materia orgánica, en comparación con los sistemas convencionales (Tang et al. 2011; Ma et al. 2013). Este tipo de sistema se utiliza generalmente como tratamiento biológico terciario de aguas residuales ricas en nitrógeno.

La principal ventaja que tienen los sistemas anammox sobre los sistemas convencionales de nitrificación-desnitrificación es el menor costo de operación, ya que no requiere aireación y el proceso ocurre en un solo reactor, además de tener menor emisión de gases de efecto invernadero (ya que el óxido nitroso N_2O no es un producto intermedio) (Ali et al. 2016;

Wang et al. 2016). En los reactores anammox, una de las estrategias desarrolladas para enriquecer en bacterias anammox los reactores y lograr un rendimiento estable es la granulación de la biomasa (Tang et al. 2017) (Figura 1.7).

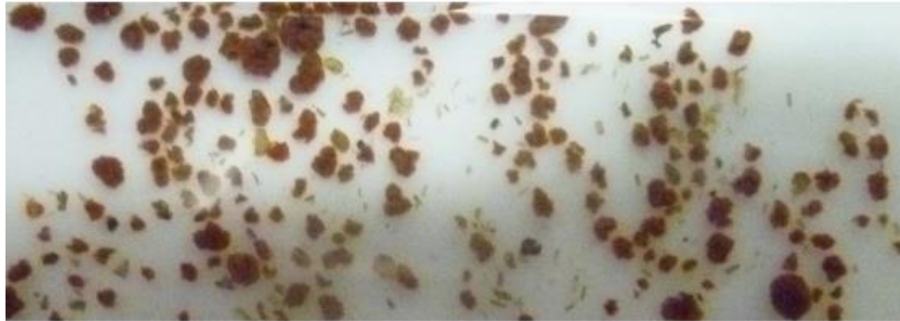

Figura 1.7. Gránulos pertenecientes a un reactor anammox (Speth et al. 2016).

El filo Chloroflexi ha sido detectado con alta abundancia en los sistemas anammox, coexistiendo junto a bacterias anammox. Inclusive, en reactores anammox alimentados con aguas residuales sintéticas que contenían NH_4^+ como único donante de electrones y en ausencia de compuestos orgánicos (Strous et al. 2006; Cho et al. 2011; Kindaichi et al. 2012). Una de las hipótesis propuestas, es que los organismos del filo Chloroflexi podrían estar metabolizando materia orgánica derivada de las células de bacterias anammox, productos solubles microbianos o exopolisacáridos, provenientes de microorganismos autótrofos (Kindaichi et al. 2012; Chu et al. 2015). Un trabajo reciente, sugiere que los microorganismos del filo Chloroflexi podrían facilitar el loop de nitrito con bacterias anammox o colaborar en la desnitrificación completa debido a la expresión de genes de la reducción disimilatoria de nitrato a amonio, óxido nítrico reductasa (norZ) y nitrito reductasas (nirK, nirS), mejorando así el rendimiento en la eliminación nitrógeno en el reactor (Lawson et al. 2017; Zhao et al. 2018).

1.3 El filo Chloroflexi

Se estima que un 99% de las especies microbianas no tienen representantes en cultivo puro, por lo tanto, su filogenia y función en diversos ecosistemas continúan siendo desconocidas (Ward et al. 1992; Torsvik et al. 1996; McCaig et al. 2001). Por lo tanto, estos microorganismos reciben la denominación de "materia microbiana oscura" (del inglés “microbial dark matter”). Sin embargo, se ha logrado demostrar mediante técnicas independientes del cultivo que dichos microorganismos están activos en sus respectivos hábitats (Hawley et al. 2017; Castelle and Banfield 2018; Vavourakis et al. 2018). El filo

Chloroflexi está representado por un grupo diverso fisiológica y filogenéticamente, donde la gran mayoría de sus organismos no tienen representantes cultivados, formando parte de dicha "materia microbiana oscura". Hasta 2001, este filo consistía en una única clase denominada Chloroflexia que contenía dos órdenes Chloroflexales y Herpetosiphonales, cada uno con una sola familia (es decir, Chloroflexaceae y Herpetosiphonaceae) y en conjunto compuesto por cinco géneros (Garrity et al. 2001; Gupta et al. 2013). Sin embargo, en los últimos 20 años, este filo bacteriano ha experimentado una enorme expansión, inicialmente, por la fusión de especies que anteriormente formaban parte de otros filos bacterianos (por ejemplo, Thermomicrobia y Actinobacteria) (Garrity et al. 2001; Hugenholtz and Stackebrandt 2004), y posteriormente por la inclusión de algunas especies aisladas o genomas obtenidos a través del ensamblado de metagenomas.

Actualmente, el filo Chloroflexi comprende 8 clases: Dehalococcoidia, Anaerolineae, Caldilineae, Ktedonobacteria, Thermomicrobia, Ardenticatenia, Thermoflexia y Chloroflexia (Figura 1.8) (Kawaichi et al. 2013; Dodsworth et al. 2014; Hanada 2014).

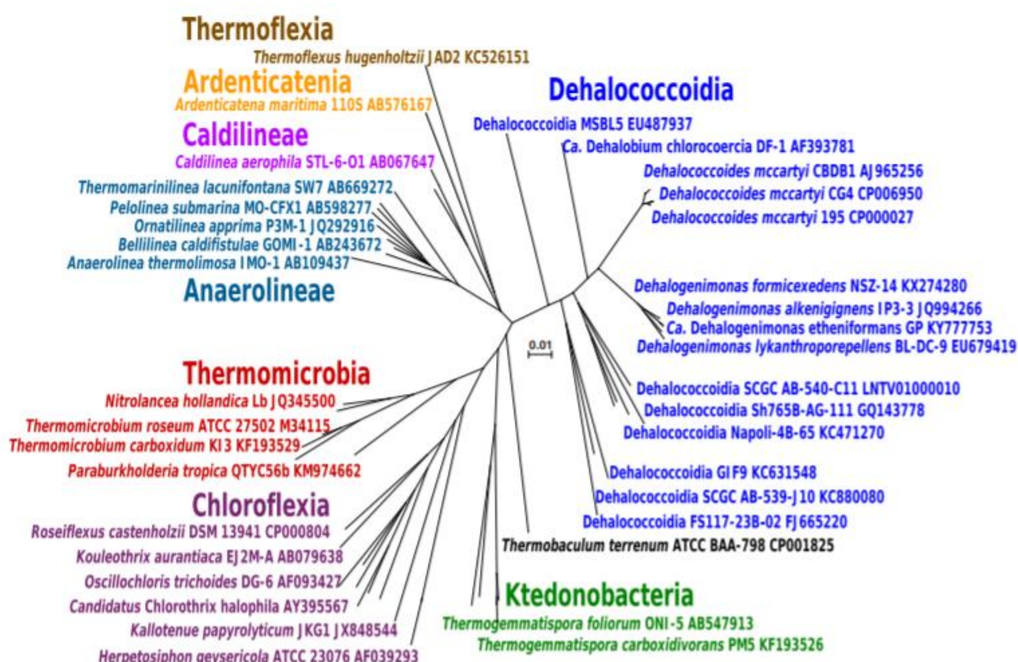

Figura 1.8. Árbol filogenético del filo Chloroflexi construido a partir del gen de ARN ribosomal de 16S de representantes cultivados y organismos *Candidatus* (Yang et al. 2020).

La gran diversidad del gen de ARNr 16S que presentan dichas clases también es evidente en los análisis filogenéticos y comparativos de genomas. Se ha sugerido que el filo Chloroflexi *sensu stricto* solo debería consistir en las clases Chloroflexia y Thermomicrobia, mientras que otros miembros deberían reclasificarse (Gupta et al. 2013). En un esfuerzo por

estandarizar la taxonomía del dominio Bacteria basada en la filogenia de genomas, (Parks et al. 2018) propuso una reclasificación del filo Chloroflexi ("Chloroflexota") en las siguientes 8 clases: Anaerolineae, Chloroflexia, Dehalococcoidia, Ktedonobacteria, UBA2235, UBA4733, UBA5177 y UBA6077. En la clasificación taxonómica utilizada por la base de datos de SILVA (versión 138), el filo Chloroflexi está dividido en las clases Anaerolineae, Chloroflexia, Dehalococcoidia, Ktedonobacteria y otras clases que aún no tienen representantes cultivados. Según esta clasificación, la clase Anaerolineae comprende los órdenes Anaerolineales, Ardenticatenales, Caldilineales, Thermoflexales y otros órdenes de los cuales aún no hay representantes cultivados.

1.3.1 Relevancia de la clase Anaerolineae en sistemas de tratamiento de aguas residuales

Dentro del filo Chloroflexi, generalmente la clase Anaerolineae conforma el grupo central de microorganismos (core) presente en reactores anaerobios (McIlroy et al. 2016; Sun et al. 2016; Xia et al. 2016; Petriglieri et al. 2018), en reactores Anammox (Kindaichi et al. 2012; Cao et al. 2016; Wang et al. 2016) y en reactores de lodos activados (Björnsson et al. 2002; Kragelund et al. 2007; Speirs et al. 2009; Yamada and Sekiguchi 2009; Speirs et al. 2011; Nierychlo et al. 2018). Desde el 2005, solo han sido aisladas 15 especies de la clase Anaerolineae (Tabla 3), de las cuales 6 fueron recuperadas de sistemas de tratamiento anaerobio (5 especies aisladas de reactores de aguas residuales y 1 especie aislada de un reactor de residuos sólidos).

En los últimos años, los esfuerzos por ensamblar genomas a partir de metagenomas en sistemas de aguas residuales han aumentado debido a la dificultad de obtener cepas de Chloroflexi en cultivo puro. Recientemente, se ensamblaron genomas de la clase Anaerolineae a partir de metagenomas en sistemas de tratamiento aerobios y anaerobios correspondientes a los órdenes Anaerolineales (*Candidatus* Brevefilum fermentans), Ardenticatenales (*Candidatus* Promineofilum) y Caldilineales (*Candidatus* Amarolinea aalborgensis) (McIlroy et al. 2016; McIlroy et al. 2017; Andersen et al. 2018) (Tabla 1.3). A partir del último genoma, los autores proponen la creación de una nueva familia, la cual denominaron Amarolineaceae. Además del ensamblado de genomas, se han logrado la caracterización fisiológica *in situ* mediante la técnica de FISH-MAR de algunos organismos de la clase Anaerolineae como *Ca.* Villigracilis y *Ca.* Sarcinithrix (Nierychlo et al. 2019a).

Tabla 1.3. Resumen de cepas aisladas y genomas ensamblados de la clase Anaerolineae.

Family	*Species*	Isolation source	Temperature optimum	Physiology	Electron acceptor conditio	Carbon sources/electron	References
Ardenticatenia	*Ardenicaena mariima*	Coastal hydrothermal field	Thermophile	Chemoheterotroph	Facultative anaerobe: O_2 NO_2^- (to NH_3 or N_2*); Fe(III)	Carbohydrates; proteins; fatty acids	(Kawaichi et al. 2013)
	Ca. Promineofilum breve	Activated sludge	Mesophile	Chemoheterotroph; Autotroph?; Fermenter	Facultative anaerobe: O_2 NO_2^- (to NH_3*); N_2O*	Carbohydrates	(McIlroy et al. 2016)
SJA-15	*Ca. Amarolinea aalborgensis*	Activated sludge	Mesophile	Chemoheterotroph; Fermenter	Facultative anaerobe: O_2 NO_2^- (to NH_3*)	Carbohydrates	(Andersen et al. 2018)
Anaerolineae	*Ca. Brevefilum fermenans*	Anaerobic digester	Mesophile	Chemoheterotroph; Fermenter	Strict anaerobe	Carbohydrates; proteins**	(McIlroy et al. 2017)
	Bellilinea caldifisulae	Thermophilic anaerobic digester	Thermophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins^w	(Yamada et al. 2007a)
	Flexiliniea Flocculi						(Sun et al. 2016)
	Levilinea saccharolyica	Mesophilic anaerobic digester	Mesophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Yamada et al. 2006)
	Longilinea arvoryzae	Rice paddy soil	Mesophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Yamada et al. 2007a)
	Lepolinea ardivialis	Mesophilic anaerobic digester	Mesophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Yamada et al. 2007a)
	Pelolinea submarina	Marine sediment	Mesophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Imachi et al. 2014)
	Ornailinea apprima	Hot water bath microbial mat	Mesophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Podosokorskaya et al. 2016)
	Anaerolinea hermophila	Thermophilic anaerobic digester	Thermophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins^w	(Sekiguchi et al. 2001a)
	Anaerolinea hermolimosa	Thermophilic anaerobic digester	Thermophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates; proteins	(Yamada et al. 2006)
	hermomarinilineae lacunofonalis	Hydrothermal vent	Thermophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Proteins	(Nunoura et al. 2013)
	Aggregailinea lena	Marine subsurface sediment	Mesophile	Chemoheterotroph; Fermenter	Facultative anaerobe	Fumarato, piruvato	(Nakahara et al. 2019)
	hermanaerohrix daxensis	Deep hot aquifier	Thermophile	Chemoheterotroph; Fermenter	Obligate anaerobe	Carbohydrates	(Grégoire et al. 2011b)
Caldilineae	*Caldilinea aerophila*	Hot spring biomat	Thermophile	Chemoheterotroph; Fermenter	Facultative anaerobe: O_2	Carbohydrates; proteins; fatty acids	(Sekiguchi et al. 2003)
	Caldilinea arbellica	Deep hot aquifier	Thermophile	Chemoheterotroph; Fermenter	Anaerobe (tolerates <1.5% O_2)	Carbohydrates; proteins	(Grégoire et al. 2011a)
	Liorilinea aerophila	Intertidal hot spring	Thermophile	Chemoheterotroph	Aerobe: O_2	Carbohydrates; proteins	(Kale et al. 2013)

Tanto las especies aisladas de la clase Anaerolineae en cultivo puro, como los genomas ensamblados y las especies caracterizadas *in situ*, comparten la mayor parte de las características fenotípicas y fisiológicas. Los microorganismos de las especies aisladas son heterótrofos, presentan morfología filamentosa, pueden ser anaerobios estrictos o facultativos, mesófilos a moderadamente termófilos, y utilizan carbohidratos y aminoácidos (o péptidos) como fuente de carbono (Tabla 1.3).

A pesar de haber avanzado gracias a las técnicas de biología molecular, el rol ecológico de la clase Anaerolineae sigue siendo desconocido en sistemas de tratamiento debido a la escasez de aislamientos. Todas las especies aisladas de Anaerolineae, excepto *Ornatilinea apprima*, son microorganismos de crecimiento lento con tiempos de duplicación de 10 a 100 horas en condiciones óptimas (Yamada and Sekiguchi 2009; Grégoire et al. 2011; Podosokorskaya et al. 2016). El crecimiento relativamente lento de este grupo y la coexistencia de diversas bacterias heterótrofas de rápido crecimiento en diversos ambientes podrían ser la causa de la dificultad de enriquecer y aislar especies de Anaerolineae (Nunoura et al. 2013). Debido a la dificultad existente para obtener representantes cultivados, y la ausencia de la posibilidad de clasificar los organismos no cultivados detectados mediante distintas técnicas de secuenciación en sistemas de lodos activados, recientemente se desarrolló una base de datos específica para sistemas de tratamiento de aguas residuales denominada MiDAS (Microbial Diversisty Activated Sludge) (Nierychlo et al. 2020). Gracias a un gran esfuerzo colaborativo

entre varios países de todo el mundo (en el cual Uruguay fue partícipe contribuyendo en la recolección de muestras de varias plantas de tratamiento de lodos activados), se logró reclutar muestras de sistemas de tratamiento, a partir de las cuales se secuenció el gen de ARNr 16S completo. MiDAS incluye además de estas secuencias específicas de comunidades microbianas en sistemas de tratamiento, la base de datos de referencia del Silva, por lo tanto se basa en la misma clasificación taxonómica. MiDAS facilita la comprensión de la diversidad y función de la comunidad en sistemas de tratamiento de aguas residuales. Una de las ventajas que presenta esta base de datos, es que permite la comparación de microorganismos no cultivados entre distintos trabajos, ya que a cada secuencia de referencia a la cual no es posible clasificar taxonómicamente, se le asigna un código alfanumérico, por lo tanto será identificada, y podrá ser comparada entre distintas investigaciones.

1.3.2 Rol de organismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales

Debido a la dificultad de aislar estos microorganismos en cultivo puro y a la poca información existente sobre su metabolismo, aún no se conoce el rol de estos microorganismos en los diferentes sistemas de tratamiento de aguas residuales. Existen varias hipótesis generadas a partir de los pocos aislamientos o las caracterizaciones *in situ* sobre la posible predominancia de la clase Anaerolineae en sistemas de tratamiento. Estas hipótesis tienen que ver con su metabolismo hidrolítico y fermentador, la posible relación sintrófica con Archaeas metanogénicas en sistemas anaerobios, su morfología filamentosa y propiedades de adherencia y el potencial rol en la formación de gránulos y flóculos. Las cepas aisladas y caracterizadas en cultivo puro, así como también, las especies caracterizadas *in situ*, tienen la capacidad de fermentar diversos compuestos orgánicos como azúcares, proteínas y aminoácidos. Esto favorecería su crecimiento en estos sistemas, metabolizando la materia orgánica proveniente de residuos sólidos o líquidos, así como también degradando restos celulares. Otra hipótesis planteada es que debido a que todos los microorganismos descriptos tienen morfología filamentosa, podrían estar implicados en la formación del esqueleto de gránulos y flóculos, jugando un papel importante en la formación, compactación y buena sedimentación los mismos (Björnsson et al. 2002; Yamada et al. 2005b). Sumado a esto, se ha postulado una posible asociación sintrófica entre algunas de las cepas aisladas de la clase Anaerolineae con Archaeas metanogénicas. Por ejemplo, algunas de las especies de Anaerolineae aisladas de reactores anaerobios, requieren co-cultivo con metanógenos para un crecimiento eficiente (Sekiguchi et al. 2001a; Yamada et al. 2005a; Yamada et al. 2006).

Además de los trabajos donde se ensamblaron 3 genomas de la clase Anaerolineae a partir de los cuales se determinaron sus potenciales funciones metabólicas, recientemente un estudio realizado en un reactor termófilo, determinó la existencia de genes relacionados a la adhesividad celular en los genomas recuperados (Xia et al. 2016). Esta característica podría conferir a algunos organismos de la clase Anaerolineae cierta ventaja para su permanencia en los reactores, así como también un rol importante en la formación de gránulos y flóculos.

1.3.3 Problemas de bulking ocasionados por sobrecrecimiento de bacterias filamentosas del filo Chloroflexi

El crecimiento excesivo de bacterias filamentosas es uno de los obstáculos más desafiantes para la estabilidad estructural de gránulos y flóculos, ya que generalmente causa el deterioro de dichas estructuras, disminuyendo su sedimentabilidad e incluso comprometiendo la operación general de los sistemas de tratamiento (Moura et al. 2018). Durante un episodio de bulking, se deben tomar medidas rápidas para controlarlo (Liu et al. 2019). Actualmente existen métodos específicos y no específicos para el control del bulking (Martins et al. 2004; Liu and Liu 2006). Sin embargo, las dificultades para conocer los factores que desencadenan estos eventos limitan la aplicación de los métodos específicos, así como también las modificaciones del reactor, la capacidad de tratamiento y las características del efluente a tratar (Guo et al. 2012a; Fan et al. 2018). Por otro lado, los métodos no específicos pueden utilizarse antes de que se identifiquen las causas que conducen a la acumulación de lodos. Los biocidas son uno de los sustratos más eficientes y más utilizados en el control de bacterias filamentosas en condiciones de bulking grave (Guo et al. 2012b). Dado que los filamentos sobresalen de los gránulos o flóculos, es posible atacarlos en su mayoría sin incidir tanto sobre las bacterias formadoras de flóculos que se encuentran más internamente en estas estructuras (Liu and Liu 2006; Guo et al. 2012b). Si bien parecería que los organismos del filo Chloroflexi podrían ser ventajosos para los sistemas de tratamiento de aguas residuales debido a su capacidad para degradar materia orgánica y posiblemente favorecer en la estructura de gránulos y flóculos, su sobrecrecimiento ha sido identificado como el agente causante de episodios de bulking, generando pérdida de biomasa del reactor y el lavado del mismo (Sekiguchi et al. 2001a; Yamada et al. 2007; Borzacconi et al. 2008; Li et al. 2008; Sekiguchi et al. 2015). Como se mencionó anteriormente, la dificultad de aislarlos en cultivo puro limita el entendimiento de los factores que desencadenan su sobrecrecimiento. Profundizar en la filogenia y metabolismo de este grupo a partir del

aislamiento y del ensamblado de genomas a partir de metagenomas, servirá como herramienta para entender su rol estructural y funcional en estos ecosistemas.

1.3.3.1 Bulking en reactores de lodos activados

El bulking filamentoso es un problema frecuente en sistemas de lodos activados. El sobrecrecimiento de bacterias filamentosas en estos sistemas genera el esponjamiento y la interconexión de los flóculos, disminuyendo su rigidez y sedimentabilidad, teniendo como resultado un clarificado deficiente (Figura 1.9).

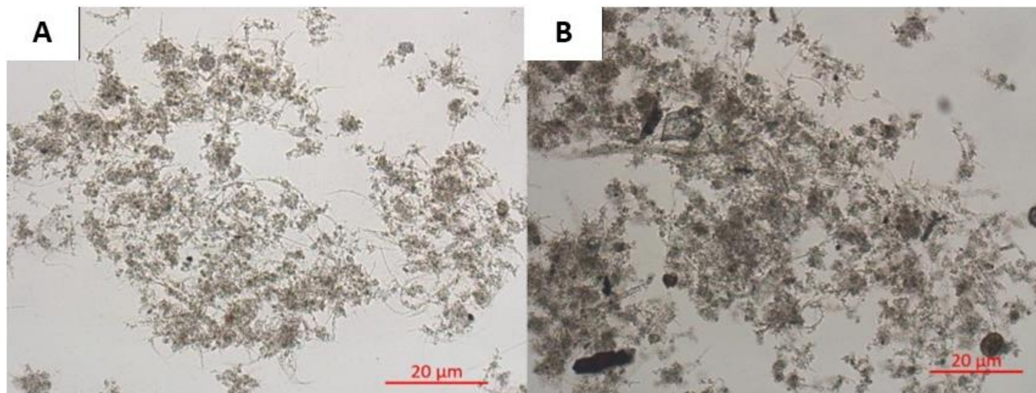

Figura 1.9. Micrografías de contraste de fases de la estructura de los flóculos: a) Presencia de bacterias filamentosas excesiva, b) Bacterias filamentosas dominantes (Deepnarain et al. 2020).

A pesar de la gran cantidad de trabajos científicos enfocados en el estudio de bacterias filamentosas y los posibles desencadenantes de su sobrecrecimiento, el bulking parece ser un problema continuo en la operación de sistemas de lodos activados. Comparado con reactores anaerobios, en sistemas de lodos activados se ha identificado una mayor diversidad de bacterias filamentosas (pertenecientes a distintos filos) que están frecuentemente relacionados con episodios de bulking, tales como *Candidatus* Microthrix (Rossetti et al. 2005), *Gordonia*, y *Skermania* (Carr et al. 2006; Kragelund et al. 2006) y *Candidatus* Nostocoida (Schade et al. 2002), así como también especies del filo Chloroflexi (Björnsson et al. 2002; Kragelund et al. 2007; Speirs et al. 2011; McIlroy et al. 2016; Nittami et al. 2017).

El sistema de identificación convencional para los microorganismos filamentosos más frecuentes en lodos activos fue publicado por Eikelboom en 1975 (Eikelboom 1975). Posteriormente, Eikelboom y Jenkins desarrollaron claves de identificación dicotómicas a partir de una serie de características morfológicas y reactivas a las tinciones diferenciales

clásicas, entre ellas Gram y Neisser (Eikelboom and van Buijsen 1983; Jenkins et al. 2003b). De esta forma, las bacterias filamentosas fueron agrupadas bajo características comunes, estableciéndose su denominación con un código de cuatro cifras precedido por la denominación 'tipo o morfotipo' y en algunos casos a nivel de género. Inicialmente, muchas de las bacterias filamentosas pertenecientes a Chloroflexi, habían sido identificadas en reactores aerobios a partir de técnicas de microscopía y definidos como morfotipos de Eikelboom (Jenkins et al. 2003b). La gran desventaja de esta identificación, es que dos especies distintas podían tener el mismo morfotipo. Pero en los últimos años, con los avances en técnicas de biología molecular su identificación ha sido facilitada. Aun así, estos microorganismos identificados mediante secuenciación de su ADN, permanecen sin un representante cultivado. Es por esto que en 1995, el Comité Internacional de Sistemática en Bacteriología recomendó que se implementara la categoría *Candidatus*, propuesta por Murray y Stackebrandt (Murray and Stackebrandt 1995). De esta forma, es posible denominar a supuestos taxones de procariotas que no han podido ser aislados, de los que se cuenta con cierta información además de su secuencia de ADN, pero para las que faltan las características requeridas para la descripción de acuerdo con el Código Internacional de Nomenclatura de Bacterias. Además de la información genómica para determinar la posición filogenética en la descripción de un taxón provisional (*Candidatus*), es necesario describir sus características estructurales, metabólicas y reproductivas, junto con el entorno natural en el que fue detectado. Algunos ejemplos de morfotipos de Eikelboom del filo Chloroflexi provenientes de reactores aerobios que fueron posteriormente nombrados como taxones provisionales son *Ca.* Amarolinea y *Ca.* Promineofilum (Eikelboom type 0092), Kouleothrix (Eikelboom type 1851), *Ca.* Sarcinithrix (Eikelboom type 0914), *Ca.* Amarithrix (type 0675) y *Ca.* Catenibacter (type 0041). Recientemente, especies pertenecientes al género *Ca.* Amarolinea, junto con especies del género *Ca.* Microthrix, fueron reportados como agentes causantes de episodios de bulking en plantas de tratamiento aerobias en Dinamarca (Nierychlo et al. 2020b). Uno de los mayores problemas en determinar qué factores desencadenan los episodios de bulking debido al sobrecrecimiento de bacterias del filo Chloroflexi es que son difíciles de aislar en cultivo puro, lo que dificulta el estudio detallado del metabolismo de estos organismos. Las condiciones operacionales de sistemas escala real en las cuales se da el episodio de bulking generalmente son escasamente documentadas, por lo tanto, se dificulta la posibilidad de correlacionar estos eventos con las condiciones operacionales de las plantas de tratamiento. Estudiando la ecofisiología de bacterias filamentosas (ya sea en cultivo puro o mediante la aplicación de técnicas *in situ*,

metagenómica y metatranscriptómica) se busca conocer el rol metabólico en estos sistemas para poder determinar los factores desencadenantes de su sobrecrecimiento.

1.3.3.2 Bulking en reactores anaerobios tipo UASB

Los sistemas anaerobios generalmente presentan menos problemas de bulking que los sistemas de lodos activados, y también presentan menor diversidad de organismos vinculados al bulking. Dentro de los grupos bacterianos más reportados como agente causante de bulking se encuentra el filo Chloroflexi (Figura 1.10 (A, B, C)) (Sekiguchi et al. 2001a; Yamada et al. 2007; Borzacconi et al. 2008; Li et al. 2008; Sekiguchi et al. 2015).

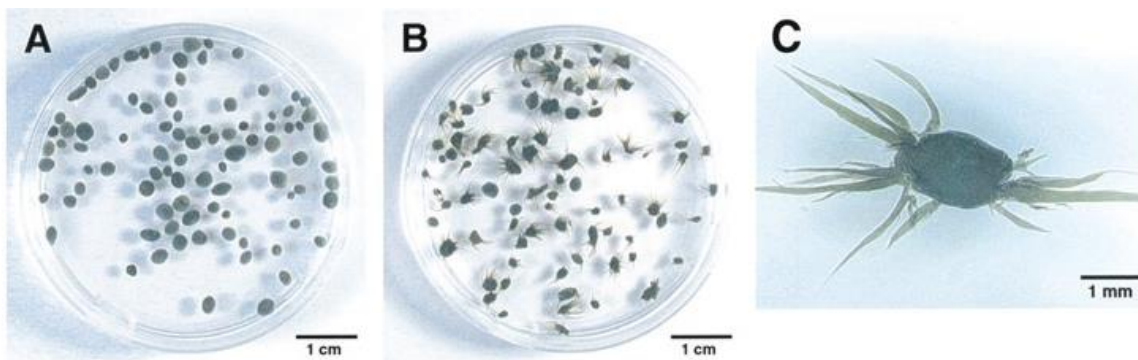

Figura 1.10. Gránulos de un reactor UASB: A) Con estructura compacta, B) Con sobrecrecimiento de bacterias filamentosas del filo Chloroflexi, C) Gránulo con sobrecrecimiento de bacterias filamentosas (Sekiguchi et al. 2001a).

El bulking filamentoso genera un aumento del volumen de los gránulos, disminuyendo su densidad y sedimentabilidad, generando la pérdida de biomasa del reactor por flotación de dichos gránulos (Appels et al. 2011; Sekiguchi et al. 2015). En Uruguay en el año 2008, se detectó una pérdida de eficiencia en un reactor metanogénico (UASB) de una industria procesadora de malta debido al sobrecrecimiento de bacterias filamentosas. Mediante la técnica de FISH con sondas específicas para el filo Chloroflexi, se determinó que este grupo era el responsable de la pérdida de densidad de los gránulos y su flotación. Se hipotetizó que su sobrecrecimiento podría deberse al alto contenido de células en el reactor, y también a la lisis celular debida a una bajada abrupta del pH (Borzacconi et al. 2008). Esta lisis celular podría haber causado la disponibilidad de materia orgánica para ser consumida por bacterias del filo Chloroflexi. La degradación de restos celulares ha sido demostrada en algunos trabajos, donde organismos del filo Chloroflexi fueron capaces de metabolizar N-

acetilglucosamina que es un componente de la pared celular (Kindaichi et al. 2004; Miura et al. 2007; Campbell et al. 2014).

1.3.3.3 Bulking en reactores anammox

Como se ha mencionado anteriormente, la buena sedimentabilidad de los gránulos depende en gran medida de su compactación y densidad. Los gránulos también ofrecen gran cantidad de polímeros extracelulares (EPS) que aumentaran la resistencia de bacterias anammox a factores adversos (Batchelor et al. 1997). Algunos trabajos han encontrado que el filo Chloroflexi presenta menor abundancia relativa en gránulos compactos (12%) y mayor en gránulos de baja sedimentación (20%) (Song et al. 2017) (Figure 1.11). Sin embargo, otros encontraron una alta abundancia relativa de Chloroflexi (28%) en gránulos que tenían una estructura compacta (Chu et al. 2015). Probablemente las poblaciones de Chloroflexi que integraban cada uno de estos reactores presentaban características distintas, por ejemplo de adhesividad, de formación de exopolisacáridos, o de posición espacial en los gránulos. Estas características podrían ser claves en el desarrollo de estructuras de gránulos no beneficiosas.

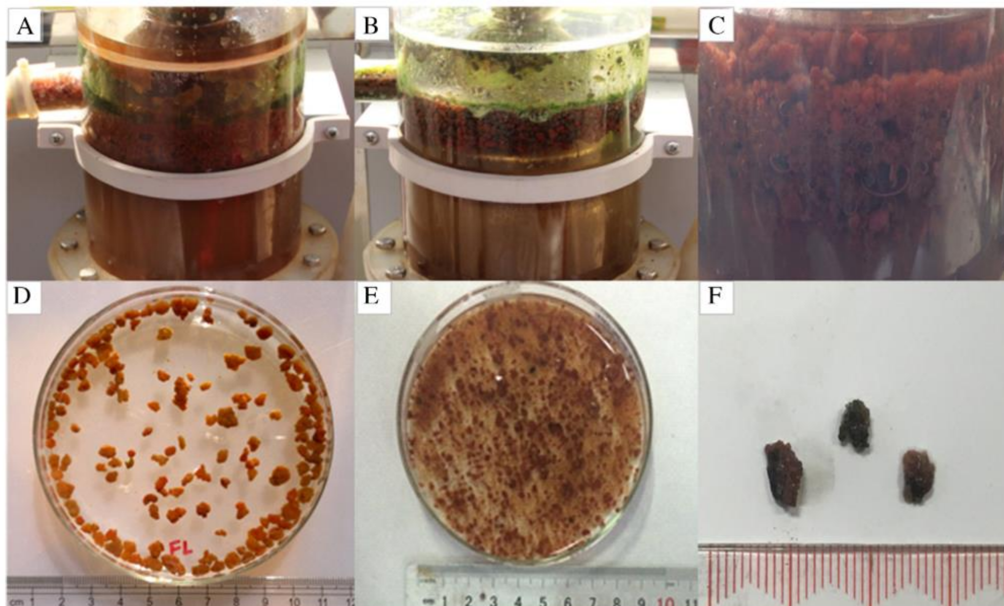

Figure 1.11. Fotografías de los gránulos anammox: A) y B) flotación de gránulos, D) gránulos con aumento de volumen (esponjamiento), E) y F) gránulos compactos (Song et al. 2017).

Por otro lado, los eventos de bulking filamentoso generado por el filo Chloroflexi no parecerían ser muy comunes en los reactores anammox, ya que en literatura solo aparecen las citas antes mencionadas.

1.4 Métodos moleculares aplicados al estudio de comunidades microbianas en sistemas de tratamiento

En los dominios de Bacteria y Archaea, el marcador evolutivo por excelencia es el gen de ARNr 16S, debido a que es ubicuo en procariotas, evoluciona a un ritmo constante (reloj filogenético), y consta de regiones altamente conservadas y regiones hipervariables (V1-V9) (Janda and Abbott 2007). Por lo tanto, a partir de la determinación de su secuencia mediante métodos de secuenciación de ADN es posible saber la relación filogenética y taxonómica de los microorganismos en estudio. Las distintas regiones variables y conservadas permiten el uso de primers específicos y primers universales, respectivamente. En las últimas décadas se han desarrollado distintas técnicas de biología molecular basadas en el estudio del gen del ARNr 16S, como por ejemplo, electroforesis en gel de gradiente desnaturalizante (DGGE), polimorfismos de longitud de fragmentos de restricción terminales (T-RFLP), hibridación *in situ* fluorescente (FISH) y varios tipos de secuenciación de ADN para determinar la composición taxonómica y funcional de las comunidades microbianas. La secuenciación del gen de ARNr 16S es el método más usado para estudiar la abundancia, taxonomía y filogenia de las comunidades microbianas complejas presentes en sistemas de tratamiento (Wojnowska-Baryła et al. 2010; Cabezas et al. 2015b). En los últimos años, con los avances tecnológicos en secuenciación de genomas, metagenomas y metatranscriptomas se ha logrado profundizar en el estudio de la diversidad, función y evolución, desde candidatos a nuevos filos hasta la descripción de nuevas especies (Albertsen et al. 2013; Matsuura et al. 2015; McIlroy et al. 2016; Andersen et al. 2018). Cada técnica apunta a contestar distintas preguntas y es recomendable utilizar varias técnicas en un mismo estudio de forma complementaria (Figura 1.12).

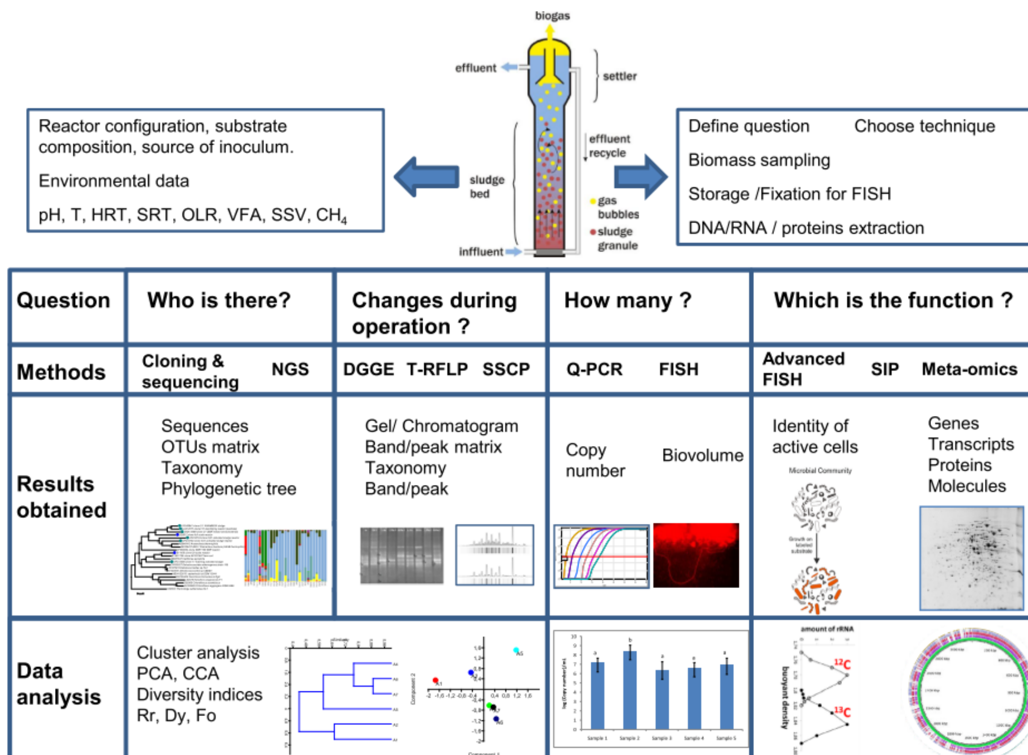

Figura 1.12. Esquema de distintas preguntas que pueden ser formuladas en un trabajo de investigación y las técnicas moleculares que pueden ser utilizadas para su abordaje (Cabezas et al. 2015b).

El ensamblado de genomas a partir de metagenomas y la secuenciación de genomas por *Single Cell Genomics* de organismos no cultivados son aproximaciones que están en constante aumento, debido a la necesidad de la reconstrucción de modelos metabólicos que intenten explicar la fisiología de los microorganismos más en detalle y su posible rol en cada ecosistema. A su vez, estos modelos metabólicos forman la base de otras técnicas moleculares como la caracterización *in situ* de células individuales, metatranscriptómica, proteómica, así como también en el diseño de medios de cultivo (Koch et al. 2014). Para comprender la ecología de los sistemas de tratamiento de aguas residuales se requiere una mejor identificación de los microorganismos que son difíciles de cultivar y la determinación de sus roles funcionales (Nielsen et al. 2010). En el caso de sistemas de tratamiento de aguas residuales, varios estudios basados en secuenciación del gen de ARNr 16S han revelado que existe un core de microorganismos para cada tipo de sistema, que son muy abundantes y que están presentes de manera estable a lo largo del tiempo (Rivière et al. 2009; Werner et al. 2011; Sundberg et al. 2013; De Vrieze et al. 2015; Saunders et al. 2016; Izadi et al. 2020). Sin embargo, queda mucho por comprender sobre la diversidad y función de las comunidades

microbianas en estos sistemas, y dada las interacciones sinérgicas de algunos microorganismos, no es posible un enfoque reduccionista para comprender la ecología del sistema (Kaeberlein et al. 2002; Fuhrman et al. 2015). A continuación se describen las técnicas de biología molecular utilizadas en esta tesis.

1.4.1 Hibridación *in situ* Fluorescente (*Fluorescent in situ hybridization* - FISH)

La hibridación *in situ* fluorescente es una técnica de hibridación específica, donde se diseñan sondas de ADN marcadas con un fluorocromo en el extremo 5', que hibridan específicamente con una secuencia de ARN blanco complementaria en células intactas (Moter and Gobel 2000). El procedimiento incluye los siguientes pasos: (i) Fijación de la muestra, (ii) Preparación de la muestra, lo que puede incluir pre-tratamientos específicos, (iii) Hibridación con sondas específicas, (iv) Lavados del preparado para remover la sonda excedente, (v) Montaje, visualización y captura de imágenes (Figura 1.13).

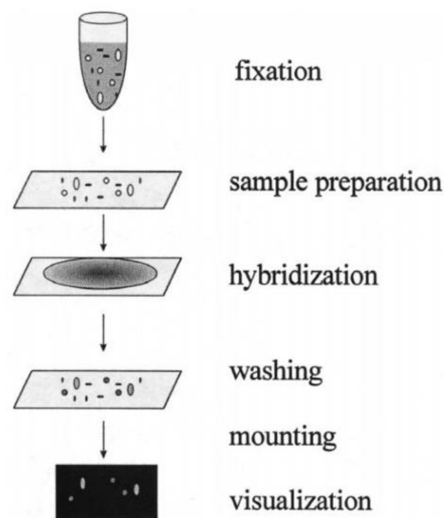

Figura 1.13. Esquema básico del procedimiento de la técnica de FISH (Moter and Gobel 2000).

En esta técnica, el fluorocromo unido al extremo 5´ es excitado por un láser a una longitud de onda específica. La emisión de fluorescencia es detectada mediante un microscopio de epifluorescencia o confocal. La técnica de FISH permite la identificación y visualización de células individuales (Amann and Fuchs 2008). En sistemas de tratamiento de aguas residuales, además de la identificación, esta técnica permite determinar la disposición espacial de bacterias filamentosas del filo Chloroflexi. Por ejemplo, en sistemas de lodos

activados permite determinar si se encuentran dentro de los flóculos, o interconectando flóculos.

1.4.2 Reacción en cadena de la polimerasa cuantitativa (q-PCR)

El q-PCR se ha empleado desde mediados de 1990 (Heid et al. 1996) y actualmente se considera el método más preciso para detectar y cuantificar secuencias específicas de ácidos nucleicos como ADN y ARN. Este método no solo detecta la presencia de secuencias de interés con primers específicos sino que cuantifica en números absolutos. La cuantificación de microorganismos facilita la vinculación de los cambios en la estructura de la comunidad con los cambios en variables ambientales o parámetros operacionales. En el PCR convencional la concentración de ADN es medida de forma semi-cuantitativa únicamente en el punto final de la amplificación. Debido a las limitaciones y sesgos del PCR la concentración de punto final no es proporcional a la concentración inicial de ADN molde (Zhang and Fang 2006). En el q-PCR se puede monitorear en tiempo real el progreso de la amplificación del ADN y visualizar la fase de amplificación exponencial (Heid et al. 1996). Esta monitorización en tiempo real, que es la clave para la cuantificación absoluta de las secuencias blanco, se consigue midiendo continuamente la fluorescencia emitida en la fase exponencial de amplificación. Entre las distintas químicas de detección utilizadas en q-PCR, el agente intercalante SYBR Green I (Figura 1.14 (A)) y las sondas TaqMan (Figura 1.14 (B)) son los métodos más utilizados.

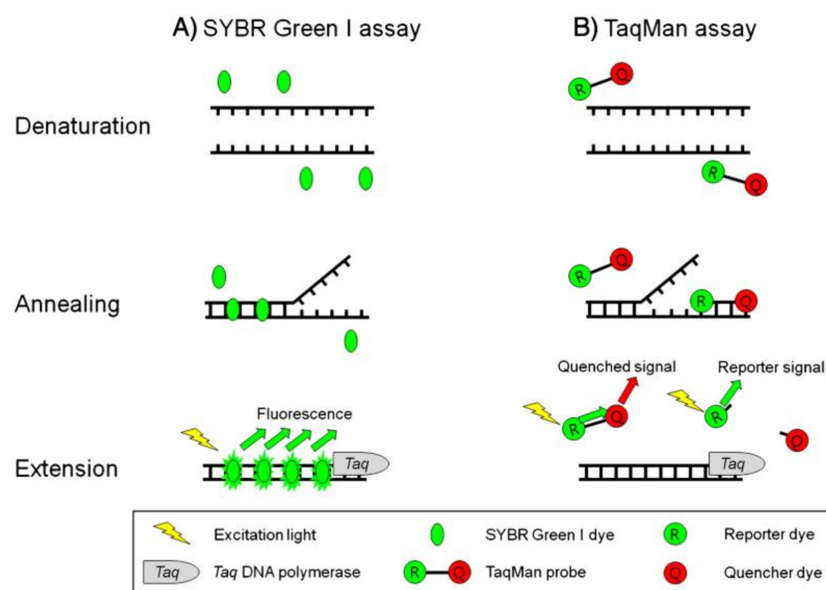

Figura 1.14. Químicas de detección más usadas en la q-PCR: (A) SYBR Green I y (B) sondas TaqMan (Kim et al. 2013).

El agente intercalante SYBR Green I se une inespecíficamente al surco menor de la doble hebra de ADN, emitiendo una fluorescencia 1000 veces mayor que cuando no está unido (Wittwer et al. 1997). Por tanto, la señal de fluorescencia aumenta a medida que se acumula más ADN de doble hebra durante la PCR permitiendo la monitorización de la reacción en tiempo real. Una de las limitaciones del uso del SYBR Green I que puede generar resultados falsos positivos, es la unión no específica al ADN doble hebra (ej, unión a dímeros de primers). En el caso de usar compuestos fluorescentes inespecíficos, la especificidad del q-PCR estará determinada únicamente por los primers elegidos. La especificidad de la amplificación se puede verificar mediante el análisis de las curvas de melting ya que un determinado producto de amplificación tiene una temperatura de melting (Tm) específica dependiente de la secuencia de ADN (Ririe et al. 1997). Por lo tanto, se puede visualizar la presencia de productos de amplificación no específicos que tienen Tm distintas a la secuencia blanco.

La química TaqMan implica el uso de un oligonucleótido adicional denominado sonda TaqMan, además de los primers utilizados en la amplificación por PCR. Dicha sonda tiene un marcado doble, un fluoróforo en el extremo 5' y una molécula quencher en el extremo 3'. En una sonda TaqMan intacta, el quencher absorbe la fluorescencia del fluoróforo a través de la transferencia de energía por resonancia de fluorescencia (FRET) (Giulietti et al. 2001). Durante la amplificación por PCR, las sondas TaqMan que se hibridan con las secuencias complementarias en el ADN blanco, se degradan por la actividad exonucleasa de la ADN polimerasa a medida que tiene lugar la extensión de la hebra de ADN. Por lo tanto, la separación del fluoróforo del quencher, permite la emisión de fluorescencia. Por lo tanto, la intensidad de la fluorescencia emitida indica la cantidad de amplicones producidos, permitiendo la monitorización en tiempo real de la amplificación por PCR. La utilización de una sonda TaqMan en conjunto con los primers, mejora significativamente la especificidad en comparación con el agente intercalante inespecífico SYBR Green I.

1.4.3 Terminal restriction fragment length polymorphism analysis (T-RFLP)

La técnica de T-RFLP permite la detección de fragmentos de ADN de distintas longitudes que son producto de una restricción enzimática. En el T-RFLP, se amplifica por PCR un gen de interés con primers universales o específicos, donde el primer forward está marcado con un fluorocromo en el extremo 5’. En un segundo paso, el producto de PCR es sometido a

digestión enzimática, generando fragmentos de restricción (T-RF) de distinto tamaño de acuerdo al sitio de corte de la enzima utilizada (Figura 1.15).

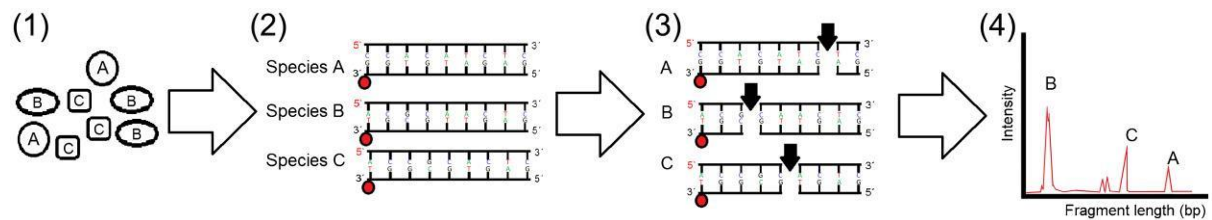

Figura 1.15. Esquema de la técnica de T-RFLP (Walker et al. 2017).

Los T-RF obtenidos son separados y detectados con un secuenciador de ADN del método de Sanger con un capilar automático. La diversidad microbiana en una muestra se puede estimar analizando el número y la altura de los fragmentos de T-RF obtenidos en el cromatograma, ya que distintos microorganismos, con diferente secuencia, tienen distintos sitios de restricción. La comparación del número y tipo de picos en las muestras puede dar una idea de la composición y diversidad de la comunidad microbiana en el tiempo (Liu et al. 1997; Dunbar et al. 2001). Sin embargo, la técnica es más adecuada para comunidades microbianas con una riqueza de baja a intermedia, ya que especies distintas pueden tener un mismo sitio de corte. Al igual que todas las técnicas basadas en amplificación de ADN por PCR, el T-RFLP tiene sesgos de amplificación y de la elección de los primers.

1.4.4 Técnicas de secuenciación masiva de segunda generación

Las tecnologías de secuenciación de segunda generación, permiten la secuenciación masiva de muestras en paralelo, reduciendo costos y tiempo en comparación a la secuenciación Sanger.

Las técnicas de secuenciación utilizadas en esta tesis se describen brevemente a continuación:

Pirosecuenciación o tecnología 454: fue la primera plataforma de secuenciación masiva de fragmentos cortos de ADN en salir al mercado entre 2004 y 2005 (Margulies et al. 2005). En esta tecnología, cada secuencia de ADN es ligada mediante un adaptador a una perla, la cual quedará embebida en una emulsión de aceite-agua que contiene todos los reactivos necesarios para la amplificación en emulsión por PCR. El diseño de este sistema permite que una única molécula de ADN se una a cada perla, generando una amplificación clonal del ADN (Figura 1.16 (A)). Las perlas junto a los reactivos de PCR quedarán ubicados en los pocillos de un microchip, donde se llevará a cabo la secuenciación del ADN por síntesis (Figura 1.16 (B)).

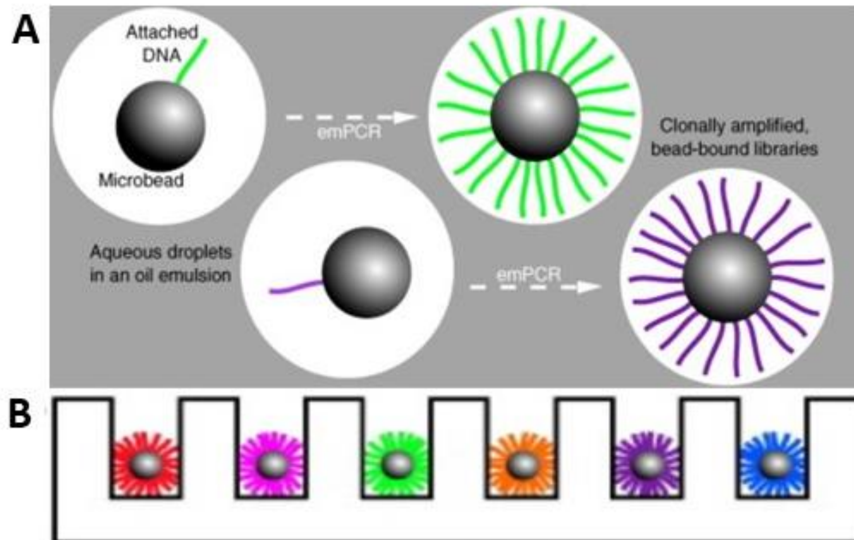

Figure 1.16. Esquema del proceso de Pirosecuenciación (Heather and Chain 2016).

La estrategia utilizada en pirosecuenciación, es la detección de bioluminiscencia producida por la acción de la luciferasa cada vez que se incorpora un dNTP en la extensión del ADN. La incorporación de un dNTP libera un grupo pirofosfato (Ppi), la ATP sulfurilasa convierte el Ppi en ATP en presencia de ADP, el ATP junto con la luciferasa cataliza la conversión de luciferina a oxiluciferina, emitiendo luz. La cantidad de luz emitida corresponde al número y orden de cada dNTP añadido durante la reacción.

Ion Torrent: La tecnología de secuenciación Ion Torrent salió al mercado en 2010 y es muy similar a la plataforma 454, difieren en el método de detección utilizado. En esta plataforma, se detectan los protones liberados por la polimerasa durante de extensión de la cadena de ADN luego de cada adición de nucleótidos (A, G, C o T). La liberación de protones resultante genera una variación de pH que provoca la acumulación de carga positiva en la placa de detección. La variación de pH refleja cuántos nucleótidos se agregaron en la elongación de la cadena. Al terminar con la incorporación y lectura de un determinado nucleótido, se efectúa un lavado y se agrega otro nucleótido distinto para continuar con la secuenciación.

Illumina: La tecnología Illumina fue lanzada al mercado en 2006 (www.illumina.com), y difiere de las anteriores principalmente en que: (1) La librería de ADN con los adaptadores se depositan en una flow-cell que contiene en su superficie adaptadores complementarios, (2) El método de secuenciación es mediante terminadores de cadena reversibles (método similar a la

secuenciación de Sanger). El paso previo consiste en la unión complementaria de los adaptadores del ADN a la flow-cell. La amplificación en Illumina se denomina amplificación "por puente" (otra de las diferencias con las plataformas anteriores), donde se obtienen secuencias clonales del fragmento inicial (Figura 1.17 (A, B)).

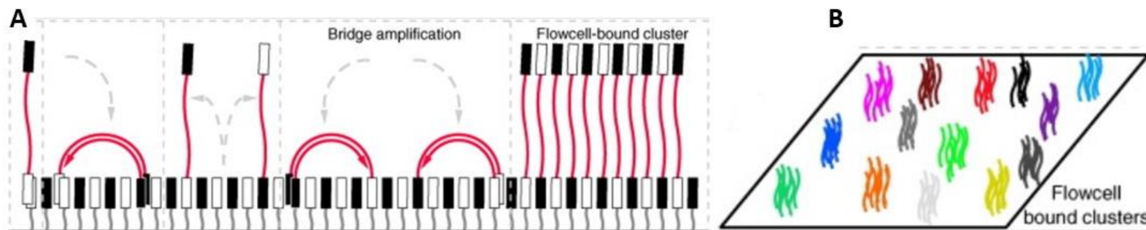

Figura 1.17. Esquema del proceso Illumina (Heather and Chain 2016).

1.5.5 Procesamiento de datos obtenidos mediante secuenciación masiva de amplicones

Las principales preguntas enfocadas en muestras ambientales refieren a cuál es la composición taxonómica de la comunidad y la abundancia de sus individuos, y cómo se relacionan con el ambiente. Para llegar a obtener estos resultados, es necesario comenzar con el pre-procesamiento de las secuencias crudas. Este paso es especialmente necesario para datos de secuenciación de amplicones. En este proceso, se eliminan secuencias erróneas y de baja calidad, así como artefactos de amplificación (quimeras), lo que mejora significativamente la precisión del análisis (Bakker et al. 2012; D'Amore et al. 2016). Además, el control de calidad es esencial para evitar una sobreestimación de la diversidad de especies de la comunidad. Si este paso se omite, es probable que se introduzcan errores durante la agrupación de secuencias similares en la construcción de las "unidades taxonómicas operacionales" (OTUs) (Huse et al. 2010) o en la determinación de las 'amplicon sequence variants' (ASV) (Callahan et al. 2017). Brevemente, el control de calidad generalmente comprende (no necesariamente en este orden): [i] eliminación de adaptadores de secuenciación, primers y barcodes (código de barras que identifica cada secuencia con su respectiva muestra, introducido en la etapa de preparación de la librería), filtrado y/o corte de secuencias de baja calidad, de-replicación (identificación de secuencias idénticas y selección de una secuencias representativa con un recuento de abundancia) para disminuir el volumen de datos, y la detección y eliminación de secuencias quiméricas (generadas por la hibridación cruzada de diferentes fragmentos de ADN durante el PCR).

Después del pre-procesamiento, el análisis se puede dividir en tres bloques principales: (1) determinación de OTU o ASV, (2) clasificación taxonómica y (3) análisis multivariados. En referencia al paso (1), los avances recientes han hecho posible analizar datos de

secuenciación sin recurrir a la construcción habitual OTUs, definidas como un grupo de secuencias que difieren entre ellas en menos de un umbral de disimilitud fijo, generalmente del 97%. Los nuevos métodos de análisis controlan los errores lo suficiente como para que las variantes de las secuencias de amplicones (ASV) tengan una resolución de un solo nucleótido entre secuencias (Callahan et al. 2017). La clasificación taxonómica de las OTUs o ASVs está basada en búsquedas de similitud de sus secuencias, con secuencias de bases de datos de referencia de genes de ARNr 16S (ej: RDP, SILVA, Greengenes, MiDAS).

1.5.6 Aplicaciones de la secuenciación masiva de amplicones del gen de ARNr 16S para el estudio de comunidades microbianas

El uso de la secuenciación masiva de amplicones para el estudio de las comunidades microbianas complejas ha aumentado drásticamente en los últimos años. Las razones incluyen la disminución en los costos de secuenciación y la capacidad de caracterizar con una mayor profundidad las comunidades microbianas, lo que permite la detección de taxones "raros" (<0.1%) (Sogin et al. 2006). En el caso del estudio de la diversidad y abundancia de bacterias y Archaeas presentes en las comunidades microbianas, los primers seleccionados en general están dirigidos a alguna de las nueve regiones hipervariables del gen de ARNr 16S (V1-V9) (Pace et al. 1986; Hugenholtz and Pace 1996; Clooney et al. 2016). Desafortunadamente, no existe un set de primers más apropiado para secuenciar eficientemente todos los taxones y tipos de ecosistemas, esta decisión en general se toma en base a consejos de la literatura publicada o pruebas *in silico* (Clooney et al. 2016). Una de las principales limitaciones de la secuenciación de amplicones es la longitud de los mismos, que actualmente ronda los 200-500pb. Considerando que el gen de ARNr 16S contiene 1542 pb, en la mayoría de los casos solo es posible alcanzar una clasificación taxonómica hasta nivel de género o familia, alcanzando pocas veces a una clasificación a nivel de especie (Zhang et al. 2011). Aunque algunas regiones hipervariables del gen de ARNr 16S proporcionan una buena aproximación de la diversidad de la comunidad, la mayoría no distingue polimorfismos entre taxones estrechamente relacionados, lo que podría evitarse mediante la secuenciación completa del gen de ARNr 16S (Johnson et al. 2019). Además, la presencia de múltiples copias intragenómicas de este gen, podrían llevar a sustituciones sutiles de nucleótidos sobreestimando la abundancia microbiana (Johnson et al. 2019). Dentro de las limitaciones de esta técnica se encuentran los sesgos asociados con la técnica de PCR y la elección de los

primers, estos últimos limitan la resolución de una fracción sustancial de la diversidad en la comunidad (Hong et al. 2009; Sharpton et al. 2011; Logares et al. 2014).

La secuenciación de amplicones del gen del ARNr 16S proporciona información sobre la composición taxonómica de la comunidad, no pudiendo resolver directamente las funciones biológicas asociadas con estos taxones. En algunos casos, la reconstrucción filogenética se puede utilizar para inferir las funciones biológicas que están codificadas en un genoma que contiene una secuencia del gen del ARNr de16S similar (Langille et al. 2013). Pero, la precisión con la que estos métodos estiman la verdadera diversidad funcional de una comunidad está ligada a qué tan bien la diversidad genómica de la comunidad está representada por los genomas disponibles en las bases de datos. Por último, debido a que puede haber transferencia horizontal del gen de ARNr 16S entre taxones relacionados pero lejanamente, el análisis de las secuencias del gen de ARNr 16S puede resultar en sobre-estimaciones de la diversidad de la comunidad (Acinas et al. 2004).

1.5.7 Procesamiento de datos obtenidos mediante secuenciación de metagenomas

A diferencia de los estudios basados en secuenciación de amplicones, la metagenómica evita los sesgos relacionados a utilización de primers y amplificación por PCR. El control de calidad sobre las secuencias obtenidas en los metagenomas, se lleva a cabo de forma análoga a los pasos aplicados en análisis de secuenciación de amplicones, omitiendo los pasos que tratan en particular con los errores y artefactos relacionados con la amplificación.

A continuación, se describen algunos de los enfoques para abordar el análisis de datos de metagenómica (Figura 1.18):

(1) Clasificación taxonómica de las secuencias del ARNr 16S presentes en el metagenoma. Es adecuado para responder preguntas relacionadas con la composición taxonómica de una muestra o relacionadas con la presencia o ausencia de organismos.

(2) Ensamblado de las secuencias de forma contigua (por alineamiento de secuencia), generando secuencias de mayor longitud denominadas contigs. A partir de este punto existen dos caminos principales: a) La anotación funcional de los contigs para determinar el perfil metabólico de la comunidad total; b) Los contigs son sometidos al proceso de binning, donde son agrupados según algunas características como contenido en GC o frecuencias de tetrámeros, generando bines (o genomas). Posteriormente, cada bin es anotado para determinar su potencial metabólico y/o se determinan relaciones filogenómicas a partir del alineamiento de genes marcadores de copia única presentes en cada bin.

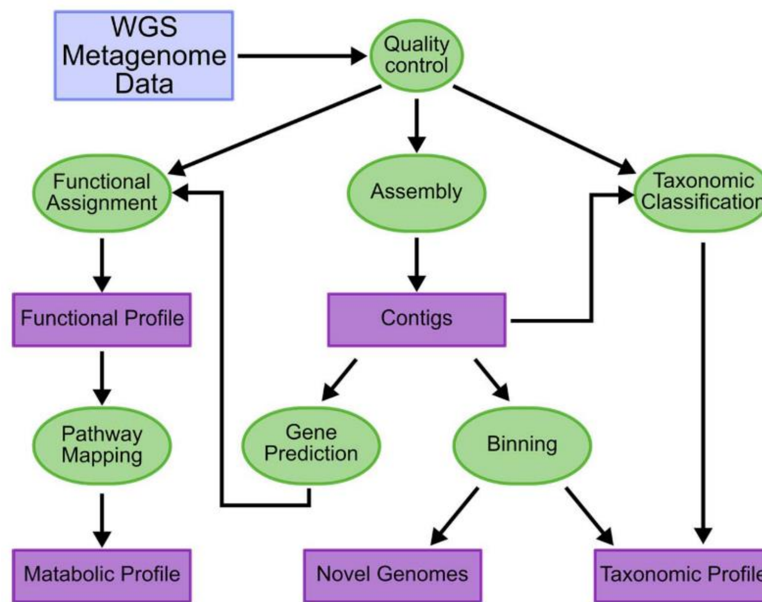

Figura 1.18. Esquema simplificado de los principales pasos en el análisis de metagenomas. Los rectángulos indican resultados, y los óvalos representan pasos del procesamiento (Jünemann et al. 2017). WGS= whole generation sequence.

1.5.8 Aplicaciones de la Metagenómica al estudio de comunidades microbianas

Mediante la secuenciación de metagenomas se obtiene información sobre el potencial metabólico de la comunidad completa (procariotas, eucariotas, virus, etc) (Su et al. 2012; Shakya et al. 2013). El análisis de metagenomas puede tener varios objetivos, como determinar la diversidad y abundancia de las comunidades, determinar sus funciones metabólicas, recuperar genomas presentes en la muestra y determinación de filogenia y metabolismo (Sharpton 2014). Debido a que todo el ADN genómico presente en la muestra es secuenciado, es posible explorar dos aspectos de una comunidad microbiana: ¿qué microorganismos están presentes y que son capaces de hacer?

El análisis de datos de metagenómica presenta varios desafíos. La complejidad y tamaño de los datos dificulta el análisis bioinformático, y la mayoría de las comunidades son tan diversas que los genomas recuperados generalmente no están completamente representados por las secuencias obtenidas. Como resultado, es posible que dos secuencias pertenecientes al mismo gen no se superpongan y no puedan ser ensambladas (Schloss and Handelsman 2008; Sharpton et al. 2011). Cuando las secuencias se superponen, no siempre es evidente si son del mismo genoma o de distintos genomas, lo que puede desafiar el ensamblado de secuencias (Mavromatis et al. 2007; Mende et al. 2012).

1.5 Marco del trabajo y objetivos

1.5.1 Marco del trabajo: persistencia de microorganismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales

La principal línea de investigación de nuestro laboratorio es el estudio de la microbiología de sistemas de tratamiento de aguas residuales. Para ello aplicamos técnicas de biología molecular que nos permiten estudiar las comunidades sin aislar los microorganismos. Se trabaja en colaboración con grupos de Ingeniería de reactores de nuestro país (BioProA, Facultad de Ingeniería, UDELAR) y de otros países de Latinoamérica y Europa. Los colegas ingenieros diseñan y operan los reactores y nuestro laboratorio se centra en estudiar la microbiología involucrada en los procesos con la finalidad de poder optimizarlos y conocer cómo funcionan. Se han estudiado sistemas anaerobios metanogénicos a escala real (Callejas et al. 2019; Cabezas et al. 2020), sistemas de laboratorio de producción de hidrógeno (Castelló et al. 2009; Perna et al. 2013; Ferraz Júnior et al. 2015) y sistemas de remoción de nitrógeno convencionales (Cabezas et al. 2009) y sistemas anammox (Costa et al. 2014). A lo largo de estos años fue sorprendente encontrar que, salvo en los reactores de producción de hidrógeno (Etchebehere et al. 2016), en todos los otros reactores estudiados tanto aerobios como anaerobios el filo Chloroflexi estuvo presente en las comunidades microbianas estudiadas. Sin embargo, hasta el momento se conoce poco del rol que juegan estos microorganismos en los sistemas de tratamiento de aguas residuales.

El interés de nuestro laboratorio por el estudio del filo Chloroflexi comenzó antes del año 2001, en el trabajo de Tesis de Doctorado de la Dra. Etchebehere. En esta Tesis se estudió la microbiología de un reactor UASB desnitrificante alimentado con lixiviado de relleno sanitario. Con las técnicas de PCR, clonado y secuenciación de Sanger de genes del ARNr 16S, se observó que dos clones que representaban el 15 % de la comunidad se clasificaron dentro del filo Green-non-Sulfur Bacteria (posteriormente renombrado como Chloroflexi). En ese trabajo se postuló que estos microorganismos podrían estar relacionados con la degradación de compuestos recalcitrantes presentes en el reactor (Etchebehere et al. 2002).

En el año 2008, nuestros colegas del laboratorio BioProA de Facultad de Ingeniería (liderado por la Dra. Liliana Borzacconi) detectaron pérdida de eficiencia en un reactor metanogénico de una industria procesadora de malta debido al sobrecrecimiento de bacterias filamentosas. Mediante la técnica de FISH con sondas específicas para el filo Chloroflexi, se determinó que este grupo era el responsable de la pérdida de densidad de los gránulos y su flotación (Figura 5). Se hipotetizó que su sobrecrecimiento podría deberse al alto contenido de células en el

reactor producida por lisis celular debida a una bajada abrupta del pH. Los resultados además de buscar soluciones al problema de la industria permitieron publicar un trabajo en conjunto (Borzacconi et al. 2008).

En el año 2012 obtuvimos financiación del Fondo Clemente Estable (FCE_2_2011_1_7062) para estudiar la diversidad y abundancia de Chloroflexi en reactores metanogénicos escala real en nuestro país. Con este proyecto se inició este posgrado de Biotecnología el cual se financió con beca de Maestría y Doctorado de la ANII y de finalización de CAP. Por otro lado, la Dra. Angela Cabezas realizó una pasantía en el laboratorio de la Dra. Anne Kristin Kaster, en ese entonces directora del grupo de *Single Cell Genomics* de DSMZ (Alemania), para secuenciar genomas de Chloroflexi detectados en un sistema aerobio de una planta vitivinícola de Uruguay mediante Fluorescence-activated cell sorting (FACS) y *Single Cell Genomics*. Este trabajo fue recientemente publicado en Frontiers in Microbiology (Dam et al. 2020).

En 2018 el proyecto Fondo Vaz Ferreira (I/FVF2017/135) liderado por la Dra. Angela Cabezas tenía como objetivo secuenciar metagenomas y ensamblar genomas de Chloroflexi de un reactor anaerobio y de un reactor aerobio. Al mismo tiempo, Patricia Bovio obtuvo financiación a través de la beca de movilidad ANII (MOV_CA_2017_1_138200) para su capacitación en el análisis de metagenomas en el Instituto de Investigaciones en Ingeniería Genética y Biología Molecular "Dr Héctor N. Torres" (INGEBI-CONICET, Buenos Aires) bajo la orientación del Dr. Leonardo Erijman y el Dr. Leandro Guerrero. Esto permitió ensamblar y analizar 8 genomas de especies nuevas de organismos del filo Chloroflexi distribuidos en distintas familias. Gracias a la formación de Patricia Bovio en el análisis de metagenomas, recibimos en 2019 una estudiante de Doctorado de la Universidad Autónoma de Barcelona, para formar en el análisis de metagenomas de reactores anammox generando nuevas colaboraciones.

En esta Tesis se abordó un tema muy poco explorado de Microbiología básica aplicado a entender cómo funcionan los sistemas de tratamiento de aguas residuales. A lo largo del trabajo se fueron aplicando nuevas metodologías siendo un gran desafío para este trabajo de Tesis la puesta a punto de las mismas. Se generó conocimiento en un tema muy poco conocido y con gran potencial de aplicación en Biotecnología Ambiental.

1.5.2 Objetivo general

Conocer la composición taxonómica y función de microorganismos del filo Chloroflexi en sistemas de tratamiento de aguas residuales y residuos sólidos.

Preguntas a responder en este trabajo de Posgrado en Biotecnología

1- ¿Cuál es la composición taxonómica de los microorganismos del filo Chloroflexi en reactores anaerobios metanogénicos a escala real?

2- ¿Son los mismos grupos que predominan en todos los reactores metanogénicos o son diferentes dependiendo del sustrato y tipo de reactor?

3- ¿Cuál es su abundancia y de qué depende?

6- ¿Los microorganismos del filo Chloroflexi presentes en reactores anaerobios metanogénicos, aerobios de lodos activados y anammox son los mismos?

7- ¿Tienen el mismo rol en los diferentes tipos de reactores?

8- ¿Es posible conocer el rol a partir de datos de metagenomas?

1.5.3 Objetivos específicos

1. Conocer la diversidad de organismos del filo Chloroflexi en reactores metanogénicos escala real instalados en Uruguay.
2. Conocer la diversidad y distribución global de organismos del filo Chloroflexi en reactores metanogénicos de tratamiento de aguas residuales y de residuos sólidos a escala real, a partir del meta-análisis de datos recuperados de bases de datos.
3. Comparar la composición taxonómica y el metabolismo de organismos del filo Chloroflexi en reactores de lodos activados, anammox y metanogénicos utilizando técnicas de ensamblado de genomas.

1.5.4 Estrategia de investigación

Objetivo específico 1: Conocer la diversidad, composición taxonómica y abundancia de organismos del filo Chloroflexi en reactores metanogénicos escala real instalados en Uruguay

Un primer abordaje fue estudiar específicamente reactores metanogénicos a escala real que estaban operando en nuestro país y que podían ser monitoreados a lo largo del tiempo. Debido a la dificultad en el aislamiento de organismos del filo Chloroflexi, el estudio de este grupo de bacterias se decidió abordar con distintas técnicas de biología molecular como T-RFLP, FISH, q-PCR y secuenciación masiva del gen del ARNr 16S. En las tres primeras

técnicas se usaron primers específicos para filo Chloroflexi. Se determinó la diversidad, abundancia y morfología de de este grupo en 5 reactores UASB escala real instalados en Uruguay. Los resultados de este trabajo fueron publicados en el año 2019 en *Journal of Applied Microbiology* bajo el título de "Preliminary analysis of Chloroflexi populations in full-scale UASB methanogenic reactors" (DOI: 10.1111/jam.14115), y se presentan en el Capítulo 2.

Objetivo específico 2- Conocer la diversidad, abundancia y distribución global de organismos del filo Chloroflexi en reactores metanogénico de aguas residuales y de residuos sólidos escala real, a partir del meta-análisis de datos recuperados de bases de datos
Con el objetivo de ampliar la diversidad de configuraciones de reactores metanogénicos y de sustratos, nos planteamos realizar un meta-análisis a partir de sets de datos de secuenciación masiva del gen de ARNr 16S depositados en bases de datos. En este trabajo logramos incluir datos de 17 artículos publicados que representaban 62 reactores de tratamiento de aguas residuales y digestores de sólidos. Este tipo de análisis presentó el gran desafío de analizar datos crudos provenientes de distintas plataformas de secuenciación, utilizando distintos set de primers y por lo tanto, regiones amplificadas del gen de ARNr 16S no superpuestas. Los resultados de este trabajo se publicaron en el año 2021 en *Frontiers in Microbiology* bajo el título de "Database Mining to Unravel the Ecology of the Phylum Chloroflexi in Methanogenic Full Scale Bioreactors" (DOI: 10.3389/fmicb.2020.603234), y se presentan en el Capítulo 3.

Objetivo específico 3- Comparar la composición taxonómica y el metabolismo de organismos del filo Chloroflexi en reactores de lodos activados, anammox y metanogénicos utilizando técnicas de ensamblado de genomas
En los capítulos 2 y 3 se estudió principalmente la diversidad y abundancia del filo Chloroflexi en reactores metanogénicos. El siguiente paso, además de conocer más sobre la diversidad y abundancia, fue abordar la pregunta de qué rol cumple el filo Chloroflexi en distintos sistemas de tratamiento de aguas residuales como, por ejemplo, un reactor metanogénico, un reactor de lodos activados y un reactor anammox. Y además, determinar si las vías metabólicas presentes en los genomas eran distintas entre los distintos reactores. Mediante la secuenciación de metagenomas de los distintos reactores y el ensamblado de los genomas de Chloroflexi, se determinó el potencial metabólico de este grupo en cada sistema. Los resultados se presentan en el capítulo 4 y el manuscrito se encuentra en preparación con

el título: Functional redundancy of diverse Anaerolineae genomes in activated sludge, methanogenic reactors and anammox reactors.

1.6 Referencias introducción general

Acinas SG, Marcelino LA, Klepac-Ceraj V, Polz MF (2004) Divergence and Redundancy of 16S rRNA Sequences in Genomes with Multiple rrn Operons. J Bacteriol 186:2629–2635. doi: 10.1128/JB.186.9.2629-2635.2004

Adhikari JR, Lohani SP (2019) Design, installation, operation and experimentation of septic tank – UASB wastewater treatment system. Renew Energy 143:1406–1415. doi: 10.1016/j.renene.2019.04.059

Ahn J, McIlroy S, Schroeder S, Seviour R (2009) Biomass granulation in an aerobic:anaerobic-enhanced biological phosphorus removal process in a sequencing batch reactor with varying pH. J Ind Microbiol Biotechnol 36:885–893. doi: 10.1007/s10295-009-0566-3

Ahn YH (2006) Sustainable nitrogen elimination biotechnologies: A review. Process Biochem. 41:1709–1721

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 31:533–538. doi: 10.1038/nbt.2579

Ali M, Chai LY, Min XB, Tang CJ, Afrin S, Liao Q, Wang HY, Peng C, Song YX, Zheng P (2016) Performance and characteristics of a nitritation air-lift reactor under long-term HRT shortening. Int Biodeterior Biodegrad 111:45–53. doi: 10.1016/j.ibiod.2016.04.003

Amani T, Nosrati M, Sreekrishnan TR (2010) Anaerobic digestion from the viewpoint of microbiological, chemical, and operational aspects — a review. Environ Rev 18:255–278 . doi: 10.1139/A10-011

Amann R, Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. Nat. Rev. Microbiol. 6:339–348

Andersen MH, McIlroy SJ, Nierychlo M, Nielsen PH, Albertsen M (2018) Genomic insights into *Candidatus* Amarolinea aalborgensis gen. nov., sp. nov., associated with settleability problems in wastewater treatment plants. Syst Appl Microbiol 1–8. doi: 10.1016/j.syapm.2018.08.001

Angenent LT, Karim K, Al-Dahhan MH, Wrenn BA, Domíguez-Espinosa R (2004) Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends Biotechnol 22:477–485. doi: 10.1016/j.tibtech.2004.07.001

Appels L, Baeyens J, Degrève J, Dewil R (2008) Principles and potential of the anaerobic digestion of waste-activated sludge. Prog. Energy Combust. Sci. 34:755–781

Appels L, Lauwers J, Degrve J, Helsen L, Lievens B, Willems K, Van Impe J, Dewil R (2011) Anaerobic digestion in global bio-energy production: Potential and research challenges. Renew. Sustain. Energy Rev. 15:4295–4301

Bakker MG, Tu ZJ, Bradeen JM, Kinkel LL (2012) Implications of Pyrosequencing Error Correction for Biological Data Interpretation. PLoS One 7:44357. doi: 10.1371/journal.pone.0044357

Batchelor SE, Cooper M, Chhabra SR, Glover LA, Stewart GS, Williams P, Prosser JI (1997) Cell density-regulated recovery of starved biofilm populations of ammonia-oxidizing bacteria. Appl Environ Microbiol 63:

Björnsson L, Hugenholtz P, Tyson GW, Blackall LL (2002) Filamentous Chloroflexi (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. Microbiology 148:2309–2318

Borzacconi L, López I, Passeggi M, Etchebehere C, Barcia R (2008) Sludge deterioration in a full scale UASB reactor after a pH drop working under low loading conditions. Water Sci Technol 57:797–802. doi: 10.2166/wst.2008.071

Bovio P, Cabezas A, Etchebehere C (2019) Preliminary analysis of Chloroflexi populations in full-scale UASB methanogenic reactors. J Appl Microbiol 126:667–683. doi: 10.1111/jam.14115

Cabezas A, Bovio P, Etchebehere C (2020) Commercial formulation amendment transiently affects the microbial composition but not the biogas production of a full scale methanogenic UASB reactor. Environ Technol (United Kingdom) 41:3119–3133. doi: 10.1080/09593330.2019.1600042

Cabezas A, de Araujo JC, Callejas C, Galès A, Hamelin J, Marone A, Sousa DZ, Trably E, Etchebehere C (2015b) How to use molecular biology tools for the study of the anaerobic digestion process? Rev Environ Sci Bio/Technology 14:555–593. doi: 10.1007/s11157-015-9380-8

Cabezas A, Draper P, Etchebehere C (2009) Fluctuation of microbial activities after influent load variations in a full-scale SBR: Recovery of the biomass after starvation. Appl Microbiol Biotechnol 84:1191–1202. doi: 10.1007/s00253-009-2138-x

Callahan BJ, McMurdie PJ, Holmes SP (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. ISME J 11:2639–2643. doi: 10.1038/ismej.2017.119

Callejas C, Fernández A, Passeggi M, Wenzel J, Bovio P, Borzacconi L, Etchebehere C (2019) Microbiota adaptation after an alkaline pH perturbation in a full-scale UASB anaerobic reactor treating dairy wastewater. Bioprocess Biosyst Eng 42:2035–2046. doi: 10.1007/s00449-019-02198-3

Campbell AG, Schwientek P, Vishnivetskaya T, Woyke T, Levy S, Beall CJ, Griffen A, Leys E, Podar M (2014) Diversity and genomic insights into the uncultured Chloroflexi from the human microbiota. Environ Microbiol 16:2635–2643. doi: 10.1111/1462-2920.12461

Cao S, Du R, Li B, Ren N, Peng Y (2016) High-throughput profiling of microbial community structures in an ANAMMOX-UASB reactor treating high-strength wastewater. Appl Microbiol Biotechnol 100:6457–6467. doi: 10.1007/s00253-016-7427-6

Cardenas E, Orellana LH, Konstantinidis KT, Mohn WW (2018) Effects of timber harvesting on the genetic potential for carbon and nitrogen cycling in five North American forest ecozones. Sci Rep 8:1–13. doi: 10.1038/s41598-018-21197-0

Carr EL, Eales KL, Seviour RJ (2006) Substrate uptake by Gordonia amarae in activated sludge foams by FISH-MAR. Water Sci Technol 54:39–45. doi: 10.2166/wst.2006.369

Castelle CJ, Banfield JF (2018) Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. Cell 172:1181–1197

Castelló E, García y Santos C, Iglesias T, Paolino G, Wenzel J, Borzacconi L, Etchebehere C (2009) Feasibility of biohydrogen production from cheese whey using a UASB reactor: Links between microbial community and reactor performance. Int J Hydrogen Energy 34:5674–5682. doi: 10.1016/j.ijhydene.2009.05.060

Chan YJ, Chong MF, Law CL, Hassell DG (2009) A review on anaerobic-aerobic treatment of industrial and municipal wastewater. Chem Eng J 155:1–18. doi: 10.1016/j.cej.2009.06.041

Chen S, Dong X (2005) Proteiniphilum acetatigenes gen. nov., sp. nov., from a UASB reactor treating brewery wastewater. Int J Syst Evol Microbiol 55:2257–2261. doi: 10.1099/ijs.0.63807-0

Chernicharo CAL (2006) Post-treatment options for the anaerobic treatment of domestic wastewater. Rev. Environ. Sci. Biotechnol. 5:73–92

Cho S, Fujii N, Lee T, Okabe S (2011) Development of a simultaneous partial nitrification and anaerobic ammonia oxidation process in a single reactor. Bioresour Technol 102:652–659. doi: 10.1016/j.biortech.2010.08.031

Chong S, Sen TK, Kayaalp A, Ang HM (2012) The performance enhancements of upflow anaerobic sludge blanket (UASB) reactors for domestic sludge treatment - A State-of-the-art review. Water Res 46:3434–3470. doi: 10.1016/j.watres.2012.03.066

Chu Z rui, Wang K, Li X kun, Zhu M ting, Yang L, Zhang J (2015) Microbial characterization of aggregates within a one-stage nitritation-anammox system using high-throughput amplicon sequencing. Chem Eng J 262:41–48. doi: 10.1016/j.cej.2014.09.067

Clooney AG, Fouhy F, Sleator RD, O'Driscoll A, Stanton C, Cotter PD, Claesson MJ (2016) Comparing apples and oranges?: Next generation sequencing and its impact on microbiome analysis. PLoS One 11:1–16. doi: 10.1371/journal.pone.0148028

Conrad R (1999) Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments. FEMS Microbiol Ecol 28:193–202. doi: 10.1111/j.1574-6941.1999.tb00575.x

Costa MCMS, Carvalho L, Leal CD, Dias MF, Martins KL, Garcia GB, Mancuelo ID, Hipólito T, Macconell EFA, Okada D, Etchebehere C, Chernicharo CAL, Calabria Araujo J (2014) Impact of inocula and operating conditions on the microbial community structure of two anammox reactors. Environ Technol (United Kingdom) 35:1811–1822. doi: 10.1080/09593330.2014.883432

D'Amore R, Ijaz UZ, Schirmer M, Kenny JG, Gregory R, Darby AC, Shakya M, Podar M, Quince C, Hall N (2016) A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics 17:55. doi: 10.1186/s12864-015-2194-9

Daffonchio D, Thaveesri J, Verstraete W (1995) Contact angle measurement and cell hydrophobicity of granular sludge from upflow anaerobic sludge bed reactors. Appl Environ Microbiol 61:3676–3680. doi: 10.1128/aem.61.10.3676-3680.1995

Dam, H. T., Vollmers, J., Sobol, M. S., Cabezas, A., & Kaster, A. K. (2020). Targeted cell sorting combined with single cell genomics captures low abundant microbial dark matter with higher sensitivity than metagenomics. Frontiers in microbiology, 11, 1377.

de Bok FAM, Harmsen HJM, Plugge CM, de Vries MC, Akkermans ADL, de Vos WM, Stams AJM (2005) The first true obligately syntrophic propionate-oxidizing bacterium, Pelotomaculum schinkii sp. nov., co-cultured with Methanospirillum hungatei, and emended description of the genus Pelotomaculum. Int J Syst Evol Microbiol 55:1697–1703. doi: 10.1099/ijs.0.02880-0

De Vrieze J, Saunders AM, He Y, Fang J, Nielsen PH, Verstraete W, Boon N (2015) Ammonia and temperature determine potential clustering in the anaerobic digestion microbiome. Water Res 75:312–323. doi: 10.1016/j.watres.2015.02.025

Deepnarain N, Nasr M, Kumari S, Stenström TA, Reddy P, Pillay K, Bux F (2020) Artificial intelligence and

multivariate statistics for comprehensive assessment of filamentous bacteria in wastewater treatment plants experiencing sludge bulking. Environ Technol Innov 19:100853. doi: 10.1016/j.eti.2020.100853

Demirel B, Scherer P (2008) The roles of acetotrophic and hydrogenotrophic methanogens during anaerobic conversion of biomass to methane: A review. Rev. Environ. Sci. Biotechnol. 7:173–190

Demirel B, Yenigun O, Onay TT (2005) Anaerobic treatment of dairy wastewaters: a review. Process Biochem 40:2583–2595. doi: 10.1016/j.procbio.2004.12.015

Dodsworth JA, Gevorkian J, Despujos F, Cole JK, Murugapiran SK, Ming H, Li WJ, Zhang G, Dohnalkova A, Hedlund BP (2014) Thermoflexus hugenholtzii gen. nov., sp. nov., a thermophilic, microaerophilic, filamentous bacterium representing a novel class in the Chloroflexi, Thermoflexia classis nov., and description of Thermoflexaceae fam. nov. and Thermoflexales ord. nov. Int J Syst Evol Microbiol 64:2119–2127. doi: 10.1099/ijs.0.055855-0

Dunbar J, Ticknor LO, Kuske CR (2001) Phylogenetic Specificity and Reproducibility and New Method for Analysis of Terminal Restriction Fragment Profiles of 16S rRNA Genes from Bacterial Communities Phylogenetic Specificity and Reproducibility and New Method for Analysis of Terminal Restrictio. Appl Environ Microbiol 67:190–197. doi: 10.1128/AEM.67.1.190

Eikelboom DH (1975) Filamentous organisms observed in activated sludge. Water Res 9:365–388. doi: 10.1016/0043-1354(75)90182-7

Eikelboom DH, van Buijsen HJJ (1983) Microscopic sludge investigation manual. https://www.worldcat.org/title/microscopic-sludge-investigation-manual/oclc/11554105. Accessed 27 Mar 2021

Etchebehere C, Castelló E, Wenzel J, del Pilar Anzola-Rojas M, Borzacconi L, Buitrón G, Cabrol L, Carminato VM, Carrillo-Reyes J, Cisneros-Pérez C, Fuentes L, Moreno-Andrade I, Razo-Flores E, Filippi GR, Tapia-Venegas E, Toledo-Alarcón J, Zaiat M (2016) Microbial communities from 20 different hydrogen-producing reactors studied by 454 pyrosequencing. Appl Microbiol Biotechnol 100:3371–3384. doi: 10.1007/s00253-016-7325-y

Etchebehere C, Errazquin MI, Dabert P, MuxÃ- L (2002) Community analysis of a denitrifying reactor treating landfill leachate. FEMS Microbiol Ecol 40:97–106. doi: 10.1111/j.1574-6941.2002.tb00941.x

Fan N, Wang R, Qi R, Gao Y, Rossetti S, Tandoi V, Yang M (2018) Control strategy for filamentous sludge bulking: Bench-scale test and full-scale application. Chemosphere 210:709–716 . doi: 10.1016/j.chemosphere.2018.07.028

Ferraz Júnior ADN, Etchebehere C, Zaiat M (2015) High organic loading rate on thermophilic hydrogen production and metagenomic study at an anaerobic packed-bed reactor treating a residual liquid stream of a Brazilian biorefinery. Bioresour Technol 186:81–88 . doi: 10.1016/j.biortech.2015.03.035

Fotidis IA, Karakashev D, Angelidaki I (2014) The dominant acetate degradation pathway/methanogenic composition in full-scale anaerobic digesters operating under different ammonia levels. Int J Environ Sci Technol 11:2087–2094. doi: 10.1007/s13762-013-0407-9

Fuhrman JA, Cram JA, Needham DM (2015) Marine microbial community dynamics and their ecological interpretation. Nat. Rev. Microbiol. 13:133–146

Garrity GM, Holt JG, Castenholz RW, Pierson BK, Keppen OI, Gorlenko VM (2001) Phylum BVI. Chloroflexi phy. nov. In: Bergey’s Manual® of Systematic Bacteriology. Springer New York, pp 427–446

Gerardi MH (2002) Nitrification and Denitrification in the Activated Sludge Process. John Wiley & Sons, Inc., Hoboken, NJ, USA

Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, Mathieu C (2001) An overview of real-time quantitative PCR: Applications to quantify cytokine gene expression. Methods 25:386–401. doi: 10.1006/meth.2001.1261

Grégoire P, Fardeau ML, Joseph M, Guasco S, Hamaide F, Biasutti S, Michotey V, Bonin P, Ollivier B (2011) Isolation and characterization of Thermanaerothrix daxensis gen. nov., sp. nov., a thermophilic anaerobic bacterium pertaining to the phylum “Chloroflexi”, isolated from a deep hot aquifer in the Aquitaine Basin. Syst Appl Microbiol 34:494–497. doi: 10.1016/j.syapm.2011.02.004

Grotenhuis JTC, Plugge CM, Stams AJM, Zehnder AJB (1992) Hydrophobicities and electrophoretic mobilities of anaerobic bacterial isolates from methanogenic granular sludge. Appl. Environ. Microbiol. 58:1054–1056

Gujer W, Zehnder AJB (1983) Conversion processes in anaerobic digestion. In: Water Science and Technology. IWA Publishing, pp 127–167

Guo J, Peng Y, Wang S, Yang X, Wang Z, Zhu A (2012a) Stable limited filamentous bulking through keeping the competition between floc-formers and filaments in balance. Bioresour Technol 103:7–15. doi: 10.1016/j.biortech.2011.08.114

Guo J, Peng Y, Wang Z, Yuan Z, Yang X, Wang S (2012b) Control filamentous bulking caused by chlorine-resistant Type 021N bacteria through adding a biocide CTAB. Water Res 46:6531–6542. doi: 10.1016/j.watres.2012.09.037

Gupta RS, Chander P, George S (2013) Phylogenetic framework and molecular signatures for the class Chloroflexi and its different clades; Proposal for division of the class Chloroflexi class. nov. into the suborder Chloroflexineae subord. nov., consisting of the emended family Oscillochlorida. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 103:99–119. doi: 10.1007/s10482-012-9790-3

Hanada S (2014) The phylum Chloroflexi, the family chloroflexaceae, and the related phototrophic families oscillochloridaceae and roseiflexaceae. In: The Prokaryotes: Other Major Lineages of Bacteria and The Archaea. Springer-Verlag Berlin Heidelberg, pp 515–532

Hattori S (2008) Syntrophic acetate-oxidizing microbes in methanogenic environments. Microbes Environ. 23:118–127

Hawley AK, Nobu MK, Wright JJ, Durno WE, Morgan-Lang C, Sage B, Schwientek P, Swan BK, Rinke C, Torres-Beltrán M, Mewis K, Liu WT, Stepanauskas R, Woyke T, Hallam SJ (2017) Diverse Marinimicrobia bacteria may mediate coupled biogeochemical cycles along eco-thermodynamic gradients. Nat Commun 8:1–10. doi: 10.1038/s41467-017-01376-9

Heather JM, Chain B (2016) The sequence of sequencers: The history of sequencing DNA. Genomics 107:1–8 . doi: 10.1016/j.ygeno.2015.11.003

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. Genome Res 6:986–994 . doi: 10.1101/gr.6.10.986

Hong S, Bunge J, Leslin C, Jeon S, Epstein SS (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. ISME J 3:1365–1373. doi: 10.1038/ismej.2009.89

Hugenholtz P, Pace NR (1996) Identifying microbial diversity in the natural environment: A molecular

phylogenetic approach. Trends Biotechnol. 14:190–197

Hugenholtz P, Stackebrandt E (2004) Reclassification of Sphaerobacter thermophilus from the subclass Sphaerobacteridae in the phylum Actinobacteria to the class Thermomicrobia (emended description) in the phylum Chloroflexi (emended description). Int J Syst Evol Microbiol 54:2049–2051. doi: 10.1099/ijs.0.03028-0

Hulshoff Pol LW, De Castro Lopes SI, Lettinga G, Lens PNL (2004a) Anaerobic sludge granulation. Water Res 38:1376–1389. doi: 10.1016/j.watres.2003.12.002

Hulshoff Pol LW, De Castro Lopes SI, Lettinga G, Lens PNL (2004b) Anaerobic sludge granulation. Water Res 38:1376–1389. doi: 10.1016/j.watres.2003.12.002

Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol 12:1889–1898. doi: 10.1111/j.1462-2920.2010.02193.x

Izadi P, Izadi P, Eldyasti A (2020) Influence of microbial population dynamics on accelerated development of anammox micro-granular consortium. J Water Process Eng 36:101384 . doi: 10.1016/j.jwpe.2020.101384

Janda JM, Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. J. Clin. Microbiol. 45:2761–2764

Jenkins D, Richard M, Daigger G (2003a) Manual on the Causes and Control of Activated Sludge Bulking, Foaming, and Other Solids Seperation Problems, 3rd Edition

Jenkins D, Richard MG, Daigger GT (2003b) Manual on the Causes and Control of Activated Sludge Bulking, Foaming, and Other Solids Separation Problems

Jetten MSM (2008) The microbial nitrogen cycle. Environ Microbiol 10:2903–2909 . doi: 10.1111/j.1462-2920.2008.01786.x

Ji B, Yang K, Zhu L, Jiang Y, Wang H, Zhou J, Zhang H (2015) Aerobic denitrification: A review of important advances of the last 30 years. Biotechnol. Bioprocess Eng. 20:643–651

Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, Weinstock GM (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun 10:1–11 . doi: 10.1038/s41467-019-13036-1

Jünemann S, Kleinbölting N, Jaenicke S, Henke C, Hassa J, Nelkner J, Stolze Y, Albaum SP, Schlüter A, Goesmann A, Sczyrba A, Stoye J (2017) Bioinformatics for NGS-based metagenomics and the application to biogas research. J Biotechnol 261:10–23. doi: 10.1016/j.jbiotec.2017.08.012

Kaeberlein T, Lewis K, Epstein SS (2002) Isolating "uncultivabte" microorganisms in pure culture in a simulated natural environment. Science (80-) 296:1127–1129. doi: 10.1126/science.1070633

Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y (2013) Ardenticatena maritima gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum "Chloroflexi" isolated from an iron-rich coastal hydrothermal field, and description of Ardenticatenia classis nov. Int J Syst Evol Microbiol 63:2992–3002. doi: 10.1099/ijs.0.046532-0

Khan AA, Gaur RZ, Tyagi VK, Khursheed A, Lew B, Mehrotra I, Kazmi AA (2011) Sustainable options of post treatment of UASB effluent treating sewage: A review. Resour. Conserv. Recycl. 55:1232–1251

Kim J, Lim J, Lee C (2013) Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. Biotechnol Adv 31:1358–1373 . doi: 10.1016/j.biotechadv.2013.05.010

Kim M, Jeong SY, Yoon SJ, Cho SJ, Kim YH, Kim MJ, Ryu EY, Lee SJ (2008) Aerobic Denitrification of Pseudomonas putida AD-21 at Different C/N Ratios. J Biosci Bioeng 106:498–502 . doi: 10.1263/jbb.106.498

Kindaichi T, Ito T, Okabe S (2004) Ecophysiological Interaction between Nitrifying Bacteria and Heterotrophic Bacteria in Autotrophic Nitrifying Biofilms as Determined by Microautoradiography-Fluorescence *In situ* Hybridization. Appl Environ Microbiol 70:1641–1650 . doi: 10.1128/AEM.70.3.1641-1650.2004

Kindaichi T, Yuri S, Ozaki N, Ohashi A (2012) Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor. Water Sci Technol 66:2556–2561 . doi: 10.2166/wst.2012.479

Koch H, Galushko A, Albertsen M, Schintlmeister A, Gruber-Dorninger C, Lücker S, Pelletier E, Le Paslier D, Spieck E, Richter A, Nielsen PH, Wagner M, Daims H (2014) Microbial metabolism: Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. Science (80-) 345:1052–1054 . doi: 10.1126/science.1256985

Kragelund C, Kong Y, van der Waarde J, Thelen K, Eikelboom D, Tandoi V, Thomsen TR, Nielsen PH (2006) Ecophysiology of different filamentous Alphaproteobacteria in industrial wastewater treatment plants. Microbiology 152:3003–3012 . doi: 10.1099/mic.0.29249-0

Kragelund C, Levantesi C, Borger A, Thelen K, Eikelboom D, Tandoi V, Kong Y, Van Der Waarde J, Krooneman J, Rossetti S, Thomsen TR, Nielsen PH (2007) Identity, abundance and ecophysiology of filamentous Chloroflexi species present in activated sludge treatment plants. FEMS Microbiol Ecol 59:671–682 . doi: 10.1111/j.1574-6941.2006.00251.x

Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Clemente JC, Burkepile DE, Vega Thurber RL, Knight R, Beiko RG, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31:814–821 . doi: 10.1038/nbt.2676

Lawson CE, Wu S, Bhattacharjee AS, Hamilton JJ, McMahon KD, Goel R, Noguera DR (2017) Metabolic network analysis reveals microbial community interactions in anammox granules. Nat Commun 8:1–12 . doi: 10.1038/ncomms15416

Lettinga G, Hulshoff Pol LW (1991) USAB-process design for various types of wastewaters. In: Water Science and Technology. IWA Publishing, pp 87–107

Lettinga G, van Velsen AFM, Hobma SW, de Zeeuw W, Klapwijk A (1980) Use of the upflow sludge blanket (USB) reactor concept for biological wastewater treatment, especially for anaerobic treatment. Biotechnol Bioeng 22:699–734 . doi: 10.1002/bit.260220402

Li J, Hu B, Zheng P, Qaisar M, Mei L (2008) Filamentous granular sludge bulking in a laboratory scale UASB reactor. Bioresour Technol 99:3431–3438 . doi: 10.1016/j.biortech.2007.08.005

Liu J, Li J, Xie K, Sellamuthu B (2019) Role of adding dried sludge micropowder in aerobic granular sludge reactor with extended filamentous bacteria. Bioresour Technol Reports 5:51–58 . doi: 10.1016/j.biteb.2018.12.002

Liu WT, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63:4516–22 . doi: 0099-2240/97/$04.00?0

Liu Y, Liu QS (2006) Causes and control of filamentous growth in aerobic granular sludge sequencing batch

reactors. Biotechnol. Adv. 24:115–127

Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X, Bai Y (2020) A practical guide to amplicon and metagenomic analysis of microbiome data. Protein Cell. doi: 10.1007/s13238-020-00724-8

Locey, K. J., & Lennon, J. T. (2016). Scaling laws predict global microbial diversity. Proceedings of the National Academy of Sciences, 113(21), 5970-5975.

Logares R, Sunagawa S, Salazar G, Cornejo-Castillo FM, Ferrera I, Sarmento H, Hingamp P, Ogata H, de Vargas C, Lima-Mendez G, Raes J, Poulain J, Jaillon O, Wincker P, Kandels-Lewis S, Karsenti E, Bork P, Acinas SG (2014) Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. Environ Microbiol 16:2659–2671 . doi: 10.1111/1462-2920.12250

López IM (2016) The potential of biogas production in Uruguay. Renew Sustain Energy Rev 54:1580–1591 . doi: 10.1016/j.rser.2015.10.099

Luo G, Li J, Li Y, Wang Z, Li WT, Li AM (2016) Performance, kinetics behaviors and microbial community of internal circulation anaerobic reactor treating wastewater with high organic loading rate: Role of external hydraulic circulation. Bioresour Technol 222:470–477 . doi: 10.1016/j.biortech.2016.10.023

Ma B, Peng Y, Zhang S, Wang J, Gan Y, Chang J, Wang S, Wang S, Zhu G (2013) Performance of anammox UASB reactor treating low strength wastewater under moderate and low temperatures. Bioresour Technol 129:606–611 . doi: 10.1016/j.biortech.2012.11.025

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes X V., Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380 . doi: 10.1038/nature03959

Martins AMP, Pagilla K, Heijnen JJ, Van Loosdrecht MCM (2004) Filamentous bulking sludge - A critical review. Water Res 38:793–817 . doi: 10.1016/j.watres.2003.11.005

Matsuura N, Tourlousse DM, Ohashi A, Sekiguchi Y (2015) Draft Genome Sequences of Anaerolinea thermolimosa IMO-1, Bellilinea caldifistulae GOMI-1, Leptolinea tardivitalis YMTK-2, Levilinea saccharolytica KIBI-1, Longilinea arvoryzae KOME-1, Previously Described as Members of the Class Anaerolineae (Chloroflexi. Genome Announc 3:3–4 . doi: 10.1128/genomeA.00975-15.Copyright

Mavromatis K, Ivanova N, Barry K, Shapiro H, Goltsman E, McHardy AC, Rigoutsos I, Salamov A, Korzeniewski F, Land M, Lapidus A, Grigoriev I, Richardson P, Hugenholtz P, Kyrpides NC (2007) Use of simulated data sets to evaluate the fidelity of metagenomic processing methods. Nat Methods 4:495–500 . doi: 10.1038/nmeth1043

McCaig AE, Grayston SJ, Prosser JI, Glover LA (2001) Impact of cultivation on characterisation of species composition of soil bacterial communities. FEMS Microbiol Ecol 35:37–48 . doi: 10.1111/j.1574-6941.2001.tb00786.x

McCarty PL, Bae J, Kim J (2011) Domestic wastewater treatment as a net energy producer-can this be

achieved? Environ Sci Technol 45:7100–7106 . doi: 10.1021/es2014264

McIlroy SJ, Karst SM, Nierychlo M, Dueholm MS, Albertsen M, Kirkegaard RH, Seviour RJ, Nielsen PH (2016) Genomic and *in situ* investigations of the novel uncultured Chloroflexi associated with 0092 morphotype filamentous bulking in activated sludge. Isme J (In press):1–12 . doi: 10.1038/ismej.2016.14

McIlroy SJ, Kirkegaard RH, Dueholm MS, Fernando E, Karst SM, Albertsen M, Nielsen PH (2017) Culture-independent analyses reveal novel anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. Front Microbiol 8: . doi: 10.3389/fmicb.2017.01134

Mende DR, Waller AS, Sunagawa S, Järvelin AI, Chan MM, Arumugam M, Raes J, Bork P (2012) Assessment of Metagenomic Assembly Using Simulated Next Generation Sequencing Data. PLoS One 7:e31386 . doi: 10.1371/journal.pone.0031386

Merlin G, Hervé B (2013) Anaerobic Digestion of Agricultural Waste: State of the Art and Future Trends. In: Anaerob. Dig. Types, Process. Environ. Impact

Miura Y, Miura Y, Watanabe Y, Watanabe Y, Okabe S, Okabe S (2007) Significance of Chloroflexi in Performance of Submerged Membrane Bioreactors (MBR) Treating Municipal Wastewater. Environ Sci Technol 41:7787–7794 . doi: 10.1021/es071263x

Moter A, Gobel UB (2000) Fluorescence *in situ* hybridization (FISH) for direct visualization of microorganisms. 41:85–112

Moura LL, Duarte KLS, Santiago EP, Mahler CF, Bassin JP (2018) Strategies to re-establish stable granulation after filamentous outgrowth: Insights from lab-scale experiments. Process Saf Environ Prot 117:606–615 . doi: 10.1016/j.psep.2018.06.005

Münch E V., Lant P, Keller J (1996) Simultaneous nitrification and denitrification in bench-scale sequencing batch reactors. Water Res 30:277–284 . doi: 10.1016/0043-1354(95)00174-3

Murray RGE, Stackebrandt E (1995) Taxonomic note: Implementation of the provisional status *Candidatus* for incompletely described procaryotes. Int J Syst Bacteriol 45:186–187 . doi: 10.1099/00207713-45-1-186

Nealson KH (1997) SEDIMENT BACTERIA: Who’s There, What Are They Doing, and What’s New? Annu Rev Earth Planet Sci 25:403–434 . doi: 10.1146/annurev.earth.25.1.403

Nielsen PH, Kragelund C, Seviour RJ, Nielsen JL (2009) Identity and ecophysiology of filamentous bacteria in activated sludge. FEMS Microbiol. Rev. 33:969–998

Nielsen PH, Mielczarek AT, Kragelund C, Nielsen JL, Saunders AM, Kong Y, Hansen AA, Vollertsen J (2010) A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants. Water Res 44:5070–5088 . doi: 10.1016/j.watres.2010.07.036

Nierychlo M, McIlroy SJ, Kucheryavskiy S, Jiang C, Ziegler AS, Kondrotaite Z, Stokholm-Bjerregaard M, Nielsen PH (2020) *Candidatus* Amarolinea and *Candidatus* Microthrix Are Mainly Responsible for Filamentous Bulking in Danish Municipal Wastewater Treatment Plants. Front Microbiol 11:1–17 . doi: 10.3389/fmicb.2020.01214

Nierychlo M, Miłobędzka A, Petriglieri F, McIlroy B, Nielsen PH, McIlroy SJ (2019a) The morphology and metabolic potential of the Chloroflexi in full-scale activated sludge wastewater treatment plants. FEMS Microbiol Ecol 95:1–11 . doi: 10.1093/femsec/fiy228

Nierychlo M, Miłobędzka A, Petriglieri F, McIlroy B, Nielsen PH, McIlroy SJ (2019b) The morphology and metabolic potential of the Chloroflexi in full-scale activated sludge wastewater treatment plants. FEMS

Microbiol Ecol 95: . doi: 10.1093/femsec/fiy228

Nierychlo M, Miłobędzka A, Petriglieri F, McIlroy B, Nielsen PH, McIlroy SJ (2018) The ecology of the Chloroflexi in full-scale activated sludge wastewater treatment plants. bioRxiv 335752 . doi: 10.1101/335752

Nittami T, Speirs LBM, Yamada T, Suzuki I, Fukuda J, Kurisu F, Seviour RJ (2017) Quantification of Chloroflexi Eikelboom morphotype 1851 for prediction and control of bulking events in municipal activated sludge plants in Japan. Appl Microbiol Biotechnol 101:3861–3869 . doi: 10.1007/s00253-016-8077-4

Nunoura T, Hirai M, Miyazaki M, Kazama H, Makita H, Hirayama H, Furushima Y, Yamamoto H, Imachi H, Takai K (2013) Isolation and characterization of a thermophilic, obligately anaerobic and heterotrophic marine Chloroflexi bacterium from a Chloroflexi-dominated microbial community associated with a Japanese shallow hydrothermal system, and proposal for Thermomarinilin. Microbes Environ 28:228–35 . doi: 10.1264/jsme2.ME12193

Örmeci B, Vesilind PA (2000) Development of an improved synthetic sludge: A possible surrogate for studying activated sludge dewatering characteristics. Water Res 34:1069–1078 . doi: 10.1016/S0043-1354(99)00251-1

Pace NR, Stahl DA, Lane DJ, Olsen GJ (1986) The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences. Springer, Boston, MA, pp 1–55

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996 . doi: 10.1038/nbt.4229

Perna V, Castelló E, Wenzel J, Zampol C, Fontes Lima DM, Borzacconi L, Varesche MB, Zaiat M, Etchebehere C (2013) Hydrogen production in an upflow anaerobic packed bed reactor used to treat cheese whey. Int J Hydrogen Energy 38:54–62 . doi: 10.1016/j.ijhydene.2012.10.022

Petriglieri F, Nierychlo M, Nielsen PH, McIlroy SJ (2018) *In situ* visualisation of the abundant Chloroflexi populations in full-scale anaerobic digesters and the fate of immigrating species. PLoS One 13:1–14 . doi: 10.1371/journal.pone.0206255

Podosokorskaya OA, Novikov AA, Kolganova T V, Kublanov I V (2016) Ornatilinea apprima gen. nov., sp. nov., a cellulolytic representative of the class Anaerolineae. 86–92 . doi: 10.1099/ijs.0.041012-0

Pol LWH, Heijnekamp K, Lettinga G (1988) The selection pressure as a driving force behind the granulation of anaerobic sludge. 153–161

Qiao JT, Qiu YL, Yuan XZ, Shi XS, Xu XH, Guo RB (2013) Molecular characterization of bacterial and archaeal communities in a full-scale anaerobic reactor treating corn straw. Bioresour Technol 143:512–518 . doi: 10.1016/j.biortech.2013.06.014

Ririe KM, Rasmussen RP, Wittwer CT (1997) Product differentiation by analysis of DNA melting curves during the polymerase chain reaction. Anal Biochem 245:154–160 . doi: 10.1006/abio.1996.9916

Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, Li T, Camacho P, Sghir A (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME J 3:700–714 . doi: 10.1038/ismej.2009.2

Rossetti S, Tomei MC, Nielsen PH, Tandoi V (2005) “Microthrix parvicella ”, a filamentous bacterium causing

bulking and foaming in activated sludge systems: a review of current knowledge. FEMS Microbiol Rev 29:49–64 . doi: 10.1016/j.femsre.2004.09.005

Saby S, Djafer M, Chen GH (2003) Effect of low ORP in anoxic sludge zone on excess sludge production in oxic-settling-anoxic activated sludge process. Water Res 37:11–20 . doi: 10.1016/S0043-1354(02)00253-1

Sancho I, Lopez-Palau S, Arespacochaga N, Cortina JL (2019) New concepts on carbon redirection in wastewater treatment plants: A review. Sci Total Environ 647:1373–1384 . doi: 10.1016/j.scitotenv.2018.08.070

Satoh H, Miura Y, Tsushima I, Okabe S (2007) Layered structure of bacterial and archaeal communities and their *in situ* activities in anaerobic granules. Appl Environ Microbiol 73:7300–7307 . doi: 10.1128/AEM.01426-07

Saunders AM, Albertsen M, Vollertsen J, Nielsen PH (2016) The activated sludge ecosystem contains a core community of abundant organisms. ISME J 10:11–20 . doi: 10.1038/ismej.2015.117

Schade M, Beimfohr C, Lemmer H (2002) Phylogenetic and physiological characterization of a “Nostocoida limicola”-like organism isolated from activated sludge. In: Water Science and Technology. IWA Publishing, pp 91–97

Schink B (1997) Energetics of syntrophic cooperation in methanogenic degradation. Microbiol Mol Biol Rev 61:262–280 . doi: 1092-2172/97/$04.0010

Schloss PD, Handelsman J (2008) A statistical toolbox for metagenomics: Assessing functional diversity in microbial communities. BMC Bioinformatics 9:34 . doi: 10.1186/1471-2105-9-34

Schmidt I, Sliekers O, Schmid M, Bock E, Fuerst J, Kuenen JG, Jetten MSM, Strous M (2003) New concepts of microbial treatment processes for the nitrogen removal in wastewater. FEMS Microbiol Rev 27:481–492 . doi: 10.1016/S0168-6445(03)00039-1

Schmidt JE, Ahring BK (1996) Granular sludge formation in upflow anaerobic sludge blanket (UASB) reactors. Biotechnol Bioeng 49:229–246 . doi: 10.1002/(SICI)1097-0290(19960205)49:3<229::AID-BIT1>3.0.CO;2-M

Schnürer A, Nordberg Å (2008) Ammonia, a selective agent for methane production by syntrophic acetate oxidation at mesophilic temperature. Water Sci Technol 57:735–740 . doi: 10.2166/wst.2008.097

Sekiguchi Y, Kamagata Y, Nakamura K, Ohashi A, Harada H (1999) Fluorescence *in situ* hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Appl Environ Microbiol 65:1280–1288

Sekiguchi Y, Ohashi A, Parks DH, Yamauchi T, Tyson GW, Hugenholtz P (2015) First genomic insights into members of a candidate bacterial phylum responsible for wastewater bulking. PeerJ 3:e740 . doi: 10.7717/peerj.740

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001a) *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl Environ Microbiol 67:5740–5749 . doi: 10.1128/AEM.67.12.5740-5749.2001

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001b) *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I.

Appl Environ Microbiol 67:5740–5749 . doi: 10.1128/AEM.67.12.5740-5749.2001

Shakya M, Quince C, Campbell JH, Yang ZK, Schadt CW, Podar M (2013) Comparative metagenomic and rRNA microbial diversity characterization using archaeal and bacterial synthetic communities. Environ Microbiol 15:1882–1899 . doi: 10.1111/1462-2920.12086

Sharpton TJ (2014) An introduction to the analysis of shotgun metagenomic data. Front Plant Sci 5:1–14 . doi: 10.3389/fpls.2014.00209

Sharpton TJ, Riesenfeld SJ, Kembel SW, Ladau J, O'Dwyer JP, Green JL, Eisen JA, Pollard KS (2011) PhylOTU: A High-Throughput Procedure Quantifies Microbial Community Diversity and Resolves Novel Taxa from Metagenomic Data. PLoS Comput Biol 7:e1001061 . doi: 10.1371/journal.pcbi.1001061

Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere." Proc Natl Acad Sci U S A 103:12115–12120 . doi: 10.1073/pnas.0605127103

Song YX, Liao Q, Yu C, Xiao R, Tang CJ, Chai LY, Duan CS (2017) Physicochemical and microbial properties of settled and floating anammox granules in upflow reactor. Biochem Eng J 123:75–85 . doi: 10.1016/j.bej.2017.04.002

Sonune A, Ghate R (2004) Developments in wastewater treatment methods. Desalination 167:55–63 . doi: 10.1016/j.desal.2004.06.113

Speirs L, Nittami T, McIlroy S, Schroeder S, Seviour RJ (2009) Filamentous bacterium Eikelboom Type 0092 in activated sludge plants in Australia is a member of the phylum Chloroflexi. Appl Environ Microbiol 75:2446–2452 . doi: 10.1128/AEM.02310-08

Speirs LBM, Mcilroy SJ, Petrovski S, Seviour RJ (2011) The activated sludge bulking filament Eikelboom morphotype 0914 is a member of the Chloroflexi. Environ Microbiol Rep 3:159–165 . doi: 10.1111/j.1758-2229.2010.00201.x

Speth DR, In'T Zandt MH, Guerrero-Cruz S, Dutilh BE, Jetten MSM (2016) Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. Nat Commun 7: . doi: 10.1038/ncomms11172

Strous M, Heijnen JJ, Kuenen JG, Jetten MSM (1998) The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. Appl Microbiol Biotechnol 50:589–596 . doi: 10.1007/s002530051340

Strous M, Pelletier E, Mangenot S, Rattei T, Lehner A, Taylor MW, Horn M, Daims H, Bartol-Mavel D, Wincker P, Barbe V, Fonknechten N, Vallenet D, Segurens B, Schenowitz-Truong C, Médigue C, Collingro A, Snel B, Dutilh BE, Op Den Camp HJM, Van Der Drift C, Cirpus I, Van De Pas-Schoonen KT, Harhangi HR, Van Niftrik L, Schmid M, Keltjens J, Van De Vossenberg J, Kartal B, Meier H, Frishman D, Huynen MA, Mewes HW, Weissenbach J, Jetten MSM, Wagner M, Le Paslier D (2006) Deciphering the evolution and metabolism of an anammox bacterium from a community genome. Nature 440:790–794 . doi: 10.1038/nature04647

Su C, Lei L, Duan Y, Zhang KQ, Yang J (2012) Culture-independent methods for studying environmental microorganisms: Methods, application, and perspective. Appl Microbiol Biotechnol 93:993–1003 . doi: 10.1007/s00253-011-3800-7

Su C, Liu Y, Yang X, Li H (2020) Effect of Hydraulic Retention Time on Biohydrogen Production from

Glucose in an Internal Circulation Reactor. Energy and Fuels. doi: 10.1021/acs.energyfuels.9b03316

Sun L, Toyonaga M, Ohashi A, Matsuura N, Tourlousse DM, Meng X, Tamaki H, Hanada S, Cruz R, Yamaguchi T, Sekiguchi Y (2016) Isolation and characterization of Flexilinea flocculi gen . nov ., sp . nov ., a filamentous , anaerobic bacterium belonging to the class Anaerolineae in the phylum Chloroflexi. Int J Syst Evol Microbiol 66:988–996 . doi: 10.1099/ijsem.0.000822

Sundberg C, Al-Soud WA, Larsson M, Alm E, Yekta SS, Svensson BH, Sørensen SJ, Karlsson A (2013) 454 Pyrosequencing Analyses of Bacterial and Archaeal Richness in 21 Full-Scale Biogas Digesters. FEMS Microbiol Ecol 85:612–626 . doi: 10.1111/1574-6941.12148

Tang CJ, Duan CS, Yu C, Song YX, Chai LY, Xiao R, Wei Z, Min XB (2017) Removal of nitrogen from wastewaters by anaerobic ammonium oxidation (ANAMMOX) using granules in upflow reactors. Environ. Chem. Lett. 15:311–328

Tang CJ, Zheng P, Wang CH, Mahmood Q, Zhang JQ, Chen XG, Zhang L, Chen JW (2011) Performance of high-loaded ANAMMOX UASB reactors containing granular sludge. Water Res 45:135–144 . doi: 10.1016/j.watres.2010.08.018

Torsvik V, Sørheim R, Goksøyr J (1996) Total bacterial diversity in soil and sediment communities - A review. J. Ind. Microbiol. Biotechnol. 17:170–178

Van Lier JB (2008) High-rate anaerobic wastewater treatment: Diversifying from end-of-the-pipe treatment to resource-oriented conversion techniques. Water Sci Technol 57:1137–1148 . doi: 10.2166/wst.2008.040

Vavourakis CD, Andrei AS, Mehrshad M, Ghai R, Sorokin DY, Muyzer G (2018) A metagenomics roadmap to the uncultured genome diversity in hypersaline soda lake sediments 06 Biological Sciences 0605 Microbiology 06 Biological Sciences 0604 Genetics. Microbiome 6:168 . doi: 10.1186/s40168-018-0548-7

Veiga MC, Jain MK, Wu WM, Hollingsworth RI, Zeikus JG (1997) Composition and role of extracellular polymers in methanogenic granules. Appl Environ Microbiol 63:403–407 . doi: 10.1128/aem.63.2.403-407.1997

Verstraete W, Morgan-Sagastume F, Aiyuk S, Waweru M, Rabaey K, Lissens G (2005) Anaerobic digestion as a core technology in sustainable management of organic matter. Water Sci Technol 52:59–66 . doi: 10.2166/wst.2005.0498

Von Sperling M, de Lemos Chernicharo C (2015) Biological Wastewater Treatment in Warm Climate Regions Volume I. Water Intell Online 4:9781780402734–9781780402734 . doi: 10.2166/9781780402734

Walker LM, Hoppe T, Silliker ME (2017) Molecular Techniques and Current Research Approaches. In: Myxomycetes: Biology, Systematics, Biogeography and Ecology. Elsevier Inc., pp 145–173

Wang S, Guo J, Lian J, Ngo HH, Guo W, Liu Y, Song Y (2016) Rapid start-up of the anammox process by denitrifying granular sludge and the mechanism of the anammox electron transport chain. Biochem Eng J 115:101–107 . doi: 10.1016/j.bej.2016.09.001

Ward DM, Bateson MM, Weller R, Ruff-Roberts AL (1992) Ribosomal RNA Analysis of Microorganisms as They Occur in Nature. Springer, Boston, MA, pp 219–286

Wei D, Xue X, Yan L, Sun M, Zhang G, Shi L, Du B (2014) Effect of influent ammonium concentration on the shift of full nitritation to partial nitrification in a sequencing batch reactor at ambient temperature. Chem Eng J 235:19–26 . doi: 10.1016/j.cej.2013.09.005

Weiland P (2010) Biogas production: Current state and perspectives. Appl. Microbiol. Biotechnol. 85:849–860

Werner JJ, Knights D, Garcia ML, Scalfone NB, Smith S, Yarasheski K, Cummings TA, Beers AR, Knight R, Angenent LT (2011) Bacterial community structures are unique and resilient in full-scale bioenergy systems. Proc Natl Acad Sci U S A 108:4158–63 . doi: 10.1073/pnas.1015676108

Wheeldon I, Caners C, Karan K, Peppley B (2007) Utilization of biogas generated from Ontario wastewater treatment plants in solid oxide fuel cell systems: A process modeling study. Int J Green Energy 4:221–231 . doi: 10.1080/15435070601015585

Wiegant WM, de Man AWA (1986) Granulation of biomass in thermophilic upflow anaerobic sludge blanket reactors treating acidified wastewaters. Biotechnol Bioeng 28:718–727 . doi: 10.1002/bit.260280511

Wittebolle L, Marzorati M, Clement L, Balloi A, Daffonchio D, Heylen K, De Vos P, Verstraete W, Boon N (2009) Initial community evenness favours functionality under selective stress. Nature 458:623–626 . doi: 10.1038/nature07840

Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques 22:130–138 . doi: 10.2144/97221bi01

Wojnowska-Baryła I, Cydzik-Kwiatkowska A, Zielińska M (2010) The application of molecular techniques to the study of wastewater treatment systems. Methods Mol. Biol. 599:157–183

Xia Y, Wang Y, Wang Y, Chin FYL, Zhang T (2016) Cellular adhesiveness and cellulolytic capacity in Anaerolineae revealed by omics-based genome interpretation. Biotechnol Biofuels 9:1–13 . doi: 10.1186/s13068-016-0524-z

Yamada T, Sekiguchi Y (2009) Cultivation of uncultured Chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi "subphylum i" with natural and biotechnological relevance. Microbes Environ 24:205–216 . doi: 10.1264/jsme2.ME09151S

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the . Int J Syst Evol Microbiol 56:1331–1340 . doi: 10.1099/ijs.0.64169-0

Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005a) Diversity, localization, and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl Environ Microbiol 71:7493–7503 . doi: 10.1128/AEM.71.11.7493-7503.2005

Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005b) Diversity, localization, and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl Environ Microbiol 71:7493–7503 . doi: 10.1128/AEM.71.11.7493-7503.2005

Yamada T, Yamauchi T, Shiraishi K, Hugenholtz P, Ohashi A, Harada H, Kamagata Y, Nakamura K, Sekiguchi Y (2007) Characterization of filamentous bacteria, belonging to candidate phylum KSB3, that are associated with bulking in methanogenic granular sludges. ISME J 1:246–255 . doi: 10.1038/ismej.2007.28

Yang Y, Sanford R, Yan J, Chen G, Li X, Loffler F (2020) Roles of Organohalide-Respiring Dehalococcoidia in Carbon Cycling. Appl Environ Sci 5:

Yu T, Qi R, Li D, Zhang Y, Yang M (2010a) Nitrifier characteristics in submerged membrane bioreactors under different sludge retention times. Water Res 44:2823–2830 . doi: 10.1016/j.watres.2010.02.021

Yu Z, Morrison M, Schanbacher FL (2010b) Production and Utilization of Methane Biogas as Renewable Fuel. In: Biomass to Biofuels. Blackwell Publishing Ltd., Oxford, UK, pp 403–433

Zhang J, Chiodini R, Badr A, Zhang G (2011) The impact of next-generation sequencing on genomics. J. Genet. Genomics 38:95–109

Zhang QL, Liu Y, Ai GM, Miao LL, Zheng HY, Liu ZP (2012) The characteristics of a novel heterotrophic nitrification-aerobic denitrification bacterium, Bacillus methylotrophicus strain L7. Bioresour Technol 108:35–44 . doi: 10.1016/j.biortech.2011.12.139

Zhang T, Fang HHP (2006) Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples. Appl. Microbiol. Biotechnol. 70:281–289

Zhao Y, Liu S, Jiang B, Feng Y, Zhu T, Tao H, Tang X, Liu S (2018) Genome-Centered Metagenomics Analysis Reveals the Symbiotic Organisms Possessing Ability to Cross-Feed with Anammox Bacteria in Anammox Consortia

Zumft WG (1997) Cell biology and molecular basis of denitrification. Microbiol Mol Biol Rev 61:533–616 . doi: 10.1128/.61.4.533-616.1997

CAPÍTULO 2

Diversidad y abundancia del filo Chloroflexi en sistemas de tratamiento metanogénicos de Uruguay

Trabajo publicado en Journal of Applied Microbiology - DOI: 10.1111/jam.14115

Title: Preliminary analysis of Chloroflexi populations in full scale UASB methanogenic reactors

Patricia Bovio[1] • Angela Cabezas[1,2] • Claudia Etchebehere[1]

[1]Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomic, Biological Research Institute "Clemente Estable", Avenida Italia 3318, CP: 11600, Montevideo, Uruguay.

[2] Department of Environmental Sciences, Uruguay Technological University (UTEC), Francisco Antonio Maciel s/n, CP: 97000, Durazno, Uruguay.

Abbreviated running headline:
Chloroflexi phylum in methanogenic reactors

* Correspondence to: Claudia Etchebehere
Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomics, Biological Research Institute "Clemente Estable", Av. Italia 3318. CP 11600. Montevideo – Uruguay.
Phone: (598) 24871616. Fax: (598) 24875461
cetchebe@gmail.com

Abstract

Aims: The phylum Chloroflexi is frequently found in high abundance in methanogenic reactors, but their role is still unclear as most of them remain uncultured and understudied. Hence, a detailed analysis was performed in samples from five Up-flow anaerobic sludge blanket (UASB) full-scale reactors fed with different industrial wastewaters. Methods and Results: Quantitative PCR shows that the phylum Chloroflexi was abundant in all UASB methanogenic reactors, with higher abundance in the reactors operated for a long period of time, which presented granular biomass. Both T-RFLP and 16S rRNA gene amplicon sequencing revealed diverse Chloroflexi populations apparently determined by the different inocula. According to the phylogenetic analysis the sequences from the dominant Chloroflexi were positioned in branches where no sequences of the cultured representative strains are placed. FISH analysis performed in two of the reactors showed filamentous morphology of the hybridizing cells. Conclusions: While members of the Anaerolineae class within phylum Chloroflexi were predominant, their diversity is still poorly described in anaerobic reactors. Due to their filamentous morphology, Chloroflexi may have a key role in the granulation in methanogenic UASB reactors. Significance and Impact of Study: Our results bring new insights about the diversity, stability, dynamics and abundance of this phylum in full scale UASB reactors which aid in understanding their function within the reactor biomass. However, new methodological approaches and analysis of bulking biomass is needed to completely unravel their role in these reactors. Combining all this knowledge with reactor operational parameters will allow to understand their participation in granulation and bulking episodes and design strategies to prevent Chloroflexi overgrowth.

key words: Anaerobic digestion, Chloroflexi, microbial community, methanogenic reactor, full scale, UASB.

2.1 Introduction

Anaerobic digestion is composed by a series of biological reactions performed by microorganisms which can be used as a form of biological treatment of organic waste and wastewater. At least four microbial trophic groups contribute to the organic waste degradation forming CH_4 and CO_2 as final products: hydrolytic bacteria, fermenting bacteria, syntrophic bacteria and methanogenic archaea (Amani et al. 2010). Anaerobic digestion has been successfully applied in various industries at full scale for several decades (Angenent et al. 2004; Verstraete et al. 2005). The anaerobic treatment of wastewater can produce energy in the form of methane, with a low biomass production, high efficiency and a low energy demand (Van Lier 2008; Angelidaki et al. 2011). In Uruguay, these systems have been applied for the treatment of several agro-industrial wastewaters from dairy, brewery and bioethanol production industries, reducing its environmental impact and achieving renewable energy generation (López 2016).

Up-flow anaerobic sludge blanket (UASB) technology has been used widely because of its ability to treat a broad range of organic waste streams at high loading rates (Lettinga 1995; Chong et al. 2012; Hinken et al. 2014). Sludge granulation is a necessary phenomenon in UASB reactors as it provides appropriate settling ability, high biomass retention time and increases the resistance to high organic loading rates and toxic shocks (Speece 1996). The granules are composed by the previously mentioned trophic groups (hydrolytic bacteria, fermenting bacteria, syntrophic bacteria and methanogenic archaea), inorganic minerals and extracellular polymers (Lettinga 1995; Sekiguchi and Kamagata 2004). Sludge granulation favors the syntrophic process between the fermentative bacteria and methanogenic archaea facilitating interspecies hydrogen transfer (Ariesyady et al. 2007).

Chloroflexi has been reported as one of the most predominant phyla present in anaerobic wastewater treatment systems and has even been defined as part of the core group together with Proteobacteria, Bacteroidetes and Firmicutes (Chouari et al. 2005; Narihiro et al. 2009; Rivière et al. 2009; Shu et al. 2015). Within the Chloroflexi phylum, members of the Anaerolineae class are predominant but their role in these systems has not been completely addressed (Yamada and Sekiguchi 2009). This is mainly due to the difficulties to cultivate

and isolate members from this class as indicated by the efforts performed by several research groups to retrieve strains within this class (see Table 2.1 for references). The cultivated members are mainly slow growing fermenting bacteria with filamentous morphology (Table 2.1).

Table 2.1 Characteristics of isolated strains from species belonging to the classes Anaerolineae, Caldilineae and Dehalococcoidia within the phylum Chloroflexi.

	Species	Isolation source	Physiology	Carbon and energy source	Reference
Class *Anaerolineae*	*Anaerolinea thermophila* UNI-1	UASB sludge §	Strictly anaerobic †	Carbohydrates, proteins *	(Sekiguchi et al., 2003)
	Anaerolinea thermolimosa IMO-1	UASB sludge §	Strictly anaerobic †	Carbohydrates, proteins *	(Yamada et al., 2006)
	Levilinea saccharolytica KIBI-1	UASB sludge ‡	Strictly anaerobic †	Carbohydrates, proteins *	(Yamada et al., 2006)
	Leptolinea tardivitalis YMTK-2	UASB sludge ‡	Strictly anaerobic †	Carbohydrates, proteins *	(Yamada et al., 2006)
	Bellilinea caldifistulae GOMI-1	Anaerobic sludge §	Strictly anaerobic †	Carbohydrates, proteins *	(Yamada et al., 2007)
	Longilinea arvoryzae KOME-1	Rice paddy soil ‡	Strictly anaerobic †	Carbohydrates, proteins *	(Yamada et al., 2007)
	Ornatilinea apprima P3M-1	Deep terrestrial hot aquifer ‡	Strictly anaerobic †	Carbohydrates; proteins	(Podosokorskaya et al., 2016)
	Thermomarinilinea lacunofontalis SW7	Hydrothermal vent §	Strictly anaerobic †	Proteins	(Nunoura et al., 2013)
	Thermanaerothrix daxensis GNS-1	Deep hot aquifer §	Strictly anaerobic †	Carbohydrates *	(Grégoire et al., 2011b)
	Pelolinea submarina MO-CFX1	Marine sediment	Strictly anaerobic †	Carbohydrates *	(Imachi et al., 2014)
	Flexilinea floccule TC1T	UASB sludge ‡	Strictly anaerobic †	Carbohydrates *	(Sun et al., 2016)
Class *Caldilineae*	*Caldilinea aerophila* STL-6-O1	Hot spring §	Facultatively aerobic †	Carbohydrates, amino acids *	(Sekiguchi et al., 2003)
	Caldilinea tarbellica D1-25-10-4	Deep hot aquifer §	Strictly anaerobic †	Carbohydrates *	(Grégoire et al., 2011a)
	Litorilinea aerophila PRI-4131	Hot spring §	Aerobic †	Carbohydrates	(Kale et al., 2013)
Class Dehalococcoidia	*Dehalogenimonas alkenigignens* IP3-3	Groundwater ‡	Strictly anaerobic. Irregular cocci	Chlorinated compounds	(Bowman et al, 2013)

* Yeast extract was required for growth

† Multicellular filamentous morphology

‡ Mesophilic

§ Thermophilic

Recently, the preliminary annotation of the CAMBI-1 genome (*Candidatus* Brevefilum fermentans) belonging to the Anaerolineae class, obtained from an anaerobic digester by metagenomics, was performed. The results indicate that this bacteria might be anaerobic chemoorganoheterotrophs with a fermentative metabolism. In addition, their high abundance in the reactor, make them likely important primary fermenters in digester systems (McIlroy et al. 2017).

Apart from their potential role as primary fermenters, it has been postulated that they are involved in granule formation in UASB reactors due to their filamentous morphology (Sekiguchi et al. 2001). But, in a previous study on the performance of an UASB reactor installed in our country, the overgrowth of filamentous Chloroflexi was related to an unusual anaerobic bulking process that caused sedimentation problems (Borzacconi et al. 2008). This bulking process seriously affects the performance of UASB reactors as it produces sudden

washout of the granular sludge biomass. Consequently, the reactor fails in performing effluent treatment at high loading rates (Sekiguchi et al. 2001; Yamada et al. 2007b, 2011; Borzacconi et al. 2008; Li et al. 2008). With the installation of full scale UASB reactors, the prevention of this problem has become important in order to obtain optimal reactor performance and avoid environmental contamination. The conditions favoring the overgrowth of these bacteria remains unknown. Increasing the knowledge about this group is essential to control their growth, sustain granule formation and prevent bulking episodes.
In order to contribute to the understanding of the role of Chloroflexi in UASB reactors we studied their taxonomic composition, diversity, abundance, dynamics and morphology in five full scale UASB reactors treating three different industrial wastewaters. Specific terminal Restriction Fragment Length Polymorphism (T-RFLP) and quantitative PCR (q-PCR) techniques were designed to analyze the diversity and abundance of the phylum Chloroflexi. The Chloroflexi taxonomic composition was studied by 16S rRNA gene amplicon sequencing. The cell morphology was observed by Fluorescent in situ Hybridization (FISH) with specific probes targeting this phylum.

2.2 Materials and methods

2.2.1 UASB methanogenic reactors and sample collection

Anaerobic sludge samples were taken from five UASB methanogenic reactors treating effluents from four industries located in Uruguay (Table 2.2).

Table 2.2 Operational parameters of the five UASB reactors.

UASB reactor	Wastewater	Stage/Sampling period (days)	Volume (m^3)	Organic loading rate ($kgCOD/m^3d$) ‡	HRT* (days)	Temperature (°C)	COD† removal (%)	CH_4 (%)	References
AL	Vinasse	Start-up/ 0- 456	100	3.0-5.4	8.60	Oct-25	60-78 ‡	68 §	(López et al., 2017)
SR	Dairy	Start-up/ 0- 334	100	1.5-6.0	0.83	Jun-26	55-90 ‡	75 §	(Fernández, 2016)
COA	Dairy	Stable operation /0-424	40	1.5-4.0	0.71	21-31	29-97 ‡	78 §	(Passeggi et al., 2009)
COB	Dairy	Stable operation /0-424	40	1.5-4.0	0.71	20-31	19-97 ‡	78 §	(Passeggi et al., 2009)
MO	Malting	Stable operation /0-542	250	4.0-5.0	0.67	13-22	35-80 ‡	77 §	(Borzacconi et al., 2006, 2008)

* HRT: Hydraulic retention time

† COD: Chemical oxygen demand

‡ Values expressed as range

§ Values expressed as average

All reactors were operated under mesophilic conditions. Reactors COA and COB, which belong to the same wastewater treatment plant, were previously studied by Passeggi et al. (2009). Both reactors were connected in parallel, receiving effluent with the same characteristics, and were installed underground, therefore, the seasonal temperature fluctuations were lower. Reactors AL and SR were previously studied by López et al. (2017) and Fernández (2016), respectively. Reactor MO was previously studied by Borzacconi et al. (2006, 2008). Reactor AL and SR were sampled during the start-up period. Reactors COA and COB were sampled during stable operation after 5 years of operation and reactor MO was sampled during stable operation after more than 10 years of operation. In reactor MO our group detected sedimentation problems due to overgrowth of Chloroflexi filaments (Borzacconi et al. 2008). The inoculum for reactors AL, COA, COB and SR were taken from the same anaerobic lagoon treating slaughterhouse wastewater but at different time points. A sample from this inoculum was taken at the moment of reactor SR inoculation and was named SR(i). Reactor MO was inoculated with sludge taken from another anaerobic lagoon also used for the treatment of slaughterhouse wastewater. Samples from all reactors were taken at different heights and then homogenized. In total, 60 samples (1 l each approximately) were obtained from the five UASB reactors. Sub samples (50 ml) were stored at -20 °C until DNA extraction. As the storage method might impact the microbial community we stored all samples under the same conditions which is highly important in order to avoid biases as shown previously for soil microbial communities (Rubin et al 2013, Lauber et al 2010). For reactors AL and SR, day 0 means the day of inoculation. For reactor MO, COA and COB, day 0 indicates the first sampling day. Information about the reactors operation condition and performance at the day of sampling were included in Table S2.1.

2.2.2 DNA extraction

Sludge samples were thawed and centrifuged (5 min, 10,000 g). Approximately 0.35 g of wet pellet was used for DNA extraction with the ZR Soil Microbe DNA MiniPrepTM kit (Zymo Research) according to the manufacturer's instructions. The quality of the extracted genomic DNA was determinate by 1% agarose gel electrophoresis (Nucleic Acid Stain, GoodView^{TM}, Beijing) and stored at -20 °C until further use.

2.2.3 Terminal restriction fragment length polymorphism (T-RFLP) analysis

A specific approach was optimized in our laboratory to determine the structure and dynamics of Chloroflexi by 16S rRNA gene T-RFLP. For this, the PCR conditions and the enzymatic restriction were optimized using an environmental Chloroflexi clone (Clone 131, AJ412677). The protocol was optimized to obtain only one peak or T-RF for the clone. Samples were amplified using PCR primer set 27 forward (Edwards et al. 1989) labeled at the 5'-end with the 6-carboxyfluorescein dye and the specific primer for Chloroflexi, designed based on the FISH probe CFX1223R (Björnsson et al. 2002) (Table 2.3).

Table 2.3 Primers and probes used in this study for q-PCR, T-RFLP, 454-pyrosequencing, Ion Torrent sequencing and FISH.

Application	Probe/ Primer set	Specificity	Sequence (5'→3')	References
q-PCR, FISH	GNSB941	phylum *Chloroflexi*	AAACCACACGCTCCGCT	Gich et al. (2001)
	CFX1223R	phylum *Chloroflexi*	CCATTGTAGGGTGTGTGTMG	Björnsson et al. (2002)
T-RFLP	27F-6FAM	universal primer	AGAGTTTGATCMTGGCTCAG	Park et al. (2006)
	CFX1223R	phylum *Chloroflexi*	CCATTGTAGGGTGTGTGTMG	Björnsson et al. (2002)
Pyrosequencing	515F	universal primer	GTGCCAGCMGCCGCGGTAA	Caporaso et al., (2012)
	806R	universal primer	GGACTACVSGGGTATCTAAT	
Ion Torrent	520F	universal primer	AYTGGGYDTAAAGNG	Claesson et al., (2009)
	802R	universal primer	TACNNGGGTATCTAATCC	

The PCR protocol was performed as previously described (Gich et al. 2002) using the PCR cycle as follows: 94 °C for 5 min, followed by 30 cycles at 94 °C for 30 s, 56 °C for 30 s, 72 °C for 45 s and a final extension at 72 °C for 5 min. Each PCR mixture (50 µl) contained 1 µl of template DNA, 0.4 µM each primer, 0.2 mM each deoxynucleoside triphosphate, 5 µl of 10x PCR buffer, 1.5 mM $MgCl_2$, and 1.25 U of Taq Polymerase (Fermentas, USA). Purification of the PCR product was performed using the DNA Clean kit & ConcentratorTM-25 (Zymo Research, Orange, CA) according to the manufacturer's instructions. Quantification was performed using Nanodrop 1000 (Thermo Scientific, Wilmington, DE, USA). The restriction enzyme digestion was done according to Pycke et al. (2011), using the enzyme MspI (Fermentas, USA) with the following modifications of the protocol: the concentration of PCR product was optimized at 50 ng μl^{-1} of DNA per reaction, and an incubation time of 16 h at 37 °C, with subsequent inactivation at 80 °C for 20 min was used. The selection of restriction enzyme was performed from in silico digestion of 16S rRNA gene sequences

belonging to the phylum Chloroflexi cultivated members obtained from the NCBI database. For this, the in silico terminal restriction fragment sizes were calculated for each sequence using the restriction site of different enzymes (MspI, Sau3A, EcoRI, TaqI, Kpnl, Sacl, Smal, BamHI, HindIII and Pstl). The MspI was the restriction enzyme where the highest level of discrimination between species belonging to Anaerolinea and Caldilinea genera was obtained.

Separation and detection of fragments was performed using a sequencer 3130 (Applied Biosystems) with a GeneScan marker -1200 internal Liz (Applied Biosystems) at the Pasteur Institute, Uruguay. The chromatograms were analyzed using Peak Scanner v.1.0 (Applied Biosystems, Halle, Belgium). The peak relative abundances were determined by calculating the ratio between the height of each peak and the total peak height of all peaks within one sample. Only peaks with more than 3% abundance/intensity were considered. Nonmetric multidimensional scaling (NMDS) analyses using the Bray–Curtis similarity index (considers presence/absence as well as relative abundance of peaks) were performed to plot the similarity rank order of T- RFLP profiles where distances are exactly expressed on a two-dimensional ordination (greater distances represent greater dissimilarities). The diversity of the Chloroflexi community in each reactor was calculated using the Shannon H index (Shannon et al. 1963). NMDS and Shannon H index were performed using PAST (PAleontological Statistics Version 3.11) (Hammer et al. 2001).

2.2.4 Quantitative PCR

A protocol was optimized to specifically quantify Chloroflexi using described Chloroflexi primers based on the FISH probes GNSB941 and CFX1223R (Table 3) (Björnsson et al. 2002; Gich et al. 2002). All PCR reactions were performed in a Rotor-Gene 6000 (Corbett Life Science, Sydney, Australia) with SYBR Green PCR Kit (Qiagen, Hilden, Germany). Fluorescence was detected at the end of the extension step. The q-PCR protocol was as follow: 52 °C for 2 min, 95 °C for 10 min followed by 40 cycles at 95 °C for 15 s, 60 °C for 60 s. Each PCR mixture (15 µl) contained 5 µl of diluted DNA template, 10 µM of each primer, and 10x of SYBR Green master mix. DNA concentrations were determined using the Qubit fluorometric quantitative analyser (Life Technologies, Carlsbad, CA, USA). The specificity of the amplified PCR product was assessed by performing a melting curve analysis; fluorescence readings were consecutively collected during the melting process from 60 to 95 °C at a heating rate of 5 °C s^{-1}.

Quantification was made using calibration curves prepared from serial dilutions of DNA from an environmental Chloroflexi (Clone 131, AJ412677). This PCR product was obtained by amplifying the insert from a plasmid using the primers for the plasmid (T3 and T7). To prepare the standard curve, DNA was diluted to yield 10-fold dilutions in the range of 10^6-10^9 DNA copies. All quantitative PCR reactions were carried out in duplicates. The efficiencies of the quantitative PCR assays ranged from 99 % to 106 %, and the R^2 value for each standard curve line always exceeded 0.98. Boxplots were constructed with the results from the quantifications using PAST software. In order to find relationships between the quantified Chloroflexi and the operational parameters, Pearson's correlation analyses were performed using the PAST Software considering a significance level alpha = 0.05 (Hammer *et al*. 2001).

2.2.5 Fluorescence *in situ* hybridization (FISH)

FISH analysis was performed using probes targeting the phylum Chloroflexi GNSB941 and CFX1223R labeled with Cy3 (Björnsson *et al*. 2002; Gich *et al*. 2002) (Table 3). The analysis was performed to determine the morphology of members affiliated to the Chloroflexi phylum detected in these samples. For reactor COB the sample was taken during the study (COB10). For reactor MO the fixed sample was taken two years before our study. Fixation of the samples and FISH was performed according to the protocol of Amann *et al*. (1995). Hybridization was performed according to Björnsson *et al*. (2002), with the conditions established by these authors. For cellular staining 4’, 6’-diamin-2-phenylindole (DAPI) was used. The pure culture of *Sphaerobacter thermophilus* strain DSM 20745 was used as positive control, and *Escherichia coli* strain DH5α was used as negative control. The glass slides with the samples were embedded in Citifluor AF1 (Citifluor, Canterbury, United Kingdom). Images of fluorescent cells were recorded using a confocal laser scanning microscope (LSM 800, Carl Zeiss, Oberkochen, Germany). For detection of DAPI and Cy3-labeled cells, the 405 nm and 543 nm lasers were used, respectively. Samples from AL, COA and SR were not analyzed by FISH due to an experimental error in the fixation procedure.

2.2.6 Bacterial 16S rRNA gene analysis by next generation sequencing

The taxonomic composition of the Chloroflexi community was studied by next generation sequencing of the 16S rRNA gene. For that, one sample per reactor was selected: AL320, SR77, COA286, COB286 and MO205. For samples AL320, SR77, COA286 and COB286 DNA was dehydrated with 95% ethanol and sent to the Institute for Agrobiology Rosario

(INDEAR, Rosario, Argentina) where the pyrosequencing analysis was performed using a Roche Genome Sequencer FLX Titanium. The 16S rRNA genes were amplified with primers for region V4 (set 1): 515F (5'-GTGCCAGCMGCCGCGGTAA-3') and 806R (5'-GGACTACVSGGGTATCTAAT-3') (Caporaso *et al*. 2012) (Table 3). Due to discontinuation of the pyrosequencing platform service, for sample MO205, sequencing was performed on an Ion Torrent PGM (Life Technologies) in the platform at the Biological Research Institute "Clemente Estable" (Montevideo, Uruguay), with primers which cover a similar region for region V4 (set 2): 520F (5'-AYTGGGYDTAAAGNG-3') and 802R (5'-TACNNGGGTATCTAATCC-3') (Claesson *et al*. 2009) (Table 3). The coverage of the primers was checked using the test prime tool in the Silva web page (www.arb-silva.de), setting the parameters to 1 mismatch for the Maximum number of mismatches and 5 bases for the Length of 0-mismatch zone at 3' end, as recommended by the testprime tutorial. The primers presented similar coverage for the classes within the Chloroflexi phylum (Table S2.2). However, *in silico* prediction of primer coverage might not ensure that same microorganisms are detected which limits our study. Testing primers with mock communities and compare the results from environmental samples analyzed by metagenomic approach without PCR biases would be needed to fully evaluate the effectiveness of these primers in detecting most *Chloroflexi* species.

2.2.7 Bioinformatic processing

The QIIME Pipeline Version 1.9.1 was used for amplicon sequence analysis (Caporaso *et al*. 2010). All sequence reads were sorted based on their unique barcodes. After filtering the low-quality reads (coefficient greater than 25), trimming the primers, adapters and barcodes, and filtering reads with a length of less than 200 bases, a total of 4,176 to 229,365 effective reads for all samples were obtained. The chimeras and noise in the sequencing reads were further removed, after which only high-quality reads for these samples remained (number of reads). Sequences were clustered into operational taxonomic units (OTU) using UClust algorithm (Edgar 2010), based on the 97% identity threshold (de novo-based OTU picking strategy). OTUs represented by one sequence (singletons) were removed from the analysis. For the taxonomic classification of the reads, the Silva database release 128 and confidence threshold of 80% was used. The raw data was deposited at National Center for Biotechnology Information (NCBI) under accessing number: PRJNA453814.

2.2.8 Phylogenetic analysis of selected reads

For a detailed phylogenetic analysis, a phylogenetic tree was constructed with representative sequences from OTUs classified within the phylum Chloroflexi with a relative abundance higher than 2%. The sequences were compared with sequences deposited in the NCBI database (www.ncbi.nlm.nih.gov) using the tool BLASTN. A total of 30 sequences corresponding to sequences closely related to ours as well as sequences of cultivated member within the phylum Chloroflexi were used to construct the phylogenetic tree using MEGA 7 (Kumar *et al*. 2016). Sequences were aligned with ClustalW (Thompson *et al*. 1994) and the phylogenetic tree was built applying the maximum likelihood method (Maidak *et al*. 1994) and inferring 1000 replicates of bootstrapping to estimate the reliability of phylogenetic reconstructions (Felsenstein 1985). The analyses of Evolutionary Divergence between Sequences was conducted using the Maximum Composite Likelihood model (Tamura et al 2004) with MEGA7 program. Distance matrix is the result of the number of base substitutions per site from between sequences. Three clusters (1, 2 and 3) were defined in the phylogenetic tree where a branch with high bootstrap value formed by dominant OTUs, other environmental sequences within Chloroflexi and no sequence belonging to a cultured member of the Chloroflexi phylum we detected, as this might indicate possible novel members within the Chloroflexi phylum.

2.3 Results

2.3.1 Chloroflexi were abundant and persistent in UASB reactors

Quantitative PCR showed that microorganisms from the phylum Chloroflexi were present in all samples and persistent over time with values over 10^7 Chloroflexi 16S rRNA gene copy number per ng DNA for all samples analyzed (Figure 2.1 (a-e)).

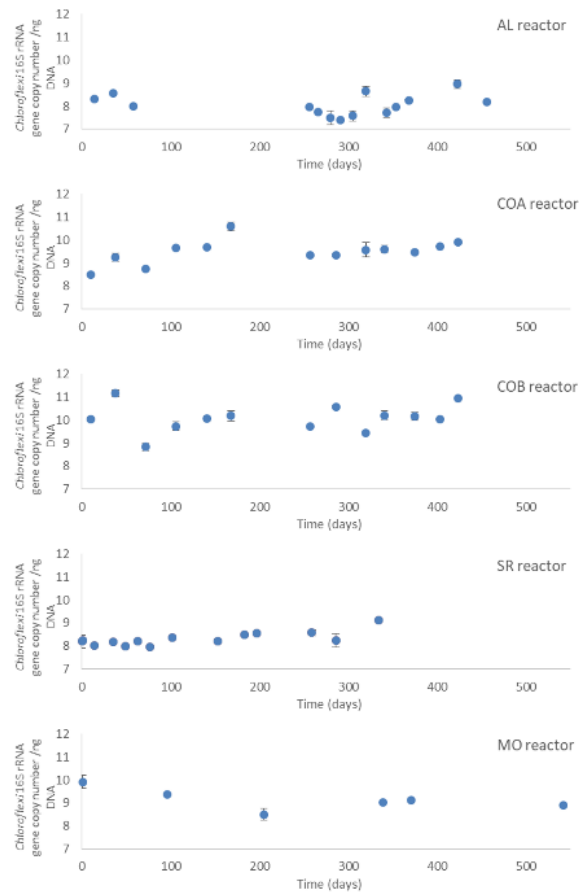

Figure 2.1 Quantification of members of phylum Chloroflexi in samples taken over time (days) for each reactor by specific q-PCR. Each point represent the mean value from duplicate reactions. The error bars represents the standard deviations. The data were presented from each reactor a AL, b COA, c COB, d SR, e MO.

In order to compare the values from the different reactors a boxplot was constructed (Figure 2.2).

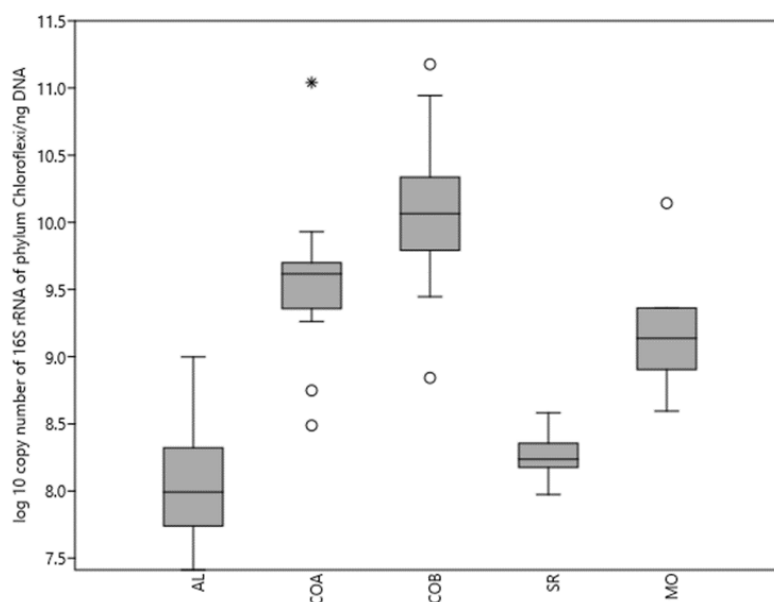

Figure 2.2 Abundance of Chloroflexi phylum in the five studied reactors. BoxPlots of Chloroflexi 16S rRNA gene copy numbers (log 10 copy number per ng of DNA) from each reactor is presented. The bottom and the top of each box present the first and third quartiles, the band inside is the median and the whiskers (short horizontal lines) represent the minimum and maximum value of each data set. Values outside the inner fences are shown as circles; values further than 3 times the box height from the box (the "outer fences") are shown as stars.

The samples from reactors AL and SR, studied during start-up periods, presented lower abundance of Chloroflexi (mean values $9.7 \times 10^7 \pm 7.9 \times 10^6$, $1.7 \times 10^8 \pm 2.5 \times 10^7$ Chloroflexi 16S rRNA gene copy number per ng DNA, respectively) than the samples from reactors COA, COB, and MO, (mean values $4.1 \times 10^9 \pm 1.5 \times 10^9$, $1.2 \times 10^{10} \pm 3.3 \times 10^9$ and $1.2 \times 10^9 \pm 2.8 \times 10^8$ Chloroflexi 16S rRNA gene copy number per ng DNA, respectively) which had years of stable operational conditions. Reactor COB presented the highest abundance while reactor AL the lowest.

Linear correlation analysis was applied to examine relationships between the abundance of the phylum Chloroflexi and the operational data of the reactors (Table 2.4).

Table 2.4 Pearson Linear correlation parameters between abundance of the phylum Chloroflexi and reactor's operational data. The data analysis was performed using PAST Software

Reactor		VSS (g/l)*	Organic load (kgCOD/m^3.d)	HTR (days) †	CH_4 (%) ‡	pH	T(C°) §
AL	r	0.397	-0.078	0.085	-0.518	-0.328	0.027
	p	0.179	0.830	0.815	0.125	0.472	0.927
COA	r	n/d	-0.101	n/d	n/d	0.585	0.267
	p	n/d	0.871	n/d	n/d	0.059	0.427
COB	r	n/d	0.897	n/d	n/d	-0.064	-0.146
	p	n/d	0.039	n/d	n/d	0.852	0.668
SR	r	0.654	-0.173	-0.231	-0.298	0.248	0.546
	p	0.040	0.711	0.551	0.516	0.752	0.056
MO	r	0.233	n/d	n/d	n/d	-0.445	-0.610
	p	0.767	n/d	n/d	n/d	0.555	0.198

* Volatile suspended solids VSS

† Hydraulic retention time HTR

‡ Biogas production (CH_4)

§ Temperature n/d: no data available for each sample date

The abundance of the phylum Chloroflexi was positively correlated with the load applied in samples from reactor COB (Pearson test, n = 13; r^2 = 0.80; *p-value* < 0.05). A positive correlation between the abundance of the phylum Chloroflexi and the amount of biomass in the reactor (volatile suspended solids, VSS) in samples from reactor SR was also detected (Pearson test, n = 14; r^2 = 0.95; *p-value* < 0.05). No other significant correlations between the abundance of Chloroflexi and reactor operational data were observed.

2.3.2 High dominance and stability of a particular Chloroflexi T-RF detected by specific T-RFLP

The results from T-RFLP analysis showed strong dominance of one Terminal Restriction Fragment (T-RF) and a low diversity for all reactors (Figure 2.3 (a)). For samples from reactors AL, COA, COB and SR, the predominant T-RF detected presented a length of 511 bp while for samples from reactor MO a T-RF of 501 bp was dominant. These predominant T-RFs presented high stability over time while T-RFs with less relative abundance were more variable.

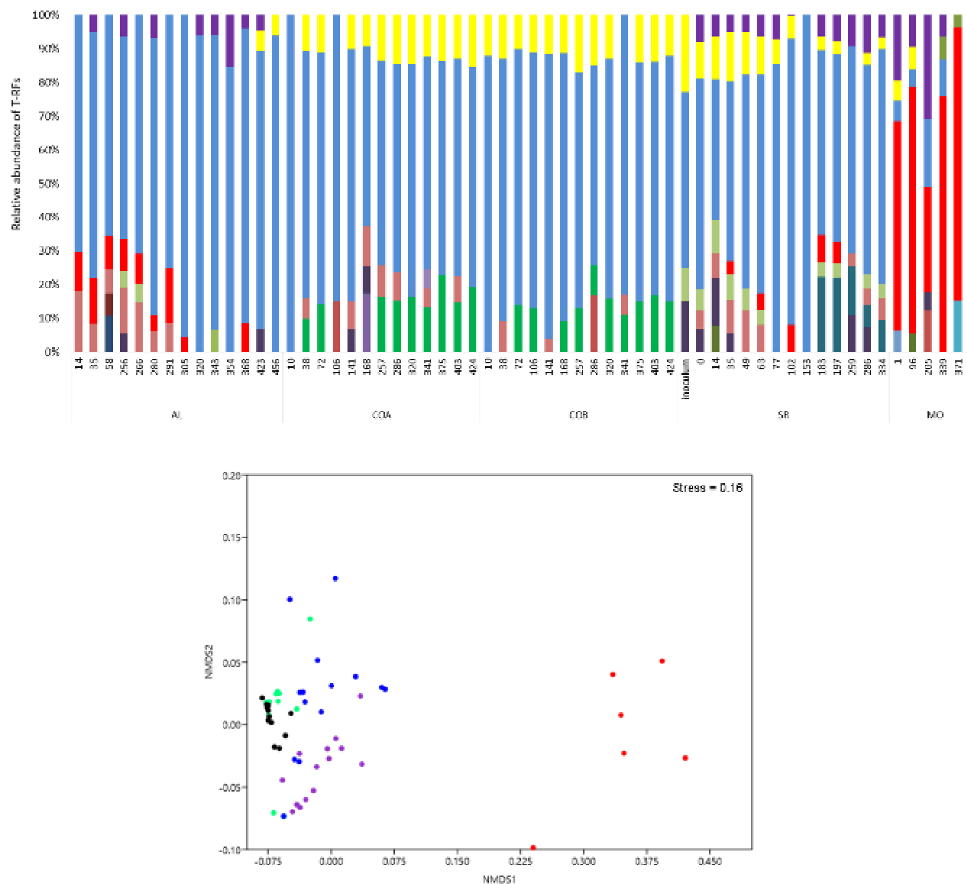

Figure 2.3 A) Specific Chloroflexi T-RFLP analysis of reactor samples taken along operation time. Relative abundances of T-RFs are indicated in different colors. Only T-RFs with relative abundance higher than 3 % for each sample were considered. (■) 52 (■) 58 (■) 70 (■) 79 (■) 87 (■) 140 (■) 148 (■) 161 (■) 167 (■) 182 (■) 200 (■) 272 (■) 369 (■) 447 (■) 457 (■) 470 (■) 501 (■) 511 (■) 520 (■) 625 (■) 633 B) NMDS ordination using T-RFLP data (●) AL (●) COA (●) COB (●) SR (●) MO. Community similarity is based on the Bray–Curtis index which includes the presence/absence as well as the relative abundances of T-RFs. The stress value of 0.16 indicates good quality of the NMDS result.

According to the Shannon similarity index the Chloroflexi community of reactor MO was the most diverse while the Chloroflexi community of reactor COB was the least diverse (mean values obtained for the samples taken from each reactor were: reactor AL: 0.84; reactor COA:1.05; reactor COB: 0.81; reactor SR: 1.32 and reactor MO: 1.37).

The NMDS analysis showed that samples from reactors AL, COA, COB and AL grouped closely while MO samples formed a separate group. This grouping was related with the

inoculum source which was the same for reactors AL, COA, COB and SR and different for reactor MO (Figure 3b).

2.3.3 Uncultured Chloroflexi dominate the communities

One sample per reactor was analyzed by 16S rRNA gene amplicon sequencing. The number of reads obtained after processing the raw data were as follows: AL320: 4,063 reads; COA286: 12,202 reads; COB286: 12,257 reads; SR77: 4,307 reads and MO205: 228,109 reads. In order to compare between samples, the numbers of sequences were normalized to the same sequencing depth (4024 reads) using QIIME.

The community composition at phylum level showed that Chloroflexi was detected in all reactors with a relative abundance higher than 20% in four of the five samples. A relative abundance of 2.8% were observed in sample MO205 (Figure 2.4 (a)). Other predominant phyla shared by the five reactor samples were: Firmicutes, Proteobacteria, Spirochaetes and Synergistetes.

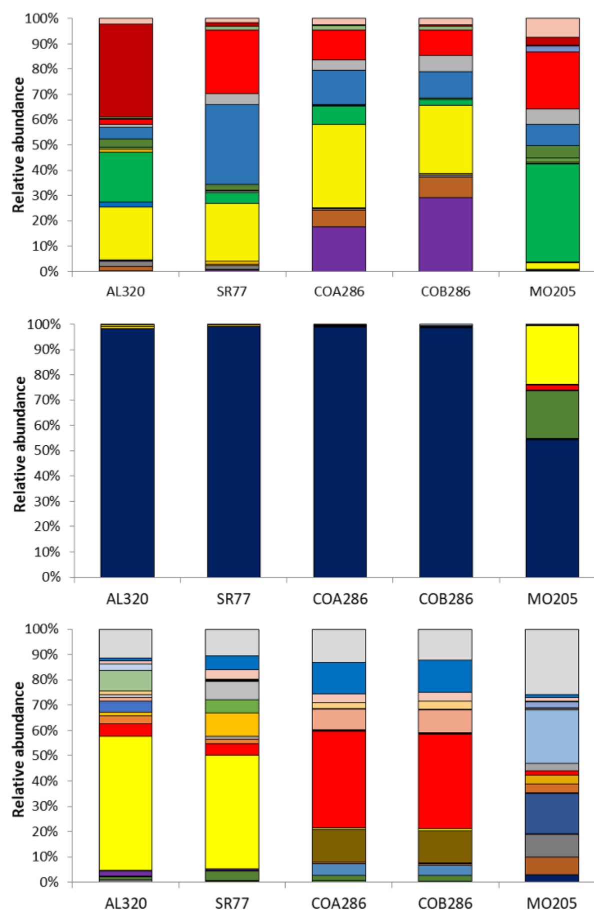

Figure 2.4 Taxonomic distribution according to the 16S rRNA gene amplicon sequencing analysis. One sample from each reactor were analyzed. The results were presented as: a at

phylum level: (■) Acidobacteria (■) Aminicenantes (■) Bacteroidetes (■) Caldiserica (■) Chlamydiae (■) Chloroflexi (■) Fibrobacteres (■) Firmicutes (■) Hydrogenedentes (■) Lentisphaerae (■) Omnitrophica (■) Parcubacteria (■) Planctomycetes (■) Proteobacteria (■) Spirochaetae (■) Synergistetes (■) Tenericutes (■) Thermotogae (■) Verrucomicrobia (■) Others lower than 1%. b at class level within the phylum Chloroflexi: (■) Anaerolineae (■) Caldilineae (■) Chloroflexia (■) Dehalococcoidia (■) Gitt-GS-136 (■) JG30-KF-CM66 (■) JG37-AG-4 (■) KD4-96 (■) Ktedonobacteria (■) SBR2076 (■) SHA-26 (■) SJA-15 (■) SJA-68 (■) Thermoflexia (■) Thermomicrobia (■) TK10 (■) unc. Bacteria. c at OTU level within the phylum Chloroflexi: (■) OTU271 (■) OTU299 (■) OTU304 (■) OTU371 (■) OTU677 (■) OTU733 (■) OTU928 (■) OTU1228 (■) OTU1353 (■) OTU1407 (■) OTU1632 (■) OTU1649 (■) OTU1650 (■) OTU1668 (■) OTU1821 (■) OTU2086 (■) OTU2360 (■) OTU2545 (■) OTU2612 (■) OTU2616 (■) OTU2620 (■) OTU2814 (■) OTU3200 (■) OTU3208 (■) OTU3509 (■) OTU4169 (■) OTU4170 (■) Others lower than 2%. The sequences showing an abundance below 1.0% for phyla a and 2% for OTUs c in all samples were grouped into ‘Others’.

In order to determine the taxonomic composition of Chloroflexi in each sample, a specific analysis of the sequencing data was performed using only the sequences affiliated to the phylum Chloroflexi (Figure 2.4 (b)). The Anaerolineae class was dominant in all reactors, with the following relative abundance: 98.2% in sample AL320; 99.3% in sample SR77; 99.0% in sample COA286, 98.7% in sample COB286; and 54.3% in sample MO205. The sample from reactor MO presented co-dominance of the classes Dehalococcoidia (19.0%) and SJA-15 (22.9%).

In all samples analyzed one particular OTU was predominant which was in accordance with T-RFLP results (Figure 2.4 (c)). This predominant OTU was shared between reactors AL and SR (OTU 1649), and between reactors COA and COB (OTU 1650). OTU 677 and OTU 2612 were co-dominant and specific for reactor MO. The OTU 1650 was the only OTU shared by all samples with relative abundance higher than 1.8 % within the Chloroflexi phylum.

A phylogenetic tree was constructed in order to get a closer insight into the phylogeny of the *Chloroflexi* sequences present in the reactors (Figure 2.5).

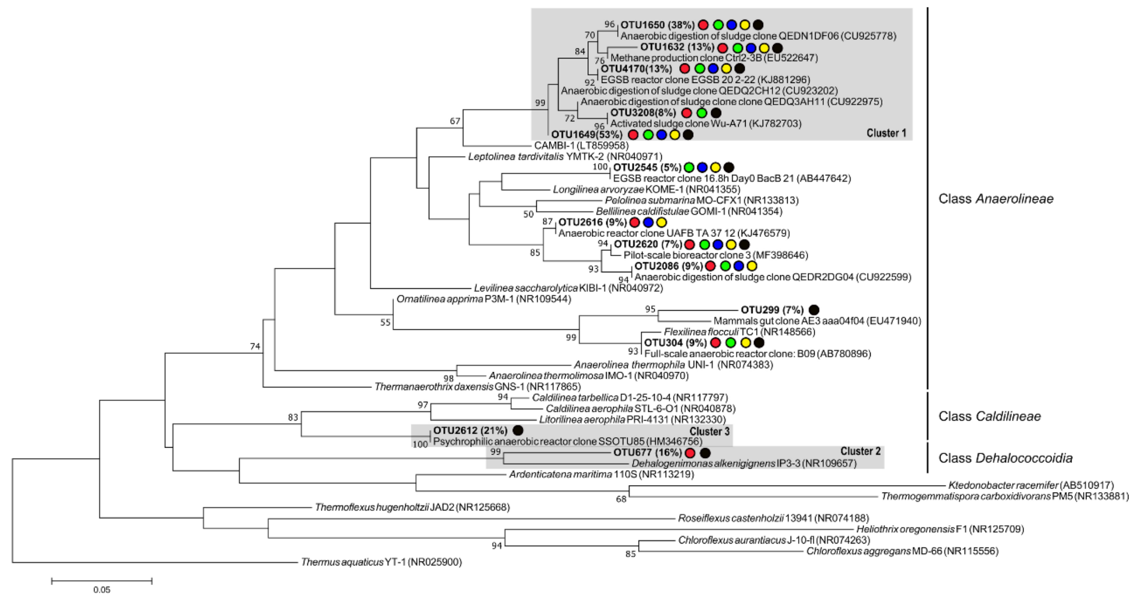

Figure 2.5 Phylogenetic distance tree of the bacterial phylum Chloroflexi constructed based on 16S rRNA gene sequence, including environmental clones, cultured representatives belonging to the phylum Chloroflexi and OTU sequences classified within the phylum Chloroflexi (🔴) AL (🔵) COA (🟡) COB (🟢) SR (⚫) MO. Bootstrap values (greater than 50%) obtained with the maximum likelihood method based on 1000 replications are shown at branch nodes. *Thermus aquaticus* YT-1 was used as outgroup. Bar indicates 0.05 substitutions per site. The relative abundance of the OTUs in the samples is show in brackets. Each reactor was identified with a different label. Only OTUs higher than 5% was used in the analysis.

The predominant OTUs in reactors AL, SR, COA and COB (OTU 1632, 1649, 1650, 3208, 4170) clustered together with sequences classified as uncultured Anaerolineae retrieved from anaerobic and aerobic reactors. These sequences form a branch without any cultured representatives (Cluster 1, Figure 2.5). The sequence from CAMBI-1, whose genome was obtained by metagenome sequencing, presented the lowest phylogenetic distance with these OTUs ranging from 0.02 to 0.03 (Figure 2.5).

The predominant OTUs detected in the sample from reactor MO clustered with a sequence from a cultured member of Dehalogenimonas genus (OTU 677), with a phylogenetic distance of 0.03 (Cluster 2), and cultured members of Caldilineae class (OTU 2612), with a phylogenetic distance of 0.04 (Cluster 3). Sequence from OTU 2612 were positioned in a branch with a sequence retrieved from a psychrophilic anaerobic sediment which was used to inoculate anaerobic psychrophilic reactors (Xing *et al*., 2010).

2.3.4 FISH revealed a filamentous morphology of Chloroflexi cells

The simultaneous hybridation with GNSB41 and CFX1223R probes was positive in samples from reactors COB and MO (Figure 2.6 (a, b)).

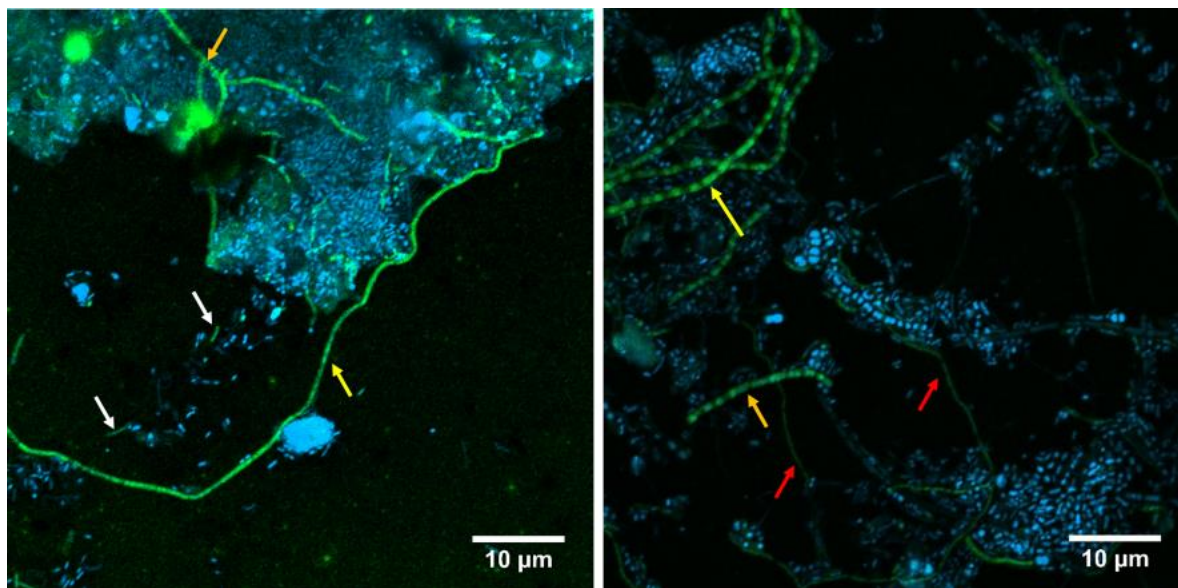

Figure 2.6 Micrographs taken from samples from COB and MO by confocal laser scanning microscopy (digitally produced images with pseudocolors). The samples were simultaneously hybridized with probes Cy3-labeled directed to the phylum Chloroflexi GNSB941 and CFX1223R (green), and DAPI staining (blue) (bar, 10 µm) a COB and b MO. Arrows indicate different types of filaments detected in each reactor.

The microscopic images showed that all labeled cells with the Chloroflexi-specific probe had filamentous morphology. The analyzed sample from reactor MO presented two types of filaments, thin and thick (Figure 2.6 (b), red arrow and yellow arrow, respectively). On the other hand, in reactor COB very short filaments (Figure 2.5 (a), white arrow) and thin filaments (Figure 2.5 (a), orange arrow) were detected.

2.4 Discussion

2.4.1 Abundance of Chloroflexi in UASB reactors

q-PCR quantification showed that *Chloroflexi* were abundant in all anaerobic UASB reactors studied during the monitored operation period. High abundances of Chloroflexi have been detected in other anaerobic reactors and it has been suggested that they are part of the core community (Chouari et al. 2005; Rivière et al. 2009). However, no Chloroflexi-specific q-

PCR quantification was performed in these samples and the abundance was calculated from the relative abundance of clone sequences.

In aerobic activated sludge reactors treating municipal wastewater, specific Chloroflexi q-PCR was used and values within the range of 5.0 x 10^4 to 3.0 x 10^5 16S rRNA gene copy number per ng of DNA were found (Yoon et al. 2010; Nittami et al. 2017). According to this, the abundance found in the anaerobic reactors studied here (between 9.7 x 10^7 ± 7.9 x10^6 and 1.2 x 10^{10} ± 3.3 x10^9 Chloroflexi 16S rRNA gene copy number ng of DNA) was higher than the abundance detected in the above mentioned aerobic reactors. Chloroflexi has also been quantified by q-PCR with the same set of primers as in our work in other anaerobic environments like marine sponges, presenting an abundance between 5.0 x10^5 and 5.2 x 10^5 Chloroflexi 16S rRNA gene copy number ng DNA which is lower than the range of values found in our reactors (Bayer et al. 2017). This further supports the importance of these microorganisms in anaerobic UASB reactors.

Regarding the abundance of Chloroflexi in the different UASB reactors studied, the sludge from reactors COA, COB and MO with granular biomass presented higher abundance of Chloroflexi than reactors AL and SR. It has to be taken into account that reactors AL and SR were studied during start-up and were inoculated with disperse biomass from a anaerobic lagoon. The relative abundance of Chloroflexi in reactors AL and SR increased after operation day 280, moment where an increase in the granulation was detected (Fernández 2016; López et al. 2017).

It has been previously proposed that filamentous Chloroflexi-like bacteria contribute to the formation and maintenance of dense and rigid granule structures by growing in layers around other microorganisms (Ariesyady et al. 2007) and play an important role in sludge granulation in mesophilic and thermophilic anaerobic reactors (Uemura et al. 1993; Sekiguchi et al. 1999, 2001; Yamada et al. 2005). The filamentous morphology favors retention in the granule as well as in the open structure of the flocs in activated sludge (Yamada et al. 2005; Yamada et al. 2007a). This is in accordance with the “spaghetti theory” of granulation, which considers the presence of filamentous bacteria important for making cores of sludge granules by constructing web-like structures (Wiegant, 1987). Due to the filamentous morphology of these bacteria, it is possible to suggest that Chloroflexi members might have a key role in granule formation. Their participation in granule formation retains Chloroflexi filaments as they form a web structure and are surrounded by other bacteria. This might be the reason for the high abundances of Chloroflexi in anaerobic reactors with granular biomass such as the UASB reactors studied here. It can also explain the higher abundance of *Chloroflexi* in

reactors with granular biomass (COA, COB and MO) compared to reactors with flocculent biomass (AL and SR), and the increase of abundance observed with the increase of granulation in start-up reactors. A deeper FISH analysis using the whole granule would give further information on their location within the granule. Moreover FISH using more specific probes targeting the predominant OTUs detected combined with metagenomic analysis would give more insights into their role in these systems. However, optimizing FISH for uncultured groups of microorganisms represents a technical challenge due to the lack of positive controls using isolated strains.

It has been previously shown that high abundance of Chloroflexi might cause bulking episodes (Sekiguchi et al. 2001; Borzacconi et al. 2008; Li et al. 2008; Yamada et al. 2011). Although there was high abundance of Chloroflexi in all reactors, no bulking episodes of the sludges were observed. No data of the abundance of this group in bulking biomass are available, then more work is needed to determine the bulking threshold values.

The positive correlation between Chloroflexi abundance and the load applied for reactor COB has been observed previously in UASB reactors where there was an increment of Chloroflexi members due to the increase in the carbohydrate content (Sekiguchi et al. 2001). In this example, the reactor presented sludge bulking problems with an organic load six times higher than in reactor COB which had good settling properties. Suggesting that the overgrowth of these microorganisms can be caused by a strong increment in the substrate concentration.

2.4.2 The biomass used as inoculum influences the Chloroflexi community structure and composition

Reactors AL, SR, COA and COB were inoculated with sludge from the same anaerobic lagoon and had the same predominant TRFs in the T-RFLP analysis, while reactor MO, inoculated with sludge from a different lagoon, had a different predominant TRF. This was also reflected in the sequencing results where the microbial composition was different in the MO reactor. However, it has to be taken into account that a different primer pair was used for this reactor.

The two predominant OTUs obtained for reactors AL, SR, COA, COB were the same, and were affiliated to the Anaerolineae class, while OTUs affiliated to Caldilineae and Dehalococcoidia were dominant in reactor MO.

Even though reactors COA and COB had been operating in stable conditions for several years and reactors AL and SR were in a transition state, the predominant Chloroflexi selected were

the same. This suggests that the inoculum source influence the predominant member of Chloroflexi. It has to be taken into account that the chemical composition of the wastewater treated in reactor AL (vinasse generated during ethanol production) differ from the wastewater treated in the dairy industries (Table S2.3). This confirms the idea that the inoculum shapes the Chloroflexi community and not the wastewater composition.

These findings are in accordance with the results from a lab scale spiral symmetry stream anaerobic bioreactor (S SAB) presented by Xu et al. (2017). The authors reported that the abundance of Chloroflexi increased over time and that the most dominant genus were Levilinea and Longilinea. They showed that these genera were present in the inoculum and were maintained throughout the whole operation, indicating that they play an important role in sludge granulation and growth (Xu et al. 2017). Moreover, it was demonstrated that the inoculum source deterministically contributes in shaping the total community structure and specific ecosystem function (Perrotta et al. 2017).

The predominant members of Chloroflexi detected in the five UASB reactors presented stability and low dynamics over time in each reactor. This is in contrast with the results obtained for the total microbial community where a high dynamics and no stability was found in UASB reactors by T-RFLP analysis (Pycke et al. 2011). This could be related to the important role of this phylum in the granule formation, the fermentative metabolism utilizing probably carbohydrates and their capacity to be retained in granules.

2.4.3 Uncultured Chloroflexi were predominant in the reactors

The Chloroflexi community dominating four of the five UASB reactors presented a high abundance of OTUs classified within the Anaerolineae class which formed a separate branch in the phylogenetic tree with no cultured representative. Moreover, the sequences in this cluster were detected with a relative abundance between 55.2 % and 66.9 % of the total Chloroflexi sequences. Anaerolineae has been reported as a group with natural and biotechnological relevance (Yamada and Sekiguchi 2009), indicating that these organisms play a relevant role in these systems. Our results indicate that probably a new Anaerolineae species dominates the Chloroflexi community in these UASB reactors. High relative abundance of Anaerolineae is frequently reported in anaerobic reactors by metagenomic sequencing or amplicon sequencing (Guo et al. 2015; Shu et al. 2015; da Silva Martins et al. 2017; McIlroy et al. 2017). Moreover, uncultivated clades related to Chloroflexi were ubiquitously found in anaerobic reactors and enrichments (Narihiro et al. 2015). These

features suggest that efforts should be made to isolate more members of this group in order to better understand their physiology. This knowledge could be used to control their overgrowth in industrial reactors.

Regarding reactor MO, predominant Chloroflexi OTUs were classified within the class Dehalococcoidia with an abundance of 19.0 % of the total Chloroflexi sequences. Members of this class were isolated from an anaerobic reactor, river sediments and contaminated groundwater (Seshadri et al. 2005; Moe et al. 2009; Yan et al. 2009; Löffler et al. 2013; Key et al. 2017). Some cultivated members of Dehalococcoidia class are implicated in the bioremediation of chlorinated solvent-contaminated sites and most rely exclusively on the anaerobic respiration of halogenated hydrocarbons (Löffler and Edwards 2006). The most abundant OTU in reactor MO was closely related with Dehalogenimonas sp., this is in accordance with the results obtained by Enitan et al. (2017) where this group was found in a full-scale UASB reactor treating brewery wastewater. The role of Dehalogenimonas related OTUs in the treatment of this substrate remains unknown and further research is needed. However, the metabolism of some members of Dehalococcoidia class was elucidated from metagenomics data, the results suggests that they are anaerobes which are likely involved in sugar and plant-derived-compound degradation to acetate (Hug et al. 2013). Therefore, these group might be involved in carbon degradation in UASB reactors instead of respiration of halogenated hydrocarbons.

A second OTU predominated in the sample from reactor MO and according to the phylogenetic tree, the sequence from this OTU was positioned in a branch with sequences from the Caldilineae class. The culture representatives of this class presents the capacity to degrade carbohydrates (Table 1). None of these strains were isolated from anaerobic reactors, but they are frequently found in aerobic reactors (Kragelund et al. 2011, Zhang et al. 2012). As some of the species can also grow in anaerobic conditions, it could be expected to find these groups also in anaerobic reactors, but the presence of this group is not frequently reported in anaerobic reactors. In conclusion, the phylum Chloroflexi was abundant in all studied UASB methanogenic reactors and exhibited various filamentous morphologies in two of the reactors. Members of the Anaerolineae class were predominant, stable over time and were not affiliated to any species from cultured representatives indicating that the diversity is still poorly described in anaerobic reactors. According to our results, the inoculum source might influence which Chloroflexi species will prevail in each reactor and therefore the Chloroflexi diversity. In UASB reactors with longer operation time and granular biomass, Chloroflexi was more abundant than in reactors in start-up period with flocculent biomass,

indicating that Chloroflexi have a key role in the granulation in methanogenic UASB reactors.

This preliminary study contributes to the understanding of the role of filamentous Chloroflexi in full scale UASB reactors. Future studies using new approaches as single cells analysis and genome reconstruction through metagenomic data may help to elucidate their metabolic potential. The knowledge about this bacterial phylum in conjunction with reactors operational parameters and analysis of bulking UASB sludge will be of great value to develop better and more efficient control strategies to prevent sludge bulking by overgrowth of members of phylum Chloroflexi.

Acknowledgments

We thank Dr. Cecilia Callejas and Dr. Liliana Borzacconi for the useful suggestions and Dr. José Sotelo for the assistance during the bioinformatic work.

Conflict of interest

No conflict of interest declared.

2.5 References

Amani T, Nosrati M, Sreekrishnan TR (2010) Anaerobic digestion from the viewpoint of microbiological, chemical, and operational aspects - a review. Environ Rev 18:255–278 doi: 10.1139/A10-011

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169 doi: 10.1016/j.jip.2007.09.009

Angelidaki I, Karakashev D, Batstone DJ, Plugge CM, Stams AJM. (2011) Biomethanation and its potential. Methods Enzymol 494:327–351 doi: 10.1016/B978-0-12-385112-3.00016-0

Angenent LT, Karim K, Al-Dahhan MH, Wrenn BA, Domíguez-Espinosa R (2004) Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends Biotechnol 22:477–485 doi: 10.1016/j.tibtech.2004.07.001

Ariesyady HD, Ito T, Okabe S (2007) Functional bacterial and archaeal community structures of major trophic groups in a full-scale anaerobic sludge digester. Water Res 41:1554–1568 doi: 10.1016/j.watres.2006.12.036

Bayer K, Kamke J, Hentschel U (2017) Quantification of bacterial and archaeal symbionts in high and low microbial abundance sponges using real-time PCR. FEMS Microbiol Ecol 89:679–690 doi: 10.1111/1574-6941.12369

Björnsson L, Hugenholtz P, Tyson GW, Blackall LL (2002) Filamentous *Chloroflexi* (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. Microbiol 148:2309–2318 doi: 10.1099/00221287-148-8-2309

Borzacconi L, López I, Passeggi M, (2006) Start-up and seady state results of a full scale UASB reactor treating malting wastewater. Water Sci Technol 54(2):261–268. PMID: 16939111

Borzacconi L, López I, Passeggi M, Etchebehere C, Barcia R (2008) Sludge deterioration in a full scale UASB reactor after a pH drop working under low loading conditions. Water Sci Technol 57:797–802 doi: 10.2166/wst.2008.071

Bowman KS, Nobre MF, da Costa MS, Rainey FA, Moe WM (2012). *Dehalogenimonas alkenigignens* sp. nov., a chlorinated alkane dehalogenating bacterium isolated from groundwater. Int J Syst Evol Microbiol 63:1492–1498 doi: 10.1099/ijs.0.045054-0

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, Mcdonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) Correspondence QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. Nat Publ Gr 7:335–336 doi: 10.1038/nmeth0510-335

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6 (8):1621–1624 doi: 10.1038/ismej.2012.8

Chong S, Sen TK, Kayaalp A, Ang HM (2012) The performance enhancements of upflow anaerobic sludge blanket (UASB) reactors for domestic sludge treatment - A State-of-the-art review. Water Res 46:3434–3470 doi: 10.1016/j.watres.2012.03.066

Chouari R, Le Paslier D, Daegelen P, Ginestet P, Weissenbach J, Sghir A (2005) Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester. Environ Microbiol 7:1104–1115 doi: 10.1111/j.1462-2920.2005.00795.x

Claesson MJ, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS One 4 doi: 10.1371/journal.pone.0006669

da Silva Martins A, Ornelas Ferreira B, Ribeiro NC, Martins R, Rabelo Leite L, Oliveira G, Colturato LF, Chernicharo CA, de Araujo JC (2017) Metagenomic analysis and performance of a mesophilic anaerobic reactor treating food waste at various load rates. Environ Technol 38:2153–2163 doi: 10.1080/09593330.2016.1247197

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–1 doi: 10.1093/bioinformatics/btq461

Edwards U, Rogall T, Blocker H, Emde M, Bottger EC (1989) Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. Nucleic Acids Res 17(19):7843–7853 PMID: 2798131

Enitan AM, Odiyo JO, Swalaha FM, Collection AS (2017) Identification of microbial community in an anaerobic reactor treating brewery wastewater. World Acad Sci Eng Technol 11:766–770

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791 doi: 10.1111/j.1558-5646.1985.tb00420.x

Fernández A (2016) Puesta en marcha de un reactor UASB modificado para el tratamiento de efluente lácteo.

Evaluación del desempeño del reactor y adaptación del inóculo. Tesis de maestría en Ingeniería Química.

Gich F, Garcia-Gil J, Overmann J (2002) Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. Arch Microbiol 177:1–10 doi: 10.1007/s00203-001-0354-6

Grégoire P, Bohli M, Cayol JL, Joseph M, Guasco S, Dubourg K, Cambar J, Michotey V, Bonin P, Fardeau, ML, Ollivier B (2011a) *Caldilinea tarbellica* sp. nov., a filamentous, thermophilic, anaerobic bacterium isolated from a deep hot aquifer in the Aquitaine Basin. Int J Syst Evol Microbiol 61:1436–1441. doi: 10.1099/ijs.0.025676-0

Grégoire P, Fardeau, ML, Joseph M, Guasco S, Hamaide F, Biasutti S, Michotey V, Bonin P, Ollivier B (2011b). Isolation and characterization of *Thermanaerothrix daxensis* gen. nov., sp. nov., a thermophilic anaerobic bacterium pertaining to the phylum "*Chloroflexi*", isolated from a deep hot aquifer in the Aquitaine Basin. Syst Appl Microbiol. 34:494–497 doi: 10.1016/j.syapm.2011.02.004

Guo J, Peng Y, Ni B-J, Han X, Fan L, Yuan Z (2015) Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing. Microb Cell Fact 14:33 doi: 10.1186/s12934-015-0218-4

Hammer Ø, Harper DT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron. 4:1–9.

Hinken L, Huber M, Weichgrebe D, Rosenwinkel KH (2014) Modified ADM1 for modelling an UASB reactor laboratory plant treating starch wastewater and synthetic substrate load tests. Water Res 64:82–93 doi: 10.1016/j.watres.2014.06.044

Hug L, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR, Williams KH, Tringe SG, Banfield JF (2013) Community genomic analyses constrain the distribution of metabolic traits across the *Chloroflexi* phylum and indicate roles in sediment carbon cycling. Microbiome 1:22 doi: 10.1186/2049-2618-1-22

Imachi H, Sakai S, Lipp JS, Miyazaki M, Saito Y, Yamanaka Y, Hinrichs KU, Inagaki F, Takai K. (2014) *Pelolinea submarina* gen. nov., sp. nov., an anaerobic, filamentous bacterium of the phylum *Chloroflexi* isolated from subseafloor sediment. Int J Syst Evol Microbiol 64:812–818 doi: 10.1099/ijs.0.057547-0

Kale V, Björnsdoóttir S.H, Friðjónsson ÓH, Pétursdóttir SK, Ómarsdóttir S, Hreggviðsson GÓ (2013) *Litorilinea aerophila* gen. nov., sp. nov., an aerobic member of class *Caldilineae*, phylum *Chloroflexi*, isolated from an intertidal hot spring. Int J Syst Evol Microbiol 63:1149–1154 doi: 10.1099/ijs.0.044115-0

Key TA, Bowman KS, Lee I, Chun J, da Costa M, Albuquerque L, Rainey FA, Moe WM (2017) Dehalogenimonas formicexedens sp. nov., a chlorinated alkane respiring bacterium isolated from contaminated groundwater. Int J Syst Evol Microbiol doi: 10.1099/ijsem.0.001819

Kragelund, C., Thomsen, T.R., Mielczarek, A.T., Nielsen, P.H., 2011. Eikelboom's morphotype 0803 in activated sludge belongs to the genus *Caldilinea* in the phylum *Chloroflexi*. FEMS Microbiol Ecol 76 (3):451-462. doi: 10.1111/j.1574-6941.2011.01065.x

Kumar S, Stecher G, Tamura K (2016) MEGA7 : Molecular Evolutionary Genetics Analysis version 7. 0 for bigger datasets. 1–11 doi: 10.1093/molbev/msw054

Lauber CL, Zhou N, Gordon JI, Knight R, Fierer N (2010) Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples. FEMS Microbiol Lett 307: 80-86 doi: 10.1111/j.1574-6968.2010.01965.x.

Li J, Hu B, Zheng P, Qaisar M, Mei L (2008) Filamentous granular sludge bulking in a laboratory scale UASB reactor. Bioresour Technol 99:3431–3438 doi: 10.1016/j.biortech.2007.08.005

Löffler FE, Edwards EA (2006) Harnessing microbial activities for environmental cleanup. Curr Opin Biotechnol 17:274–284 doi: 10.1016/j.copbio.2006.05.001

Löffler FE, Yan J, Ritalahti KM, Adrian L, Edwards EA, Konstantinidis KT, Müller JA, Fullerton H, Zinder SH, Spormann AM (2013) *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia* classis nov., order *Dehalococcoidales* ord. nov. and famil. Int J Syst Evol Microbiol 63:625–635 doi: 10.1099/ijs.0.034926-0

López I (2016) The potential of biogas production in Uruguay. Renew Sustain Energy Rev 54:1580–1591 doi: 10.1016/j.rser.2015.10.099

López I, Borzacconi L, Passeggi M (2017) Anaerobic treatment of sugar cane vinasse: Treatability and real-scale operation. J Chem Technol Biotechnol doi: 10.1002/jctb.5493

Maidak BL, Larsen N, McCaughey MJ, Overbeek R, Olsen GJ, Fogel K, Blandy J, Woese CR (1994) The Ribosomal Database Project. Nucleic Acids Res 22:3485–3487 PMID: 7524021

McIlroy SJ, Kirkegaard RH, Dueholm MS, Fernando E, Karst SM, Albertsen M, Nielsen PH (2017) Culture-independent analyses reveal novel anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. Front Microbiol 8:1134 doi: 10.3389/fmicb.2017.01134

Moe WM, Yan J, Nobre MF, da Costa MS, Rainey FA (2009). *Dehalogenimonas lykanthroporepellens* gen. nov., sp. nov., a reductively dehalogenating bacterium isolated from chlorinated solvent-contaminated groundwater. Int J Syst Evol Microbiol 59:2692–2697 doi: 10.1099/ijs.0.011502-0

Narihiro T, Terada T, Kikuchi K, Iguchi A, Ikeda M, Yamauchi T, Shiraishi K, Kamagata Y, Nakamura K, and Sekiguchi Y (2009) Comparative analysis of bacterial and archaeal communities in methanogenic sludge granules from upflow anaerobic sludge blanket reactors treating various food-processing, high-strength organic wastewaters. Microbes Environ 24:88–96 doi: 10.1264/jsme2.ME08561.

Narihiro T, Nobu MK, Kim NK, Kamagata Y, Liu WT (2015) The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. Environ Microbiol 17:1707–1720 doi: 10.1111/1462-2920.12616

Nittami T, Speirs LBM, Yamada T, Suzuki I, Fukuda J, Kurisu F, Seviour RJ (2017) Quantification of *Chloroflexi* Eikelboom morphotype 1851 for prediction and control of bulking events in municipal activated sludge plants in Japan. Appl Microbiol Biotechnol 101:3861–3869 doi: 10.1007/s00253-016-8077-4

Nunoura T, Hirai, M, Miyazaki, M, Kazama H, Makita H, Hirayama, H, Furushima Y, Yamamoto H, Imachi H. Takai K (2013). Isolation and characterization of a thermophilic, obligately anaerobic and heterotrophic marine *Chloroflexi* bacterium from a *Chloroflexi*-dominated microbial community associated with a Japanese shallow hydrothermal system, and proposal for *Thermomarinilinea lacunofontalis* gen. nov., sp. nov. Microbes Environ 28:228–235 doi: 10.1264/jsme2.ME12193.

Park, S.J., C.H. Kang, and S.K. Rhee. 2006. Characterization of the microbial diversty in a Korean solar saltern by 16S rRNA gene analysis. J Microbiol Biotechnol 16:1640-1645

Passeggi M, López I, Borzacconi L (2009) Integrated anaerobic treatment of dairy industrial wastewater and

sludge. Water Sci Technol 59:501–506 doi: 10.2166/wst.2009.010

Perrotta AR, Kumaraswamy R, Bastidas-Oyanedel JR, Alm EJ, Rodríguez J (2017) Inoculum composition determines microbial community and function in an anaerobic sequential batch reactor. PLoS One 12:1–14 doi: 10.1371/journal.pone.0171369

Podosokorskaya, O. A., Bonch-Osmolovskaya, E. A., Novikov, A. A., Kolganova, T. V., and Kublanov, I. V. (2013). *Ornatilinea apprima* gen. nov., sp. nov., a cellulolytic representative of the class *Anaerolineae*. Int. J Syst Evol Microbiol 63:86–92 doi: 10.1099/ijs.0.041012-0

Pycke BFG, Etchebehere C, Van De Caveye P, Negroni A, Verstraete W, Boon N (2011) A time-course analysis of four full-scale anaerobic digesters in relation to the dynamics of change of their microbial communities. Water Sci Technol 63:769–775 doi: 10.2166/wst.2011.307

Rivière D, Desvignes V, Pelletier E, Chaussonnerie S, Guermazi S, Weissenbach J, Li T, Camacho P, Sghir A (2009) Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME J 3:700–714 doi: 10.1038/ismej.2009.2

Rubin BER, Gibbons SM, Kennedy S, Hampton-Marcell J, Owens S, et al. (2013) Investigating the Impact of Storage Conditions on Microbial Community Composition in Soil Samples. PLOS ONE 8(7): e70460. doi: 10.1371/journal.pone.0070460.

Lettinga G (1995). Anaerobic digestion and wastewater treatment systems. Antonie Van Leeuwenhoek 67:3–28. PMID: 7741528

Sekiguchi Y, Kamagata Y, Nakamura K, Ohashi A, Harada H (1999) Fluorescence *in situ* hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Appl Environ Microbiol 65:1280–1288 PMID: 10049894

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001) *In situ* detection, isolation, and physiological properties of a thin filamentous microorganism abundant in methanogenic granular sludges: a novel isolate affiliated with a clone cluster, the green non-sulfur bacteria, subdivision I. Appl Environ Microbiol 67:5740–5749 doi: 10.1128/AEM.67.12.5740-5749.2001

Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H, Kamagata Y (2003) *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain Bacteria at the subphylum level. Int J Syst Evol Microbiol 53:1843–1851 doi: 10.1099/ijs.0.02699-0

Sekiguchi Y, and Kamagata Y (2004) Microbial community structure and functions in methane fermentation technology for wastewater treatment. Biosci Horizons 361-384

Seshadri R, Adrian L, Fouts DE, Eisen JA, Phillippy AM, Methe BA, Ward NL, Nelson WC, Deboy Robert T, Khouri HM, Kolonay JF, Dodson RJ, Daugherty SC, Brinkac LM, Sullivan SA, Madupu R, Nelson KE, Kang KH, Impraim M, Tran K, Robinson JM, Forberger HA, Fraser CM, Zinder SH, Heidelberg JF (2005) Genome sequence of the PCE-dechlorinating bacterium *Dehalococcoides ethenogenes*. Science 307:105–108 doi: 10.1126/science.1102226

Shannon CE, Weaver W (1963) The mathematical theory of communication. University of Illinois Press, Urbana.

Shu D, He Y, Yue H, Wang Q (2015) Microbial structures and community functions of anaerobic sludge in six full-scale wastewater treatment plants as revealed by 454 high-throughput pyrosequencing. Bioresour

Technol 186:163–172 doi: 10.1016/j.biortech.2015.03.072

Speece RE (1996) Anaerobic biotechnology for industrial wastewater. Archaea Press.

Sun L, Toyonaga M, Ohashi A, Matsuura N, Tourlousse DM, Meng XY, Tamaki H, Hanada S, Cruz R, Yamaguchi T, Sekiguchi T (2016) Isolation and characterization of *Flexilinea flocculi* gen. nov., sp. nov., a filamentous anaerobic bacterium belonging to the class *Anaerolineae* in the phylum *Chloroflexi*. Int J Syst Evol Microbiol 66:988–996 doi: 10.1099/ ijsem.0.000822

Tamura K, Nei M, and Kumar S (2004) Prospects for inferring very large phylogenies by using the neighbor-joining method. Proceedings of the National Academy of Sciences (USA) 101:11030-11035. doi: 10.1073/pnas.0404206101

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680 PMCID: PMC308517

Uemura S, Harada H (1993) Microbial characteristics of methanogenic sludge consortia developed in thermophilic UASB reactors. Appl Microbiol Biotechnol 39:654–660 doi: 10.1007/BF00205070

Van Lier JB (2008) High-rate anaerobic wastewater treatment: Diversifying from end-of-the-pipe treatment to resource-oriented conversion techniques. Water Sci Technol 57:1137–1148 doi: 10.2166/wst.2008.040

Verstraete W, Morgan-Sagastume F, Aiyuk S, Waweru M, Rabaey K, Lissens G (2005) Anaerobic digestion as a core technology in sustainable management of organic matter. Water Sci and Technol 52(1-2): 59-66 PMID: 16180409

Wiegant WM (1987) The ‘spaghetti theory’ on anaerobic sludge formation, or the inevitability of granulation. Granul. Anaerob Sludge 146-152

Xu M, Chen X, Dai R, Xiang X, Li G, Cao J, Xue Z, Shang K (2017) Performance and bacterial community change during the start-up period of a novel anaerobic bioreactor inoculated with long-time storage anaerobic granular sludge. J Environ Eng (United States) 143:1–8 doi: 10.1061/(ASCE)EE.1943-7870.0001222

Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005) Diversity, localization, and physiological properties of filamentous microbes belonging to *Chloroflexi* subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl Environ Microbiol 71:7493–7503 doi: 10.1128/AEM.71.11.7493-7503.2005

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp nov and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. Int J Syst Evol Microbiol 56:1331–1340. doi: 10.1099/ijs.0.64169-

Yamada T, Imachi H, Ohashi A, Harada H, Hanada S, Kamagata Y, Sekiguchi Y (2007a) *Bellilinea caldifistulae* gen. nov., sp. nov and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum *Chloroflexi* isolated from methanogenic propionate-degrading consortia. Int J Syst Evol Microbiol 57:2299–2306 doi: 10.1099/ijs.0.65098-0

Yamada T, Yamauchi T, Shiraishi K, Hugenholtz P, Ohashi A, Harada H, Kamagata Y, Nakamura K, Sekiguchi Y (2007b) Characterization of filamentous bacteria, belonging to candidate phylum KSB3, that are

associated with bulking in methanogenic granular sludges. ISME J 1:246–255 doi: 10.1038/ismej.2007.28

Yamada T, Sekiguchi Y (2009) Cultivation of uncultured *Chloroflexi* subphyla: significance and ecophysiology of formerly uncultured *Chloroflexi* “subphylum I” with natural and biotechnological relevance. Microbes Environ 24:205–216 doi: 10.1264/jsme2.ME09151S

Yamada T, Kikuchi K, Yamauchi T, Shiraishi K, Ito T, Okabe S, Hiraishi A, Ohashi A, Harada H, Kamagata Y, Nakamura K, Sekiguchi Y (2011) Ecophysiology of uncultured filamentous anaerobes belonging to the phylum KSB3 that cause bulking in methanogenic granular sludge. Appl Environ Microbiol 77:2081–2087 doi: 10.1128/AEM.02475-10

Yan J, Rash BA, Rainey FA, Moe WM (2009). Isolation of novel bacteria within the *Chloroflexi* capable of reductive dechlorination of 1,2,3-trichloropropane. Environ Microbiol 11:833–843 PMID: 19396942.

Yoon DN, Park SJ, Kim SJ, Jeon CO, Chae JC, Rhee SK (2010) Isolation, characterization, and abundance of filamentous members of *Caldilineae* in activated sludge. J Microbiol 48:275–283 doi: 10.1007/s12275-010-9366-8

Zhang, T., Shao, M.-F., Ye, L., 2012. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. ISME J 6:1137-1147 doi: 10.1038/ismej.2011.188

2.6 Material suplementario

Table S2.1. Reactor’s operation data at sampling days. The temperature for reactors COA and COB were measured inside reactors. The temperature for the other reactors was determined as the average ambient temperature for the sampling day according to INUMET (https://www.inumet.gub.uy). The data presented for reactors COA, COB and MO were supplied by the personal in charge of the wastewater treatment plant. The data presented for the reactors SR and AL were retrieved from the work presented by Fernández, 2016 and López et al., 2017.

Reactor	Day of sampling	VSS ($g\ l^{-1}$)	Organic load (kg COD. $(m^3.d)^{-1}$	HRT (day)	CH_4 (%)	pH	Temperature (°C)
	14	6.6	3.34	8.6	75	nd	19
	35	19.7	2.45	8.6	66	nd	23
	58	9.9	1.11	8.6	65	nd	25
	73 to 255	Non feeding due to stop in the production of the industry					
AL	256	20.1	1.21	16.3	70	nd	18
	266	8.6	5.39	7.8	72	7.6	18
	291	24.2	1.76	18.5	70	7.5	12
	305	21.8	2.61	12.5	69	7.4	19
	320	23.5	2.89	15.5	65	7.3	13

	343	36.0	4.03	9.2	74	7.3	18
	354	14.6	5.17	7.7	74	7.1	16
	10		1.07			6.41	25
	38		1.52			6.75	23
	72		0.60			6.39	21
	106		0.88			6.33	24
	141		0.87			5.59	25
COA	168		nd			6.20	28
	257		nd			7.55	31
	286		nd			7.10	30
	320		nd			6.42	27
	341		nd			5.90	26
	375		nd			5.62	24
	10		1.07			6.50	25
	38		1.52			6.78	23
	72		0.60			6.33	20
	106		0.88			6.29	24
	141		0.87			5.61	25
	168		nd			6.23	28
COB	257		nd			7.23	31
	258	Addition of enzymes to favor the fat degradation only to COB					
	286		nd			7.12	30
	320		nd			6.48	27
	341		nd			5.92	25
	375		nd			5.60	24
	14	30.1	1.1	0.16	nd	7.72	21
	35	24.5	1.6	0.06	72	7.10	10
	49	26.0	4.0	0.04	67	nd	15
	63	22.1	3.5	0.04	72	nd	17
	77	21.2	nd	0.04	76	nd	6
SR	91	pH increase due to a NaOH spill					
	102	25.5	2.3	0.05	64	nd	14
	153	29.6	1.8	0.05	78	nd	12
	183	51.4	2.5	0.05	58	7.61	18
	197	29.7	2.0	0.04	nd	nd	25
	259	30.5	0.9	0.03	nd	nd	22

MO	1	7.7	6.7	13
	96	3.4	6.6	20
	205	nd	nd	20
	339	nd	nd	15
	371	1.0	7.7	15
	542	10.9	7.0	22

nd= non data available for this sampling day

Table S2.2. Coverage percentage of the different Chloroflexi classes using the primer pairs of each sequencing platform. Coverage was performed using the test prime tool in the Silva web page (www.arb-silva.de).

Class	Primer coverage (%)	
	454 FLX[a]	**Ion Torrent PGM[b]**
Chloroflexi	74	84
Anaerolineae	67	85
Ardenticatenia	40	88
C10-SB1A	100	100
Caldilineae	91	92
Chloroflexia	92	84
Dehalococcoidia	77	78
JG30-KF-CM66	94	91
Ktedonobacteria	97	97
SBR2076	0	0
SJA-15	89	94
SJA/68	0	0
TK10	86	70
Thermomicrobia	90	88

[a] Primer pair: 515F GTGCCAGCMGCCGCGGTAA/806R GGACTACVSGGGTATCTAAT

[b] Primer pair: 520F AYTGGGYDTAAAGNG/802R TACNNGGGTATCTAATCC

Table S2.3. Chemical composition of wastewaters used to feed the industrial reactors studied in this work. The two dairy industries studied did not include cheese whey in their wastewater. For the sugar cane vinasse the data were obtained from López et al., 2017, for the malting wastewater the data were supplied by the personal of the industry, for the dairy industry wastewater the data were obtained from Passeggi et al., 2009. The main chemical

compounds of the wastewaters were retrieved from the following bibliography: Parnaudeau et al., 2006 (vinasse), Herrera, 2003 (malting wastewater) and Demirel et al., 2005 (dairy wastewater).

	Sugar cane vinasse	Malting wastewater	Dairy wastewater
COD (g L-1)	36.00	2.53	2.52
TKN (g L-1)	0.8	nd	36
$N\text{-}NH_4^+$ (mg L-1)	nd	21.0	nd
P total (mg L-1)	nd	25.3	nd
K (g L-1)	4.5	nd	nd
SO_4^{-2} (g L-1)	1.5	nd	nd
NO_3^- (mg L-1)	nd	7.9	nd
Fats (mg L-1)	nd	nd	495
pH	4.5	7.1	7.4
Main chemical compounds	Carbohydrates (glucose, saccharose, fructose, trehalose), organic acids (acetate, oxalate, aconitate), alcohols and polyols (glycerol, arabitol, aconitol)	Proteins (legumin, fibrin), carbohydrates (maltose, arabinose)	Proteins (casein) carbohydrates (lactose), lipids (triacylglycerides)

nd= non data available *The pH was adjusted to 7 before feeding the reactor.

References

Demirel, B., Yenigun, O., & Onay, T. T. (2005). Anaerobic treatment of dairy wastewaters: A review. Process Biochemistry, 40(8), 2583–2595

Herrera A., "límites máximos permisibles reglamentados de Contaminantes en efluentes liquidos industriales y Alternativa de control del análisis in situ. (2003) Thesis Universidad Nacional De Ingenieria Facultad De Ingenleria Quimica Y Manufacturera, Perú.

López I, Borzacconi L, Passeggi M (2017) Anaerobic treatment of sugar cane vinasse: Treatability and real-scale operation. J Chem Technol Biotechnol doi: 10.1002/jctb.5493

Parnaudeau, V., N. Condom, R. Oliver, P. Cazevieille, S. Recous (2008) Vinasse organic matter quality and mineralization potential, as influenced by raw material, fermentation and concentration processes. Bioresource Technology 99 (2008) 1553–1562

Passeggi M, López I, Borzacconi L (2009) Integrated anaerobic treatment of dairy industrial wastewater and sludge. Water Sci Technol 59:501–506 doi: 10.2166/wst.2009.010

CAPÍTULO 3

Meta-análisis: determinación de la ecología del filo Chloroflexi en reactores metanogénicos escala real

Trabajo publicado en Frontiers in Microbiology - DOI: 10.3389/fmicb.2020.603234

Title: Database mining to unravel the ecology of the phylum Chloroflexi in anaerobic full scale bioreactors

Patricia Bovio-Winkler[1] • Angela Cabezas[2] • Claudia Etchebehere[1]

[1]Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomic, Biological Research Institute "Clemente Estable", Avenida Italia 3318, CP: 11600, Montevideo, Uruguay.

[2]Instituto Tecnológico Regional Centro Sur, Universidad Tecnológica, Francisco Antonio Maciel s/n, CP: 97000, Durazno, Uruguay.

Keywords: Chloroflexi, full-scale methanogenic reactors, wastewater, filamentous bacteria, 16S rRNA amplicon sequence, meta-analysis.

* Correspondence to: Claudia Etchebehere

Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomics, Biological Research Institute "Clemente Estable", Av. Italia 3318. CP 11600. Montevideo, Uruguay.

Phone: (598) 24871616. Fax: (598) 24875461

cetchebehere@iibce.edu.uy

Abstract

Chloroflexi populations have been reported as one of the most predominant phyla in full-scale anaerobic wastewater treatment systems. However, their global diversity in these systems is still poorly described. The present meta-analysis focuses on analyzing the abundance and diversity of Chloroflexi in full-scale wastewater treatment anaerobic reactors (WTARs) and solid waste treatment anaerobic reactors (STARs). We built a diversity dataset with 16S ribosomal RNA gene sequence data from 62 full-scale methanogenic reactors studied worldwide. Microbial diversity analysis revealed that a cluster within the MiDAS phylotype T78 (Anaerolineae class) dominated WTARs and STARs with different core populations. According to the phylogenetic analysis, most of the sequences detected by high-throughput sequencing formed clusters with no cultured representatives. Anaerolineae class was more abundant WTAR (granular biomass) than in STARs (disperse biomass) suggesting

that Anaerolineae contribute to the granules structure due to their filamentous morphology.. Our results show that cross-study comparisons can be fruitfully used to understand diversity of Chloroflexi. However, more efforts are needed to standardized protocols because differences in experimental procedures may produce variation that outweighs biological differences.

Key words: Chloroflexi, full-scale methanogenic reactors, filamentous bacteria, 16S rRNA amplicon sequence, meta-analysis.

3.1 Introduction

Anaerobic digestion is an efficient biological process widely applied to treat solid organic waste and wastewater, where the organic matter can be converted into a renewable energy source known as biogas. As it is a mixture containing mainly methane (CH4) and carbon dioxide (CO2), it can be used as a replacement for fossil fuels to generate heat or electricity (Angenent et al., 2004; Appels et al., 2011; Verstraete et al., 2005).

Chloroflexi has been reported as one of the most predominant phylum present in solid waste and wastewater treatment systems (Bovio et al., 2019; Petriglieri et al., 2018; Shu et al., 2015). In particular, the Anaerolineae class has been identified as one of the core microbial populations in full-scale anaerobic reactors (Bovio et al., 2019; Nelson et al., 2011; St-Pierre and Wright, 2014). From 12 isolated strains within the Anaerolineae class, five have been isolated from wastewater treatment anaerobic reactors (WTARs) and one from solid-waste treatment anaerobic reactors (STARs) (Sekiguchi et al., 2003; Sun et al., 2016; Yamada et al., 2007a, 2006). All isolated species share similar phenotypic traits such as filamentous morphology, strict anaerobic growth, and the ability to ferment carbohydrates or amino acids. However, the ecological role of the Anaerolineae remains uncertain due to the scarcity of isolates and annotated genome sequences.

In recent years, efforts to assemble genomes from metagenomes have increased due to the difficulty to obtain Chloroflexi strains in pure culture. Genomes belonging to Anaerolineales (*Candidatus* Brevefilum fermentans CAMBI-1), Ardenticatenales (*Candidatus* Promineofilum) and Caldilineales (*Candidatus* Amarolinea aalborgensis) orders have been assembled from shotgun metagenomic sequencing data obtained from aerobic and anaerobic reactors (Andersen et al., 2018; McIlroy et al., 2017, 2016). These genomes showed similar potential metabolism compared to isolated members. It has been suggested that the common

prevalence of Anaerolineae class in WTARs and STARs might be due to their advantageous cellular adhesiveness, their potential as cellulose degraders and/or because of being anaerobic syntrophs (Xia et al., 2016). For example, some Anaerolineae species require syntrophic association with hydrogenotrophic methanogens for efficient growth (Sekiguchi et al., 2001a; Yamada et al., 2006, 2005a).

The two most common types of anaerobic digestion systems currently in use are up-flow anaerobic sludge blanket (UASB) reactors and continuously stirred tank reactors (CSTR). In UASB systems, and variations such as expanded granular sludge bed (EGSB) and internal circulation (IC) reactors, the active microbial biomass form compact anaerobic granules which settle against the hydraulic up-flow inside the reactor, preventing biomass washout (Skiadas et al., 2003). Contrastingly, in CSTR systems, generally used for solid waste anaerobic treatment, the active microbial biomass typically does not exhibit granulation but grow in suspension and is constantly removed from the system (Hofman-Bang et al., 2003; Klocke et al., 2007). It has been hypothesized that Chloroflexi are relevant for the granule skeleton formation in WTARs as they grow as filaments and therefore play an important role in sludge sedimentation (Yamada et al., 2005b). On the other hand, Chloroflexi has been reported to be occasionally involved in bulking episodes in WTARs, caused by their overgrowth, generating biomass washout (Borzacconi et al., 2008; Li et al., 2008; Sekiguchi et al., 2015a, 2001b; Yamada et al., 2007b). Meanwhile, it has been suggested that, in STARs some of these Chloroflexi derive and migrate from the aerobic activated sludge community when they are coupled together, although many were identified as being exclusive to the digester (Kirkegaard et al., 2017; Petriglieri et al., 2018).

The phylum Chloroflexi has been studied in STARs and WTARs mainly by molecular methods in separate studies. However, a comparison of Chloroflexi diversity and abundance between full scale STARs and WTARs has still not been investigated. During the last decade, the amount of studies using amplicon sequencing to analyze the microbial community in STARs and WTARs have been in constant growth, but the experimental approaches or data analysis differ widely. This strongly limits our ability to compare among studies and draw general conclusions regarding their diversity or the identification of important taxa. The recent ecosystem-specific Microbial Database for Activated Sludge (MiDAS) 16S rRNA gene amplicon sequencing-based survey, facilitates the understanding of wastewater treatment ecosystem diversity and function (Nierychlo et al., 2020). Data analysis appears to be an important area where further improvements and unification of experimental procedures are necessary. The main objective of this study was to answer the following questions: Is

there a particular Chloroflexi population predominant in all anaerobic reactors or are there several groups? Is there a selection of different groups between STARs and WTARs? To address these questions, we conducted a meta-analysis of publicly available microbial datasets generated by high-throughput sequencing in 62 full scale anaerobic reactors treating 29 different solid wastes or wastewaters.

3.2 Materials and methods

3.2.1 Data collection and composition

This meta-analysis includes sequences from published papers, all of which used bacterial 16S rRNA gene sequencing to survey the microbial community in full-scale anaerobic reactors. Seventeen studies were selected that comprised 62 methanogenic reactors: 27 reactors treating wastewater (WTAR) and 35 reactors treating solid waste (STAR). The configuration of the reactors, operational data, substrates, sequencing platforms and primer sets are summarized in Table 3.1 (Supplementary Table S3.1).

Table 3.1 Studies included in this meta-analysis and operational parameters

	Reactor configuration	Amount of reactors	Substrate	Temperature	Platform	Primers	Study
WTAR	UASB	4	Brewery	Mesophilic	Roche 454 FLX	27F -338R	Werner et al. 2011
	IC	1					
	EGSB	4					
	UBF	1	High-Strength Pharmaceutical	Mesophilic	Illumina MiSeq	8F -338R	Ma et al. 2017
	UASB	2	Pulp and paper	Mesophilic	Roche 454 GS-FLX	357F-926R	Shu et al. 2015
		1	Starch				
		1	Juice				
		2	Municipal and campus domestic sewage				
		1	Potato	NA	Illumina MiSeq	515F -806R	Zhu et al. 2016
		1	Poultry slaughterhouse	NA	Ion 318™	577F -924R	Delforno et al. 2016
		1	Vinasse	Mesophilic	Roche 454 FLX	515F -806R	Bovio et al. 2018
		3	Dairy				
		1	Malting		Ion Torrent PGM	520F-802R	
		1	Paper mill	NA	Illumina MiSeq	515F -806R	Sposob et al. 2018
	AnaEG	1	Cassava, potato, maize starch	NA	Illumina MiSeq	341F-805R	Qin et al. 2018
	IC	1	Ethanol processing wastewater	Mesophilic	Illumina MiSeq	341F-805R	Qin et al. 2019
	AnaEG	1					
STAR	-	1	Municipal solid wastes	Mesophilic	Roche 454	27F -518R	Cardinali-Rezende et al. 2016
	Plug-flow	2	Dairy cattle manure suppl. with cheese	Mesophilic	Roche 454	27F -519R	St-Pierre et al. 2013
			Cattle manure with ice cream waste				
	Complete mix	1	Cattle manure with oil waste				
	CSTR	1	Food waste-recycling wastewater	Thermophilic	Roche 454	787F-1492R	Lee et al. 2016
		2	Primary and biological sewage sludge	Mesophilic	Ion Torrent PGM	515F -806R	Hao et al. 2016
		3	Mixed sludge Thin stillage Agricultural waste	Mesophilic	Illumina MiSeq	515F -805R	Liu et al. 2017
		1	Maize silage	Mesophilic	Roche 454 GS	27F -519R	Lucas et al. 2015
		7	Cattle manure	Mesophilic	Illumina MiSeq	515F -806R	Li et al. 2015
	USR	1	Cattle manure				
		1	Swine manure				
	CSTR	11	Swine manure				
		4	Sewage sludge	Mesophilic	Roche 454 GS-FLX	787F-1492R	Shin et al. 2016

PCR-generated amplicons from bacterial 16S rRNA genes were sequenced using a diversity of primer sets targeting different regions of the 16S rRNA gene of Bacteria and Archaea (Table 3.1, Supplementary Table S3.1). The data sets were retrieved from the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI, USA) or were requested from the authors directly. The sample metadata for the processing of the raw data were extracted from published articles or provided by the authors. Some other operational data such as temperature, chemical oxygen demand removal (COD), pH, hydraulic retention time (HTR) and volatile fatty acid (VFA) were also collected in this study, but these data were only present for some reactors. Accordingly, we did not include those variables in this meta-analysis.

3.2.2 Data processing, clustering and taxonomic assignments

Raw reads (FASTQ) from each study were assessed for nucleotide quality and adaptors contamination with FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). The raw reads were processed through Trimmomatic v0.38 (Bolger et al., 2014) for removal of ambiguous reads, adapters, and low-quality sequences to obtain high quality reads (slidingwindow:4:25, minlen:200). Depending on the sequencing platform (single-end or paired-end) the pre-processed raw data was imported in QIIME2 v.2019.10 (Bolyen et al., 2019). For demultiplexing the raw data, qiime cutadapt demux-paired or demux-single plugins were used with default settings. All chimeras and borderline chimeras were detected and discarded using the qiime vsearch uchime-denovo plugin (vsearch v2.7.0) with defaults settings (Rognes et al., 2016).

3.2.3 Analysis strategy of the sequencing data

Given the heterogeneity of the selected hypervariable regions targeted within the 16S rRNA gene, we used Closed-reference OTU picking method. In this method, the input sequences are clustered against a reference database as a means of being able to combine non-overlapping tags (Rideout et al., 2014). Because input sequences are not compared directly to one another (*de novo* OTU picking method), but rather to an external reference, the input sequences don't need to overlap. This is essential for this analysis because we are performing a meta-analysis including sequences derived from different amplification products of the same marker gene, such as the V1-V2 and V3-V4 regions of the 16S rRNA (e.g., as in the meta-analysis performed in (Caporaso et al., 2010). As a result, for each short read we obtained a

representative reference sequence of approximately 1500 bp, which enabled the comparison of non-overlapping tags. For this reason, it was not possible to use the ASV method in which overlapping sequences are needed. OTUs were then picked at 97% sequence identity against the SILVA 138 reference database using the plugin of vsearch cluster-features-closed-reference plugin (--p-perc-identity 0.97) (Glöckner et al., 2017; Quast et al., 2012). Taxonomy was assigned with qiime feature-classifier-sklearn (--p-confidence 0.8) plugin against the MiDAS 3.6 database (Nierychlo et al., 2020). MiDAS 3.6 database is a comprehensive ecosystem-specific reference database for activated sludge and anaerobic digesters which provides a taxonomic classification at all ranks for all sequences (Dueholm et al., 2019; Nierychlo et al., 2020). The singletons were deleted and each count table was then normalized to the lowest number of reads (561 reads per sample). Even though this sequence number is low, we decided to prioritize the inclusion of all samples and all variable regions amplified by the most frequently used primers. After the normalization, 10 reactors presented 0 Chloroflexi sequences which were the reactors with lowest relative abundance (below 0.5%) of Chloroflexi in the microbial community.

3.2.4 Data analysis and visualization

As equal variances are a prerequisite of Permutational multivariate analysis of variance (PERMANOVA) (Daniels et al., 2011), previous to this analysis we performed a variance homogeneity check (betadisper function of the 'vegan' package) (Oksanen et al., 2013). In order to decipher the significant influence of explicative factors on the microbial diversity (e.g. reactor type, platform, target region and substrate, Table 3.1), a PERMANOVA with Adonis function, and principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity distance was calculated using R software version 3.5.1 (R Core Team 2013) in R Studio environment Version 1.0.153. Correction for multiple hypothesis testing was performed with the Benjamini–Hochberg (BH) procedure using the p.adjust function in R (Benjamini and Hochberg, 1995). Two-way analysis of variance (ANOVA) performed with R was used to test the effect of reactor type, platform, target region and substrate on the relative abundance of phylum Chloroflexi. The differences and correlations were considered significant at $p < 0.05$. In order to determine significant difference in the relative abundance of Chloroflexi between STAR and WTARs, the Shapiro-Wilk (to test normal distribution) and Kruskal-Wallis (to test equality of means) tests were applied in R. The correlation between the relative abundance of Euryarcheota and Chloroflexi was performed with the

Spearman's rho correlation test in R.

The diversity indices (Shannon H, Evenness and Chao-1), and the t-test and Kruskal-Wallis test (for normal and non-normal distribution, respectively) to test the equal means over diversity indexes was performed using past software (v4.02). Boxplots and PCoA plots were generated using R package ggplot2 (Wickham, 2017).

3.2.5 Phylogenetic tree

A phylogenetic tree was constructed to evaluate the phylogenetic diversity of the phylum Chloroflexi. It included the reference OTUs obtained with the closed-reference OTU picking method with a relative abundance greater than 5% in at least one sample, reference OTUs which were part of the core microbiome (defined as Chloroflexi members composed by OTUs shared by 50% of the STARs or WTARs) and 16S rRNA sequences from assembled genomes and cultured representatives from the phylum Chloroflexi. Thermotoga spp. sequence served as an outgroup for rooting the tree. The phylogenetic tree was performed using the maximum likelihood phylogenetic tree method with the generalized time reversible (GTR) (Nei and Kumar, 2000) substitution model and GAMMA distribution model using 1,000 bootstraps in MEGA X (Kumar et al., 2018). The resultant phylogenetic trees were visualized in iTOL (Letunic and Bork, 2019).

3.2.6 Primer coverage rate

The coverage rates for Chloroflexi populations of the different primer sets used can differ with the target location within the 16S rRNA gene. The Silva database Test Prime tool (Klindworth et al., 2013) was used to evaluate primer bias over Chloroflexi populations at different taxonomic levels with settings recommended in the Silva database Test Prime tutorial (1 mismatch and 5 bases).

3.3 Results

In the present study, high-throughput 16S rRNA sequence data from 17 different studies were compiled, including 62 full-scale anaerobic reactors that treat 29 different types of wastewaters or solid wastes. The data retrieved was generated using different primer pairs, sequencing platforms and sequencing depth which required bioinformatics tools to be adapted in order to be able to compare the data properly. Moreover, the primer coverage for the target group was also determined and analyzed. Finally, the diversity, phylogeny and

abundance of the Chloroflexi population across these methanogenic reactors, applying closed-reference method, were determined.

3.3.1 In silico analysis of the coverage rate of Chloroflexi primer pairs

As different primer pairs were used to generate sequences in the different investigations, we analyzed how this might affect the recovery of Chloroflexi populations, at different taxonomic levels, frequently found in methanogenic reactors. The *in silico* analysis performed showed different coverage of the phylum when different primer pairs are used. The primer sets directed to the region between V3 and V5 hypervariable regions (forty-seven reactors) showed the highest coverage for the Chloroflexi phylum (between 78% and 91%) (Table 3.2).

Table 3.2 *In silico* primer coverage prediction for the 10 primers sets used in the studies restricted to the phylum Chloroflexi in SILVA Test Prime analysis (SSU Ref 138 NR). Color red indicates higher primer coverage, while green color indicates lower primer coverage. Numbers indicate the primer coverage as a percentage of the total sequences of the data base detected by the primer set.

		Target region	V1-V2		V1-V3		V3-V5	V3-V4	V4-V5	V4		V5-V9
		Primers	27F-338R	8F-338R	27F-518R	27F-519R	357F-926R	341F-805R	577F-924R	515F-806R	520F-802R	787F-1492R
		References								Bovio et al. 2019		
										Sposob et al. 2018		
								Qin et al. 2019		Hao et al. 2016		
						Lucas et al. 2015		Qin et al. 2018		Zhu et al. 2016		Lee et al. 2016
			Werner et al. 2011	Ma et al. 2017	Cardinali-Rezende et al. 2016	St Pierre et al. 2013	Shu et al. 2015	Liu et al. 2017	Delforno et al. 2016	Li et al. 2015	Bovio et al. 2019	Shin et al. 2016
-	-	Chloroflexi	58	59	86	87	87	78	91	84	85	70
-	-	Anaerolineae	42	43	86	87	83	74	90	80	85	70
Anaerolineales	-	-	5	5	84	84	90	71	94	72	85	76
Anaerolineales	Anaerolineaceae	-	5	5	84	84	90	71	94	72	85	76
Anaerolineales	Anaerolineaceae	Anaerolinea	0	0	92	92	97	90	98	89	93	73
Anaerolineales	Anaerolineaceae	Bellilinea	0	0	75	75	92	100	92	92	100	88
Anaerolineales	Anaerolineaceae	Flexilinea	9	9	82	82	30	30	93	90	90	70
Anaerolineales	Anaerolineaceae	Leptolinea	2	2	81	81	95	90	99	29	95	84
Anaerolineales	Anaerolineaceae	Levilinea	17	17	100	100	82	82	94	94	94	78
Anaerolineales	Anaerolineaceae	Longilinea	0	0	86	88	93	93	95	85	96	81
Anaerolineales	Anaerolineaceae	Ornatilinea	13	13	88	88	88	84	92	96	92	75
Anaerolineales	Anaerolineaceae	Pelolinea	0	4	63	63	94	94	96	96	96	75
Anaerolineales	Anaerolineaceae	RBG-16-58-14	0	0	89	89	100	100	100	100	100	100
Anaerolineales	Anaerolineaceae	Thermanaerothrix	0	0	100	100	100	50	100	50	50	100
Anaerolineales	Anaerolineaceae	Thermomarinilinea	0	0	100	100	91	100	82	91	91	80
Ardenticatenales	-	-	35	36	91	92	46	42	93	84	87	66
Ardenticatenales	Ardenticatenaceae	-	100	100	100	100	92	80	96	83	86	61
Ardenticatenales	Ardenticatenaceae	Ardenticatena	100	100	100	100	100	100	100	100	100	67
Caldilineales	-	-	87	87	91	91	96	91	96	91	92	79
Caldilineales	Caldilineaceae	-	87	87	91	91	96	91	96	91	92	79
Caldilineales	Caldilineaceae	Caldilinea	80	80	100	100	88	75	100	88	88	83
Caldilineales	Caldilineaceae	Litorilinea	86	86	100	100	91	78	91	74	78	64
RBG-13-54-9	-	-	53	54	86	86	91	67	89	69	68	70
SBR1031	-	-	82	85	86	86	82	81	84	86	85	63
SJA-15	-	-	91	91	89	91	99	99	96	96	96	86
Thermoflexales	-	-	71	71	71	71	100	86	96	86	82	26
Thermoflexales	Thermoflexaceae	-	71	71	71	71	100	86	96	86	82	26
Thermoflexales	Thermoflexaceae	Thermoflexus	71	71	71	71	100	86	96	86	82	26
-	-	Dehalococcoidia	88	92	87	88	96	81	95	84	81	67
Dehalococcoidales	-	-	95	95	90	93	97	38	92	77	38	75
Dehalococcoidales	Dehalococcoidaceae	-	97	97	97	97	98	29	94	94	28	70

Meanwhile, the coverage for families commonly found in methanogenic reactor as Anaerolineaceae, Ardenticatenaceae and Caldilineaceae was between 71% and 100%. The primer sets that covered hypervariable regions V1-V2 and V5-V9 (fifteen reactors) were the least efficient covering the Chloroflexi phylum (58-70%). For Anaerolineaceae, Ardenticatenaceae and Caldilineaceae the coverage was between 61% and 100%, with the exception of V1-V2 region which showed a coverage of 5% for Anaerolineaceae (Table 3.2).

3.3.2 Microbial community structure in STARs and WTARs

In total, 1,082,904 high quality sequences from 62 reactors, were retrieved from public databases, ranging from 561 to 164,994 (median of 11299). All samples were randomly re-sampled at the same depth of 561 sequences before analyzing the data. To be able to compare sequences from different regions, closed-reference OTUs picking method was used. In this method, input sequences are aligned against a reference database. Because input sequences are not compared directly to one another, but rather to an external reference, they do not need to overlap. The closed-reference method allows performing a meta-analysis including sequences derived from different amplification products of the same marker gene. Clustering the pruned and combined data sets from all studies resulted in 13,165 no-singletons OTUs at a similarity threshold of 97%. Bacterial community diversity at phylum level was evaluated for all reactors. Proteobacteria, Chloroflexi, Bacteroidetes, and Firmicutes were the predominant phyla in WTAR and STAR (Figure 3.1). For WTARs the most abundant phylum was Proteobacteria, while Firmicutes dominated STARs. According to the PCoA analysis, the type of reactor (solid waste vs liquid wastewater treatment) influences the total bacterial community composition. PERMANOVA tests revealed that 15% of the variation of the total community was attributable to the type of reactor (PERMANOVA, $p < 0.001$) (Figure 3.2, Table 3.2).

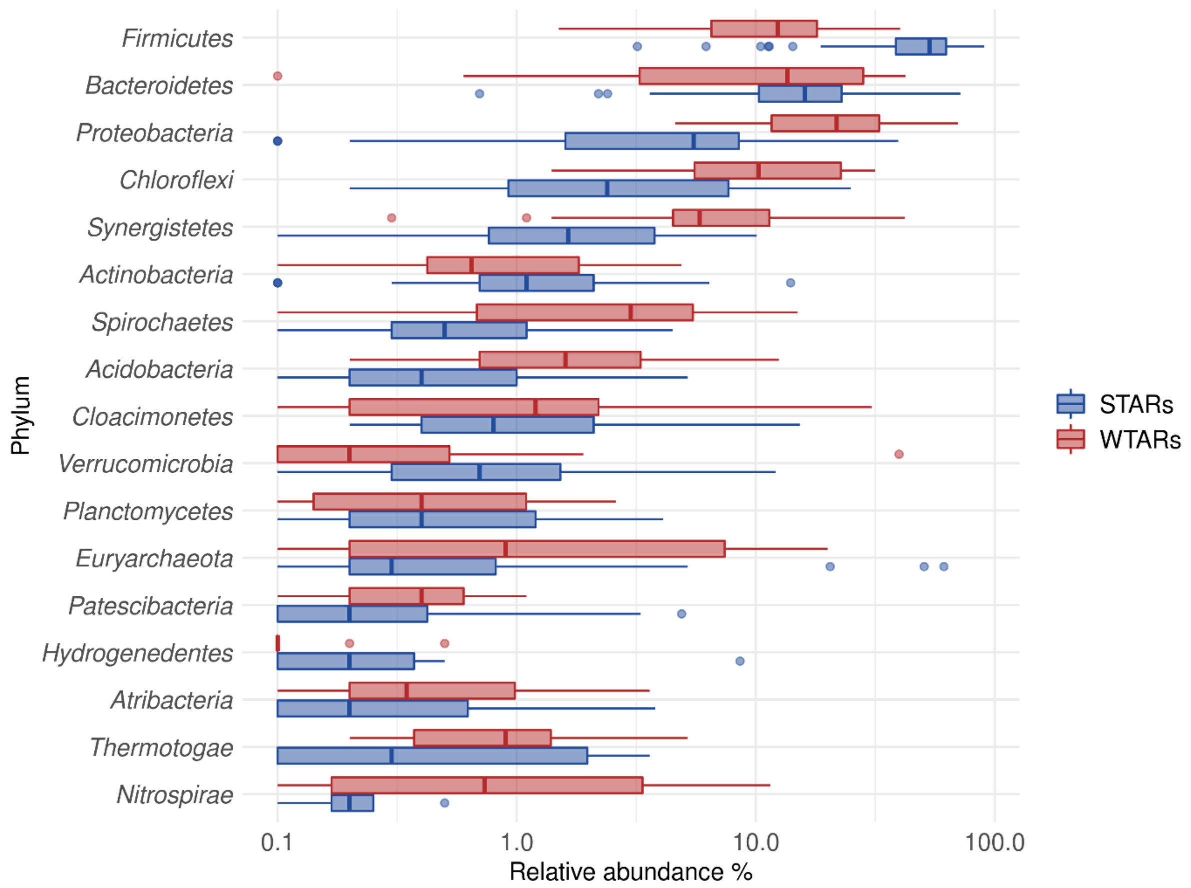

Figure 3.1 Boxplot of the total community for STARs and WTARs showing the phyla with a relative abundance higher than 2% in at least one reactor.

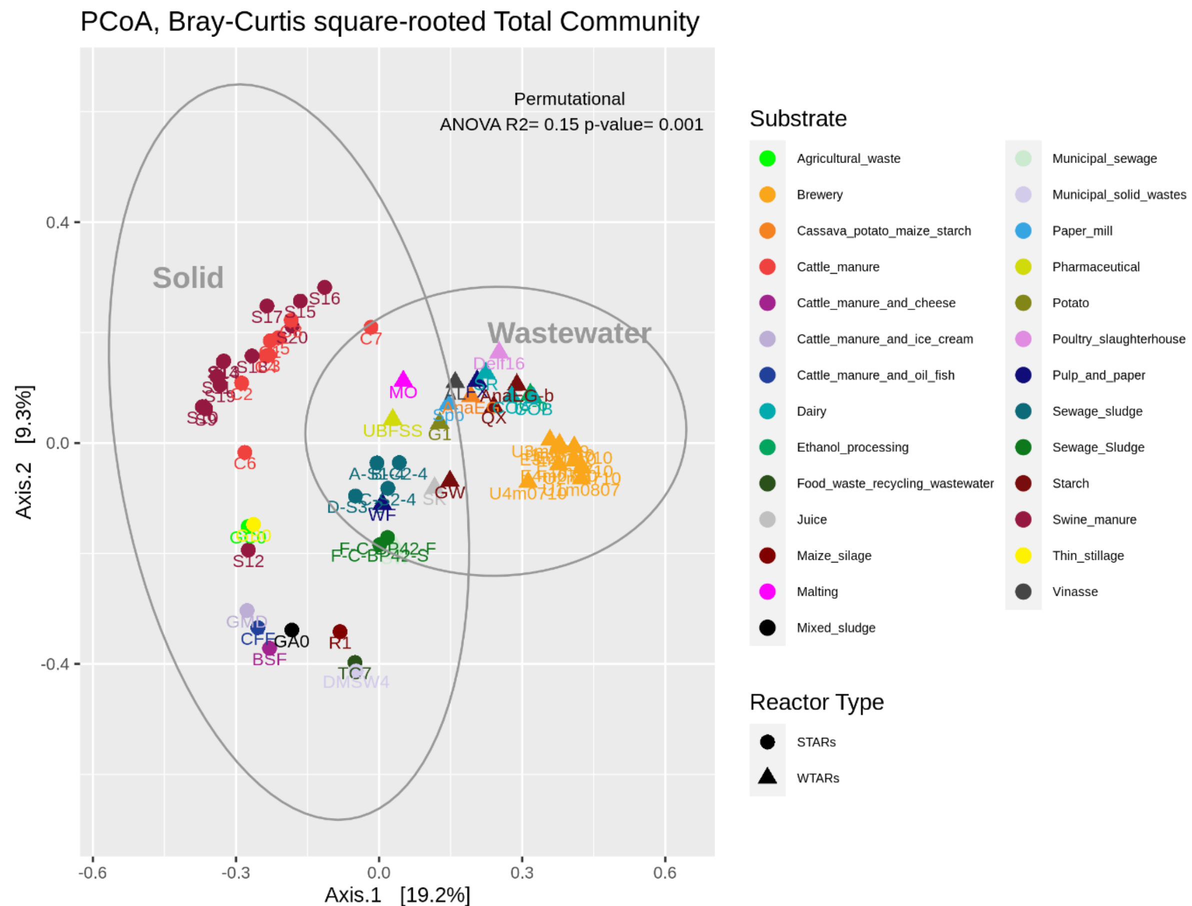

Figure 3.2 PCoA analysis and PERMANOVA test of the total community. The different colors indicate the different wastes or wastewaters used to fed the reactors. The PERMANOVA test results are shown at the top of the figure (see Table 3.3).

Regarding the abundance of the Chloroflexi phylum, we detected higher abundance in WTARs than in STARs. Moreover, Chloroflexi was detected in all reactors studied except for four STARs. In order to determine which variable significantly affected the relative abundance of Chloroflexi, two-way ANOVA analysis was applied. The variables tested were reactor type, primer pair target region, sequencing platform and substrate on the relative abundance of Chloroflexi. Our results revealed that the reactor type had the strongest impact on the abundance of Chloroflexi at phylum level (R^2=9.944, $p < 0.001$) followed by the target region (R^2=6.045, $p < 0.001$) (Supplementary Table S3.2).

Interestingly, the hypervariable region V3-V5 recovered the highest relative abundance of the phylum Chloroflexi (median value 7%) followed by V1-V2 (median value 6%) and V5-V9 (median value 0.5%) (Supplementary Figure S1). This is in accordance with the results obtained with the Silva test primer analysis (Supplementary Table S3.3).

3.3.3 Chloroflexi community composition

The classification of Chloroflexi sequences at the class level showed that Anaerolineae class was predominant in all reactors independently of the substrate, reactor type, 16S rRNA-targeted region and sequencing platform (Figure 3.3 (A)). The relative abundance of the Anaerolineae in the total community ranged between 1.4% and 31.1% (average: 13.2%) for WTARs and between 0.2% and 24.9% (average: 4.8%) for STARs. The difference in the median relative abundance of Anaerolineae was significantly different when comparing STARs and WTARs (Kruskal-Wallis test p-value < 0.001). Other minor classes detected were Chloroflexia and Dehalococcoidia present in some STARs and WTARs.

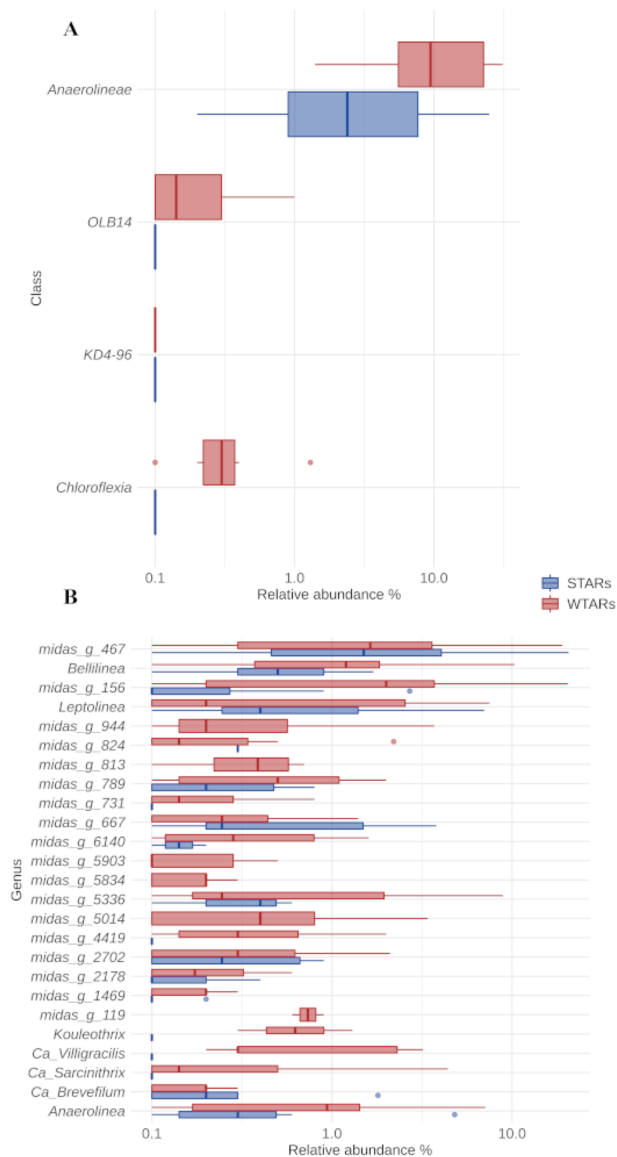

Figure 3.3 Boxplot of the Chloroflexi phylum for STARs and WTARs showing the relative abundance at: (A) class level (showing all the classes present in the systems); (B) genus level (showing the 25 most abundant genus which belong to the Anaerolinea class).

The diversity of the Chloroflexi community (Shannon index) was significantly higher in WTARs than in STARs (t-test, t=2.32, p=0.024) (Supplementary Table S3.4). The Chao1 and Evenness indices did not show significant differences (Kruskal-Wallis test, p> 0.05).

In this meta-analysis, most of the genera detected within the phylum Chloroflexi had no cultured representatives. The recent ecosystem-specific MiDAS 16S rRNA gene amplicon sequencing-based survey, facilitates the understanding of wastewater treatment ecosystem diversity and function (Nierychlo et al., 2019). The advantage of MiDAS 3.6 database is that it provides provisional alphanumeric names at genus and species levels allowing comparisons

between unclassified members from different investigations. We found that the four most abundant genera were the same in STARs and WTARs with changes in their relative abundances: midas_g_467, midas_g_156, *Bellilinea* and *Leptolinea* (Figure 3.3 (B)). While midas_g_467 was abundant in both systems, midas_g_156 was more abundant in W TARs. Despite that both belong to the T78 cluster and are therefore closely related.

PCoA analysis and PERMANOVA tests revealed that 33% of the variation within Chloroflexi populations was attributable to the primer target region (PERMANOVA, $p < 0.001$) (Figure 3.4, Table 3.3), while 10% of the variation was attributable to reactor type (Supplementary Figure S2), Table 3.3).

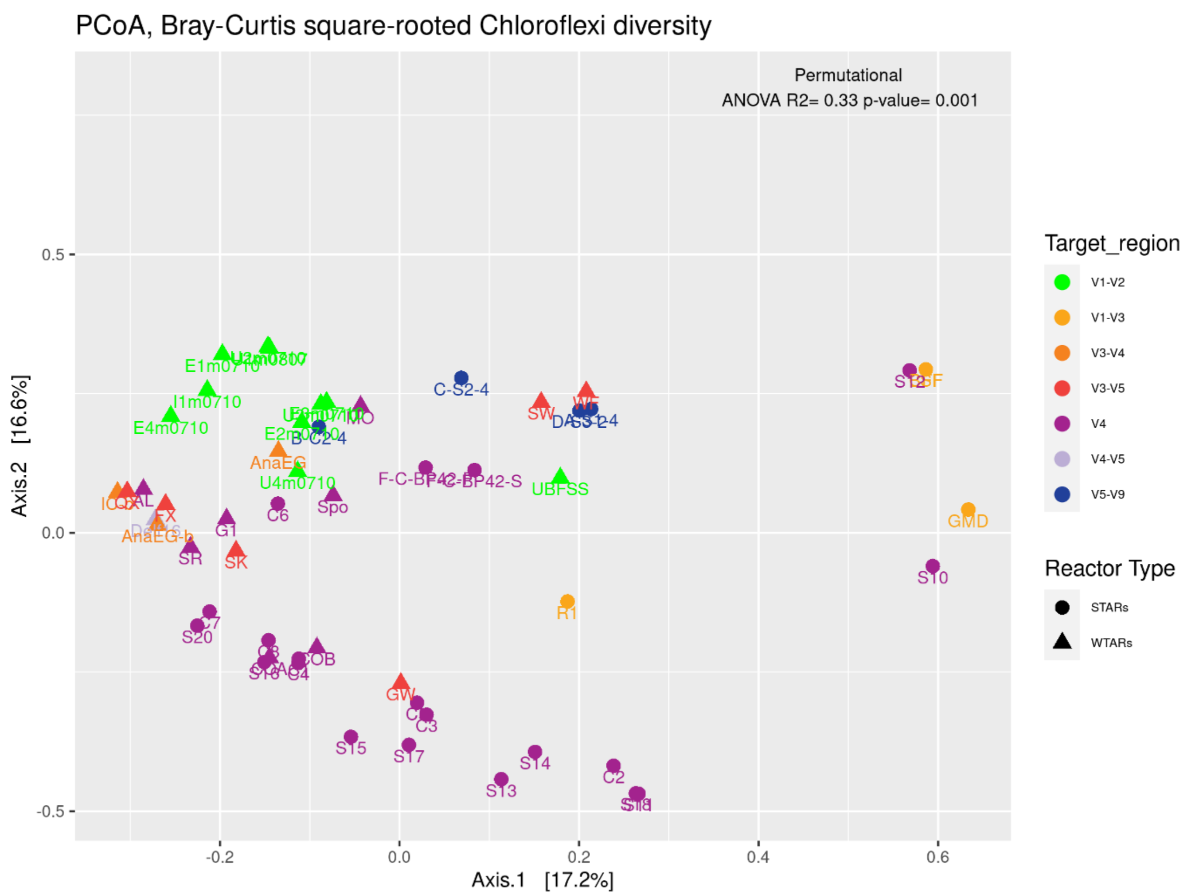

Figure 3.4 PCoA analysis and PERMANOVA test of the phylum Chloroflexi populations, the different colors indicate the different hypervariable regions. The PERMANOVA test results are shown at the top of the figure (see Table 3.3).

Table 3.3 PERMANOVA (adonis) and multivariate homogeneity of groups dispersions analysis (betadisper) results. Groups that satify the homogeneity group condition to perform PERMANOVA test are highlighted in grey. P values were corrected for multiple tests using

the Benjamini-Hochberg correction factor. Substrate refers to the feeding according to Table 3.1

	Total community					Chloroflexi population				
	betadisper (anova)		adonis		Benjamini Hochberg	betadisper (anova)		adonis		Benjamini Hochberg
	R^2	p-value	R^2	p-value	Adj. P-value	R^2	p-value	R^2	p-value	Adj. P-value
Reactor type	0.447	0.507	0.15	0.001	0.001	0.001	0.988	0.095	0.001	0.001
Platform	8.202	0.001	0.13	0.001	0.001	8.65	0.001	0.124	0.001	0.001
Target region	3.33	0.007	0.308	0.001	0.001	1.469	0.208	0.33	0.001	0.001
Substrate	7.297	0.001	0.733	0.001	0.001	4.697	0.025	0.654	0.001	0.001

When performing PERMANOVA test, only with reactors in which the V3-V5 hypervariable region was amplified to avoid the primers bias, 12% of the variation of Chloroflexi populations was attributable to the reactor type (Supplementary Table S3.2). STARs and WTARs shared 23 of the 25 most abundant OTUs. STARs were dominated by midas_s_6158 (midas_g_467), midas_s_1625 (midas_g_467), midas_s_1462 (midas_g_467), midas_s_3887 (Leptolinea) and midas_s_667 (midas_g_667) (Figure 3.5, Supplementary Table S3.5), while WTARs were dominated by midas_s_6158 (midas_g_467), midas_s_956 (midas_g_156), midas_s_6727 (Bellilinea), midas_s_3887 (Leptolinea) and midas_s_1462 (midas_g_467).
To determine if Chloroflexi populations were shared between reactors or if each reactor type had specific Chloroflexi populations we determined the minimum Chloroflexi core populations. Interestingly, we found that WTARs and STARs showed different Chloroflexi core populations as no OTU was shared by 50% of the reactors (Table 3.4).

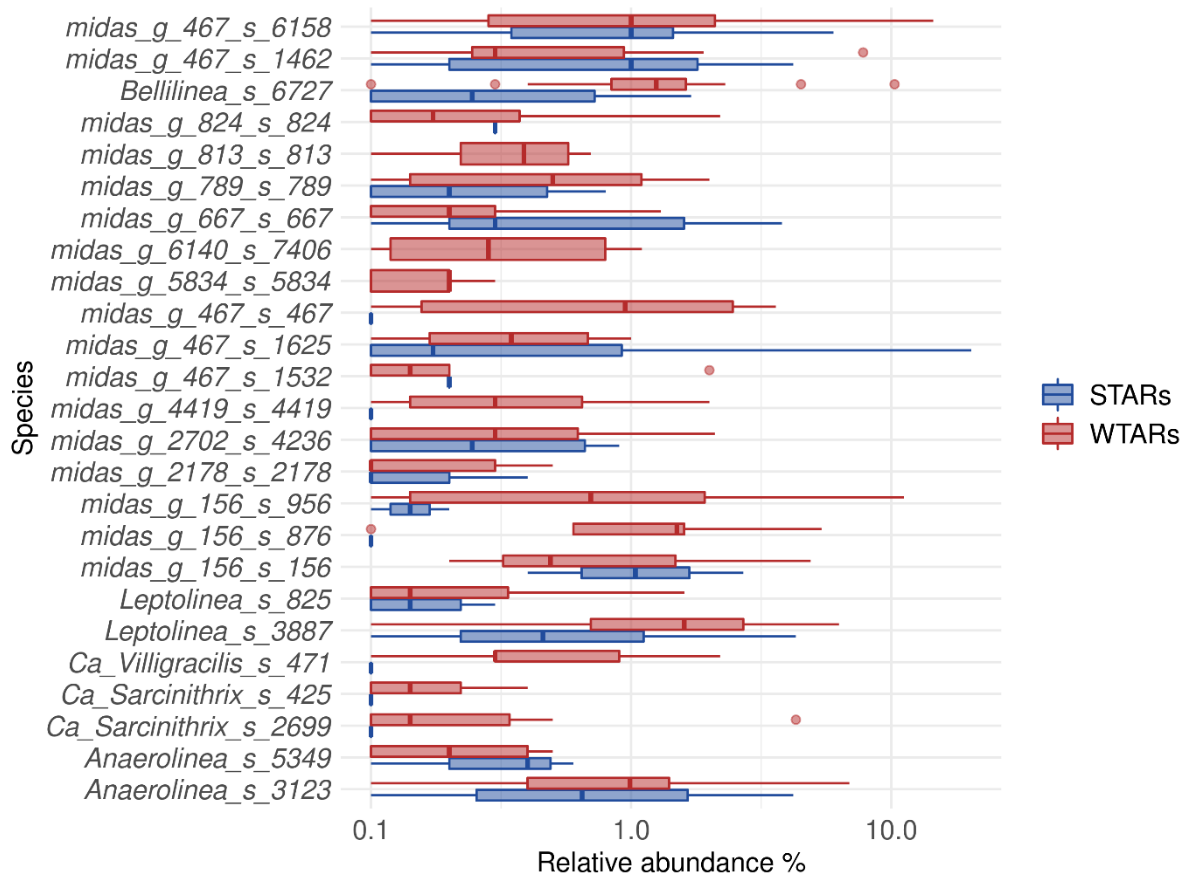

Figure 3.5 Boxplot of the Chloroflexi phylum for STARs and WTARs showing the relative abundance at species level (25 most abundant).

Table 3.4 Minimum Core microbiome obtained for STARs and WTARs. The accessing number of the sequences and the classification at the level of family, genera and species is shown. NA: not assigned

	Accession number	Family	Genera	Species
WTARs	CU922639	*Anaerolineaceae*	midas_g_156	NA
	EF515722			midas_s_956
	CU922098			
	CU926258		midas_g_467	NA
	CU918857			midas_s_467
	CU921338			midas_s_6158
	CU922464			midas_s_1462
	JQ996679		midas_g_2702	NA
	CU923106		NA	NA
	EF515566		NA	NA
	EF029839		NA	NA
	CU922316		*Leptolinea*	midas_s_3887
	CU921229			
	FQ660222			
	CU923750			
	JQ668574		*Bellilinea*	midas_s_6727
	EF515680			
	FPLM01002501	midas_f_7837	NA	NA
STARs	CU926258	*Anaerolineaceae*	midas_g_467	NA
	CU923011			midas_s_1462
	AB700375			midas_s_1532
	FJ645708			

Within each reactor type, cores were found. For WTARs the core was composed by 18 OTUs belonging to Anaerolineaceae (midas_g_156, midas_g_467, midas_g_2702, Leptolinea and Bellilinea genus) and midas_f_7837 (not assigned at genus level) families, representing at least 5 genera (Table 3.4). In STARs the core microbiome was composed by 4 OTUs belonging to the Anaerolineaceae family (midas_g_467). Only one OTU belonging to the Anaerolineaceae family (midas_g_467) was present in the core of both STARs and WTARs. It should be taken into account that these would be the most abundant OTUs. The low sequencing depth in some samples does not allow to detect OTUs with low abundances.

On the other hand, the correlation between Chloroflexi and Euryarchaeota (methanogens) was evaluated using the relative abundances obtained from Li et al. 2015, which included 18 reactors (515F-806R). Interestingly, we found a significant correlation between these two groups (rho: 0.60, $p < 0.01$) (Supplementary Table S3.6).

3.3.4 Phylogenetic analyses of the sequences affiliated to the phylum Chloroflexi

In order to determine the phylogenetic position of the most abundant Chloroflexi microorganisms, OTUs with a relative abundance greater than 5 % in at least one reactor and OTUs belonging to the minimum core microbiome (defined as Chloroflexi members composed by OTUs shared between 50% of the STARs or WTARs) were used to construct a phylogenetic tree. Most of the selected OTUs sequences were positioned within the order Anaerolineales (52 OTUs) followed by midas_o_1 (5 OTUs). Half of the sequences positioned within the Anaerolineales order clustered with sequences from the T78 genus (22/52 OTUs), including most OTUs sequences from the core microbiome (Figure 3.6).

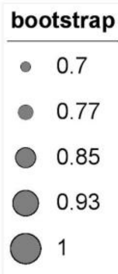

Figure 3.6 Phylogenetic tree of the phylum Chloroflexi inferred from the 16S rRNA gene sequences. The tree was reconstructed using the ML method and the GTR model. ML bootstrap values greater than or equal to 70% are shown at each node. Bar, 0.01 substitutions per site. The reference OTUs with a relative abundance greater than 5% in at least one reactor are represented by their GenBank accession numbers. Only OTUs from the core microbiome are indicated in color: from STARs are in blue color, from WTARs are in red color. Thermotoga spp. sequence served as an outgroup for rooting tree.

The genus T78 was recently splitted into midas_g_156 and midas_g_467 (Nierychlo et al., 2020). Within the Anaerolineales order, the 16S ribosomal RNA gene retrieved from the assembled genome of *Brevefilum fermentans* CAMBI-1 (LT859958) was phylogenetically closely related to the T78 cluster. A few OTUs sequences within Anaerolineales were closely related to 16S rRNA gene sequences from isolated strains or assembled genomes members (Figure 3.6). The sequences affiliated to the order midas_o_1 had no closely related sequences.

3.4 Discussion

Chloroflexi bacteria are widely distributed on Earth and particularly abundant in engineered systems such as solid waste and wastewater treatment plants. With the aim of better describing the abundance and diversity of this phylum, we performed a meta-analysis of the publicly available diversity datasets (i.e. 16S rRNA metabarcoding) in 62 full scale anaerobic reactors.

3.4.1 Anaerolineae class dominate the Chloroflexi microbial communities in the anaerobic reactors studied

Despite the bias generated from using different hypervariable regions we found that the Anaerolineae class largely dominated in both STARs and WTARs which was consistent with previous DNA based surveys (Bovio et al., 2019; Nelson et al., 2011; Rivière et al., 2009). The dominance of Anaerolineae in both types of systems might be explained through several hypotheses. First, most of the isolated strains of Anaerolineae present the potential to degrade cellulose, carbohydrates and/or proteins anaerobically playing an important role as primary and secondary fermenters, having relevance in the bottlenecking hydrolysis step of anaerobic digestion (Narihiro et al., 2015a; Sekiguchi et al., 2003, 2001a; Yamada et al., 2007a, 2006, 2005b). Thus, the fact that these systems in general treat different fermentable compounds might be favorable for the growth of Anaerolineae.

A second explanation to the predominance of Anaerolineae could be that their growth is significantly stimulated when co-cultivated with methanogens (Narihiro et al., 2015a; Sekiguchi et al., 2003, 2001a; Yamada et al., 2007a, 2006, 2005b) indicating some potential for synergistic relationships. If the Anaerolineae group was a major donor of acetate or hydrogen to methanogens (McIlroy et al., 2017; Narihiro et al., 2015b; Zamorano-López et

al., 2020), the association in a "spaghetti-like" structure could promote the metabolites transfer flux between both groups, resulting in high methane production rates. In addition, the adhesiveness properties of Methanosaeta (Angenent et al., 2004; Li et al., 2015; Wiegant, 1986; Zheng et al., 2006) and members of Anaerolineae (Xia et al., 2016) could facilitate the metabolites transfer flux improving its dominance in these systems.

Interestingly, Anaerolineae had higher relative abundance in WTARs, which generally have granular biomass, than in STARs, in which the microbial biomass typically does not exhibit granulation but has instead dispersed biomass growth (Hofman-Bang et al., 2003; Klocke et al., 2007). In a previous study, we also determined that UASB reactors with granular biomass presented higher abundance of Anaerolineae than start-up UASB reactors with disperse biomass (Bovio et al., 2019). This might be explained by the fact that the filamentous morphology of Anaerolineae contribute to the granule formation in WTARs (Satoh et al., 2007; Sekiguchi et al., 2001a; Yamada et al., 2005b) presenting more abundance in reactors with granules than in those with dispersed biomass.

Another explanation could be the fact that the higher flow velocity used in WTARs compared to STARs, may select organisms which can adhere to each other to form well-settling granular sludge, favoring the growth of Anaerolineae in these systems over other microorganisms (Xia et al., 2016).

3.4.2 Four genera within Anaerolineae predominate in the full scale anaerobic reactors studied

We found that the predominant genera and species in STARs and WTARs were the same, but changes in their relative abundances occurred. Four genera were predominant, two of them were closely related to previously described genera (*Bellilinea* and *Leptolinea*) and two of them did not harbor any cultured member and were named by the MIDAS database as midas_g_467 and midas_g_156. As for the isolated members of Anaerolineae, species belonging to the genus *Leptolinea* and *Bellilinea* present filamentous morphology and ferment carbohydrates and proteins (Yamada et al., 2007a, 2006). However, midas_g_467 and midas_g_156 previously named genus T78 in MiDAS database version 2 (Nierychlo et al., 2020), does not have representative cultured members. Despite this, some possible physiological features and ecological roles have been suggested for these genera in anaerobic reactors, which could explain the predominance of genus T78 (Jiang et al., 2020; Kirkegaard et al., 2017; Petriglieri et al., 2018; Zhu et al., 2017). Some of these previous reports suggest

that their filamentous morphology and particular position within flocs or granules, might indicate that they play a role in maintaining floc structure (Petriglieri et al., 2018) as well as the granular structure (Zhu et al., 2017). It has also been reported that genus T78 could play an important role in hydrolysis and fermentation steps in the anaerobic digestion process (Jiang et al., 2019; Yuan et al., 2015). T78 also could metabolize alcohols and carbohydrates through syntrophic interactions (Praveckova et al., 2016). In the literature we found that two species belonging to the midas_g_467 correlated positively with different operational parameters: being total ammonia nitrogen, biogas yield, temperature and organic loading rate (OLR) (Jiang et al., 2020). *Brevefilum fermentans* CAMBI-1 (LT859958) was the most closely related metagenome-assembled genome of the T78 cluster. The annotation of this genome suggested that members of the phylotype are important fermenters in mesophilic anaerobic digestion systems (McIlroy et al., 2017). Also, they are co-localized with the filamentous Archaea Methanosaeta spp. which suggests a potential undetermined synergistic relationship (McIlroy et al., 2017). Both, the potential fermenting metabolism as well as the potential syntrophism with archaea could explain the dominance of the T78 genus in anaerobic reactors. Our correlation analysis between Euryarchaeota and Chloroflexi supports this hypothesis. Hence, our main hypothesis about the predominance of the same four genera in both systems, relies on their capacity to undergo fermentative and hydrolytic metabolism. Further studies are needed to elucidate the function of these uncultured bacteria, which seem to have different responses to different conditions.

On the other hand, STARs and WTARs showed different core microbiomes of Chloroflexi members. In STARs the Chloroflexi core was less diverse comprising only midas_g_467 genus (4 OTUs) while in WTARs it was composed by at least midas_g_156, midas_g_467, midas_g_2702, Leptolinea and Bellilinea genus (18 OTUs). This result was in accordance with the differences observed for the Shannon diversity index. The existence of a core microbiome based on the type of reactor (STARs/WTARs) despite of the heterogeneity of wastewater/solids and environmental factors (e.g., pH, temperature) may suggest that this core could be functionally important in each system. Specific core of microorganism between STARS and WTARs has been detected previously (Calusinska et al., 2018) but regarding Chloroflexi, no previous reports were found. A possible explanation for the different Chloroflexi cores found could be related to the biomass structure found in each type of system. Further exploration of this hypothesis is needed which would contribute to a better understanding of the role of Chloroflexi in the structure of granules and flocs and their metabolic role. Systematic studies examining multiple anaerobic reactor designs with greater

depth of coverage and using platforms that generate long-reads or primer-free alternatives should help further define the 'core microbiome' of Chloroflexi populations in methanogenic reactors. However, it is important to keep in mind that further recovery of isolated members or interpretative genomes is required to disclose the ecological importance of Anaerolineae as a core population in the anaerobic digestion process.

3.4.3 Challenges and limitations of comparing sequences retrieved from databases generated with different primer sets and sequencing platforms

One of the bottlenecks in comparing microbial profiles is that different studies use different 16S rRNA gene regions generating an important bias. In the present study, we found an underestimation of Chloroflexi populations using primers targeting regions V1-V2 and V5-V9. An example was the underestimation of Anaerolineae class when V1-V2 regions were used recovering only a 5% of the total population. Meanwhile, primers targeting the V3-V5 region in general better reflected the abundances at all taxonomic levels. Previous studies investigating the influence of different hypervariable regions of 16S rRNA gene over the total community in activated sludge identified no perfect region concerning abundant functional genera, except for some regions with less bias, i.e. V1 and V2 (Cai et al., 2013; Guo and Zhang, 2013). However, (Albertsen et al., 2015) recommended primers targeting the V1-V3 regions which better reflected the abundances of Chloroflexi in activated sludge. More studies are needed to identify the best taxonomic profiling effectiveness between different hypervariable regions of the 16S rRNA gene over Chloroflexi populations in anaerobic reactors. As a general rule, drawing conclusions based only on one sequencing region should be avoided due to the potential false negative results, additionally amplicon sequencing of the 16S rRNA gene often limits resolution at genus level. We need primer-free alternatives to get the entire picture of the microbial diversity in anaerobic reactors (Karst et al., 2018). Using the full-length sequence has the potential to become a tool for more precise microbial community profiling that better allows global comparisons of microbiome studies and should potentially increase the accuracy and the resolution of closely related taxa. We highlight the effort to build the MiDAS reference database for microbes in wastewater treatment systems since it allows comparing uncultured microorganisms in different studies (Nierychlo et al., 2020).

Regarding the bias of different sequencing platforms, alpha diversity is significantly affected by both sequence length and depth (Singh et al., 2015). Platforms such as pyrosequencing

which were discontinued, had less sequencing depth. To be able to compare samples with different sequencing depths, it is necessary to perform normalization at the lowest value of reads per sample. This leads to a loss of information in the less dominant groups which is an important limitation that has to be taken into account. This problem becomes relevant when comparing data from studies performed in the past (using outdated sequencing platforms) with data obtained more recently where the technology allows a much deeper sequencing. Moreover, it has been reported that Ion Torrent as well as Pyrosequencing platforms had comparatively higher error rates than Illumina platform (Salipante et al., 2014; Siqueira et al., 2012). Although it has been reported that while primer choice considerably influences quantitative abundance estimations, the sequencing platform has relatively minor effects when matched primers are used (Singh et al., 2015). Singh et al. reported that beta diversity metrics are surprisingly robust to both primer and sequencing platform biases (Singh et al., 2015).

New developments in single-cell genomics and metagenomics have in recent years provided new insights into the ecology and evolution of many novel uncultured microorganisms (Albertsen et al., 2013; Dam et al., 2020; Sekiguchi et al., 2015b). The genomes have enabled the construction of metabolic models that attempt to explain the physiology of these organisms in detail. The genome-based models form the basis for more extensive investigations, such as *in situ* single-cell characterization, metatranscriptomics and proteomics (Koch et al., 2014).

3.4.4 Importance of standardized metadata repositories for full scale methanogenic reactors

A second issue on performing a meta-analysis is the lack of standardized information on key operational parameters. It is important to have more reactor performance data associated with each reactor, such as compositions of influent waste/wastewater stream, COD loading/removal rate, pH and temperature. These parameters should be fully used to infer potential roles of lineages of interest and should also be included. In our analysis it was not possible to correlate the microbial community with operational parameters data because in several articles this metadata was incomplete. For this reason, it would be essential to supply the operational parameters together with the raw data in each new study. In this sense, the creation of metadata repositories specific for anaerobic reactors including the 16S rRNA gene datasets must be a great advance. As far as we know there are some initiatives in other

ecosystems as the Microbial Antarctic Resource Systems (MARS) (http://mars.biodiversity.aq/) in Antarctic ecosystems.

3.5 Conclusions and future directions

From this meta-analysis, we managed to identify that four genera (midas_g_467, midas_g_156, Bellilinea and Leptolinea) belonging to Anaerolineae class dominated WTARs and STARs with different core populations. All the species observed had no cultured representatives, limiting our knowledge. From the few isolated species belonged from the genera detected, we could hypothesize that the growth of Anaerolineae could be favored by fermentable substrates and by the syntrophic association with methanogenic archaea. In addition, Anaerolineae was more abundant in reactors with granular biomass (WTARs), than in those in which the microbial biomass typically does not exhibit granulation but have instead dispersed biomass growth (STARs), suggesting that Anaerolineae members play an important role in the granule structure due to their filamentous morphology and/or adhesiveness properties. Despite the extensive research that has been done on microbial communities in anaerobic reactors there is no consensus about experimental protocols, bioinformatics analyses or operational data provided, which makes it difficult to perform global comparisons of microbiome studies. We think that more efforts are needed in each study to provide information on key operational parameters associated with each reactor, since it is fundamental to infer potential roles of lineages of interest. Cross-study comparisons can be fruitfully used to understand the complexity of the anaerobic digestion process. The rapidly advancing fields of metagenomics and metatranscriptomics will provide unique insights into the patterns of microbial activity in anaerobic digestion systems, even for species which have not yet been isolated in pure culture.

Acknowledgments

The authors express their gratitude to Uruguayan National Agency for Innovation and Research (ANII, POS_NAC_2015_1_109786) and the Postgraduate Academic Commission (CAP) for providing financial support to Patricia Bovio Winkler. CE and AC are fund by SNI-ANII. This work would not have been possible without the generosity and collaborative spirit of the many primary authors of the data sets included in the meta-analysis.

Author Contributions

CE and AC planned the investigation. PBW performed the bioinformatics analysis. PBW, AC and CE discussed results and wrote the manuscript. All authors read and approved the final manuscript.

Conflict of interest

The authors declare no conflict of interest.

3.6 References

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., Nielsen, P.H., 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat. Biotechnol. 31, 533–538. https://doi.org/10.1038/nbt.2579

Albertsen, M., Karst, S.M., Ziegler, A.S., Kirkegaard, R.H., 2015. Back to Basics – The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities 1–15. https://doi.org/10.1371/journal.pone.0132783

Andersen, M.H., McIlroy, S.J., Nierychlo, M., Nielsen, P.H., Albertsen, M., 2018. Genomic insights into *Candidatus* Amarolinea aalborgensis gen. nov., sp. nov., associated with settleability problems in wastewater treatment plants. Syst. Appl. Microbiol. 1–8. https://doi.org/10.1016/j.syapm.2018.08.001

Angenent, L.T., Karim, K., Al-Dahhan, M.H., Wrenn, B.A., Domíguez-Espinosa, R., 2004. Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends Biotechnol. 22, 477–485. https://doi.org/10.1016/j.tibtech.2004.07.001

Appels, L., Lauwers, J., Degrve, J., Helsen, L., Lievens, B., Willems, K., Van Impe, J., Dewil, R., 2011. Anaerobic digestion in global bio-energy production: Potential and research challenges. Renew. Sustain. Energy Rev. https://doi.org/10.1016/j.rser.2011.07.121

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bolyen, E., Rideout, J.R., Caporaso, J.G., 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat. Biotechnol. 37, 850–852. https://doi.org/10.1038/s41587-019-0190-3

Borzacconi, L., López, I., Passeggi, M., Etchebehere, C., Barcia, R., 2008. Sludge deterioration in a full scale UASB reactor after a pH drop working under low loading conditions. Water Sci. Technol. 57, 797–802. https://doi.org/10.2166/wst.2008.071

Bovio, P., Cabezas, A., Etchebehere, C., 2019. Preliminary analysis of Chloroflexi populations in full-scale UASB methanogenic reactors. J. Appl. Microbiol. 126, 667–683. https://doi.org/10.1111/jam.14115

Cai, L., Ye, L., Tong, A.H.Y., Lok, S., Zhang, T., 2013. Biased Diversity Metrics Revealed by Bacterial 16S Pyrotags Derived from Different Primer Sets. PLoS One 8, e53649. https://doi.org/10.1371/journal.pone.0053649

Calusinska, M., Goux, X., Fossépré, M., Muller, E.E.L., Wilmes, P., Delfosse, P., 2018. A year of monitoring

20 mesophilic full-scale bioreactors reveals the existence of stable but different core microbiomes in bio-waste and wastewater anaerobic digestion systems. Biotechnol. Biofuels 11, 1–19. https://doi.org/10.1186/s13068-018-1195-8

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., Mcdonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high- throughput community sequencing data. Nat. Methods 7, 335–336. https://doi.org/10.1038/nmeth0510-335

Dam, H., John, V., Sobol, M.S., Cabezas, A., Kaster, A.-K., 2020. Targeted Cell Sorting Combined with Single Cell Genomics reveals Microbial Dark Matter overlooked by Metagenomics. Front. Microbiol.

Daniels, C., Zeifman, A., Heym, K., Ritchie, K., Watson, C., Berzins, I., Breitbart, M., 2011. Spatial heterogeneity of bacterial communities in the mucus of Montastraea annularis. Mar. Ecol. Prog. Ser. 426, 29–40. https://doi.org/10.3354/meps09024

Delforno, T.P., Okada, D.Y., Faria, C. V., Varesche, M.B.A., 2016. Evaluation of anionic surfactant removal in anaerobic reactor with Fe(III) supplementation. J. Environ. Manage. 183, 687–693. https://doi.org/10.1016/j.jenvman.2016.09.026

Dueholm, M.S., Andersen, K.S., Petriglieri, F., McIlroy, S.J., Nierychlo, M., Petersen, J.F., Kristensen, J.M., Yashiro, E., Karst, S.M., Albertsen, M., Nielsen, P.H., 2019. Comprehensive ecosystem-specific 16S rRNA gene databases with automated taxonomy assignment (AutoTax) provide species-level resolution in microbial ecology. bioRxiv 672873. https://doi.org/10.1101/672873

Glöckner, F.O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G., Yarza, P., Peplies, J., Westram, R., Ludwig, W., 2017. 25 years of serving the community with ribosomal RNA gene reference databases and tools. J. Biotechnol. 261, 169–176. https://doi.org/10.1016/j.jbiotec.2017.06.1198

Guo, F., Zhang, T., 2013. Biases during DNA extraction of activated sludge samples revealed by high throughput sequencing. Appl. Microbiol. Biotechnol. 97, 4607–4616. https://doi.org/10.1007/s00253-012-4244-4

Hammer, Ø., Harper, D., Ryan, P., 2001. Past: Paleontological statistics software package for education and data analysis. Paleontol. Electron. 4, 1–9.

Hao, L., Bize, A., Conteau, D., Chapleur, O., Courtois, S., Kroff, P., Desmond-Le Quéméner, E., Bouchez, T., Mazeas, L., 2016. New insights into the key microbial phylotypes of anaerobic sludge digesters under heavy duty conditions. Water Res. https://doi.org/10.1016/j.watres.2016.06.014

Hofman-Bang, J., Zheng, D., Westermann, P., Ahring, B.K., Raskin, L., 2003. Molecular ecology of anaerobic reactor systems. Adv. Biochem. Eng. Biotechnol. https://doi.org/10.1007/3-540-45839-5_5

Jiang, C., Peces, M., Andersen, M.H., Kucheryavskiy, S., Nierychlo, M., Yashiro, E., Andersen, K.S., Kirkegaard, R.H., Hao, L., Høgh, J., Hansen, A.A., Dueholm, M.S., Nielsen, P.H., 2020. Characterizing the growing microorganisms at species level in 46 anaerobic digesters at Danish wastewater treatment plants : A six-year survey on microbiome structure and key drivers 1–27. https://doi.org/10.1101/2020.06.07.138891

Jiang, Y., Dennehy, C., Lawlor, P.G., Hu, Z., McCabe, M., Cormican, P., Zhan, X., Gardiner, G.E., 2019.

Exploring the roles of and interactions among microbes in dry co-digestion of food waste and pig manure using high-throughput 16S rRNA gene amplicon sequencing 06 Biological Sciences 0605 Microbiology 09 Engineering 0907 Environmental Engineering. Biotechnol. Biofuels 12. https://doi.org/10.1186/s13068-018-1344-0

Karst, S.M., Dueholm, M.S., McIlroy, S.J., Kirkegaard, R.H., Nielsen, P.H., Albertsen, M., 2018. Retrieval of a million high-quality, full-length microbial 16S and 18S rRNA gene sequences without primer bias. Nat. Biotechnol. 36, 190–195. https://doi.org/10.1038/nbt.4045

Kirkegaard, R.H., McIlroy, S.J., Kristensen, J.M., Nierychlo, M., Karst, S.M., Dueholm, M.S., Albertsen, M., Nielsen, P.H., 2017. The impact of immigration on microbial community composition in full-scale anaerobic digesters. Sci. Rep. 7, 1–11. https://doi.org/10.1038/s41598-017-09303-0

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41, e1. https://doi.org/10.1093/nar/gks808

Klocke, M., Mähnert, P., Mundt, K., Souidi, K., Linke, B., 2007. Microbial community analysis of a biogas-producing completely stirred tank reactor fed continuously with fodder beet silage as mono-substrate. Syst. Appl. Microbiol. 30, 139–151. https://doi.org/10.1016/j.syapm.2006.03.007

Koch, H., Galushko, A., Albertsen, M., Schintlmeister, A., Gruber-Dorninger, C., Lücker, S., Pelletier, E., Le Paslier, D., Spieck, E., Richter, A., Nielsen, P.H., Wagner, M., Daims, H., 2014. Microbial metabolism: Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. Science (80-.). 345, 1052–1054. https://doi.org/10.1126/science.1256985

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Lee, J., Han, G., Shin, S.G., Koo, T., Cho, K., Kim, W., Hwang, S., 2016. Seasonal monitoring of bacteria and archaea in a full-scale thermophilic anaerobic digester treating food waste-recycling wastewater: correlations between microbial community characteristics and process variables. Chem. Eng. J. 300, 291–299. https://doi.org/10.1016/j.cej.2016.04.097

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47, W256–W259. https://doi.org/10.1093/nar/gkz239

Li, J., Hu, B., Zheng, P., Qaisar, M., Mei, L., 2008. Filamentous granular sludge bulking in a laboratory scale UASB reactor. Bioresour. Technol. 99, 3431–3438. https://doi.org/10.1016/j.biortech.2007.08.005

Li, L., Zheng, M., Ma, H., Gong, S., Ai, G., Liu, X., Li, J., Wang, K., Dong, X., 2015. Significant performance enhancement of a UASB reactor by using acyl homoserine lactones to facilitate the long filaments of Methanosaeta harundinacea 6Ac. Appl. Microbiol. Biotechnol. 99, 6471–6480. https://doi.org/10.1007/s00253-015-6478-4

Liu, T., Sun, L., Müller, B., Schnürer, A., 2017. Importance of inoculum source and initial community structure for biogas production from agricultural substrates. Bioresour. Technol. 245, 768–777. https://doi.org/10.1016/j.biortech.2017.08.213

Ma, H., Zhang, L., Ye, L., Ding, L., Ren, H., Hu, H., 2017. Determination and Variation of Core Bacterial Community in a Two-Stage Full-Scale Anaerobic Reactor Treating High-Strength Pharmaceutical

Wastewater. J. Microbiol. Biotechnol. 27, 1808–1819. https://doi.org/10.4014/jmb.1707.07027

McIlroy, S.J., Karst, S.M., Nierychlo, M., Dueholm, M.S., Albertsen, M., Kirkegaard, R.H., Seviour, R.J., Nielsen, P.H., 2016. Genomic and *in situ* investigations of the novel uncultured Chloroflexi associated with 0092 morphotype filamentous bulking in activated sludge. Isme J (In press), 1–12. https://doi.org/10.1038/ismej.2016.14

McIlroy, S.J., Kirkegaard, R.H., Dueholm, M.S., Fernando, E., Karst, S.M., Albertsen, M., Nielsen, P.H., 2017. Culture-independent analyses reveal novel anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. Front. Microbiol. 8. https://doi.org/10.3389/fmicb.2017.01134

Narihiro, T., Nobu, M.K., Kim, N.K., Kamagata, Y., Liu, W.T., 2015a. The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. Environ. Microbiol. 17, 1707–1720. https://doi.org/10.1111/1462-2920.12616

Narihiro, T., Nobu, M.K., Kim, N.K., Kamagata, Y., Liu, W.T., 2015b. The nexus of syntrophy-associated microbiota in anaerobic digestion revealed by long-term enrichment and community survey. Environ. Microbiol. 17, 1707–1720. https://doi.org/10.1111/1462-2920.12616

Nei, M., Kumar, S., 2000. Molecular evolution and phylogenetics. Oxford University Press.

Nelson, M.C., Morrison, M., Yu, Z., 2011. A meta-analysis of the microbial diversity observed in anaerobic digesters. Bioresour. Technol. 102, 3730–3739. https://doi.org/10.1016/j.biortech.2010.11.119

Nierychlo, M., Andersen, K., Nielsen, P.H., 2019. Species-level microbiome composition of activated sludge - introducing the MiDAS 3 ecosystem-specific reference database and taxonomy.

Nierychlo, M., Andersen, K.S., Xu, Y., Green, N., Jiang, C., Albertsen, M., Dueholm, M.S., Nielsen, P.H., 2020. MiDAS 3: An ecosystem-specific reference database, taxonomy and knowledge platform for activated sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge. Water Res. 115955. https://doi.org/10.1016/j.watres.2020.115955

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Steven, M.H.H., Wagner, H., 2013. Vegan: Community Ecology Package. R Package Version. 2.0-10 [WWW Document]. URL https://www.researchgate.net/publication/258996451_Vegan_Community_Ecology_Package_R_Package_Version_20-10 (accessed 8.7.20).

Petriglieri, F., Nierychlo, M., Nielsen, P.H., McIlroy, S.J., 2018. *In situ* visualisation of the abundant Chloroflexi populations in full-scale anaerobic digesters and the fate of immigrating species. PLoS One 13, 1–14. https://doi.org/10.1371/journal.pone.0206255

Praveckova, M., Brennerova, M. V., Holliger, C., De Alencastro, F., Rossi, P., 2016. Indirect evidence link PCB dehalogenation with geobacteraceae in anaerobic sediment-free microcosms. Front. Microbiol. 7, 933. https://doi.org/10.3389/fmicb.2016.00933

Qin, X., Wu, X., Li, L., Li, C., Zhang, Z., Zhang, X., 2018. The advanced anaerobic expanded granular sludge bed (AnaEG) possessed temporally and spatially stable treatment performance and microbial community in treating starch processing wastewater. Front. Microbiol. 9, 589. https://doi.org/10.3389/fmicb.2018.00589

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596. https://doi.org/10.1093/nar/gks1219

Rideout, J.R., McDonald, D., Navas-Molina, J.A., Walters, W.A., Huse, S.M., Gilbert, J.A., Clemente, J.C., Robbins-Pianka, A., Ursell, L.K., Gonzalez, A., Chase, J., Zhou, H.-W., Gibbons, S.M., He, Y., Knight, R., Caporaso, J.G., 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. PeerJ 2, e545. https://doi.org/10.7717/peerj.545

Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P., Sghir, A., 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. ISME J. 3, 700–714. https://doi.org/10.1038/ismej.2009.2

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584

Salipante, S.J., Kawashima, T., Rosenthal, C., Hoogestraat, D.R., Cummings, L.A., Sengupta, D.J., Harkins, T.T., Cookson, B.T., Hoffman, N.G., 2014. Performance Comparison of Illumina and Ion Torrent Next-Generation Sequencing Platforms for 16S rRNA-Based Bacterial Community Profiling. https://doi.org/10.1128/AEM.02206-14

Satoh, H., Miura, Y., Tsushima, I., Okabe, S., 2007. Layered structure of bacterial and archaeal communities and their *in situ* activities in anaerobic granules. Appl. Environ. Microbiol. 73, 7300–7307. https://doi.org/10.1128/AEM.01426-07

Sekiguchi, Y., Ohashi, A., Parks, D.H., Yamauchi, T., Tyson, G.W., Hugenholtz, P., 2015a. First genomic insights into members of a candidate bacterial phylum responsible for wastewater bulking. PeerJ 3, e740. https://doi.org/10.7717/peerj.740

Sekiguchi, Y., Ohashi, A., Parks, D.H., Yamauchi, T., Tyson, G.W., Hugenholtz, P., 2015b. First genomic insights into members of a candidate bacterial phylum responsible for wastewater bulking. PeerJ 3, e740. https://doi.org/10.7717/peerj.740

Sekiguchi, Y., Takahashi, H., Kamagata, Y., Ohashi, A., Harada, H., 2001a. *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl. Environ. Microbiol. 67, 5740–5749. https://doi.org/10.1128/AEM.67.12.5740-5749.2001

Sekiguchi, Y., Takahashi, H., Kamagata, Y., Ohashi, A., Harada, H., 2001b. *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl. Environ. Microbiol. 67, 5740–5749. https://doi.org/10.1128/AEM.67.12.5740-5749.2001

Sekiguchi, Y., Yamada, T., Hanada, S., Ohashi, A., Harada, H., Kamagata, Y., 2003. Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. Int. J. Syst. Evol. Microbiol. 53, 1843–1851. https://doi.org/10.1099/ijs.0.02699-0

Shin, S.G., Koo, T., Lee, J., Han, G., Cho, K., Kim, W., Hwang, S., 2016. Correlations between bacterial populations and process parameters in four full-scale anaerobic digesters treating sewage sludge. Bioresour. Technol. 214, 711–721. https://doi.org/10.1016/j.biortech.2016.05.021

Shu, D., He, Y., Yue, H., Wang, Q., 2015. Microbial structures and community functions of anaerobic sludge in six full-scale wastewater treatment plants as revealed by 454 high-throughput pyrosequencing. Bioresour. Technol. 186, 163–172. https://doi.org/10.1016/j.biortech.2015.03.072

Singh, K., He, S., Tremblay, J., Lee, J., Tringe, S.G., Kirton, E.S., Fern, A., Dangl, J.L., Chen, F., Woyke, T., 2015. Primer and platform effects on 16S rRNA tag sequencing. Front. Microbiol. 6, 1–15. https://doi.org/10.3389/fmicb.2015.00771

Siqueira, J.F., Fouad, A.F., Rôças, I.N., 2012. Pyrosequencing as a tool for better understanding of human microbiomes. J. Oral Microbiol. 4, 1–15. https://doi.org/10.3402/jom.v4i0.10743

Skiadas, I. V., Gavala, H.N., Schmidt, J.E., Ahring, B.K., 2003. Anaerobic granular sludge and biofilm reactors. Adv. Biochem. Eng. Biotechnol. https://doi.org/10.1007/3-540-45838-7_2

Sposob, M., Cydzik-Kwiatkowska, A., Bakke, R., Dinamarca, C., 2018. Temperature-induced changes in a microbial community under autotrophic denitrification with sulfide. Process Biochem. 69, 161–168. https://doi.org/10.1016/j.procbio.2018.03.006

St-Pierre, B., Wright, A.D.G., 2014. Comparative metagenomic analysis of bacterial populations in three full-scale mesophilic anaerobic manure digesters. Appl. Microbiol. Biotechnol. 98, 2709–2717. https://doi.org/10.1007/s00253-013-5220-3

Sun, L., Toyonaga, M., Ohashi, A., Matsuura, N., Tourlousse, D.M., Meng, X., Tamaki, H., Hanada, S., Cruz, R., Yamaguchi, T., Sekiguchi, Y., 2016. Isolation and characterization of Flexilinea flocculi gen . nov ., sp . nov ., a filamentous , anaerobic bacterium belonging to the class Anaerolineae in the phylum Chloroflexi. Int. J. Syst. Evol. Microbiol. 66, 988–996. https://doi.org/10.1099/ijsem.0.000822

Verstraete, W., Morgan-Sagastume, F., Aiyuk, S., Waweru, M., Rabaey, K., Lissens, G., 2005. Anaerobic digestion as a core technology in sustainable management of organic matter. Water Sci. Technol. 52, 59–66. https://doi.org/10.2166/wst.2005.0498

Werner, J.J., Knights, D., Garcia, M.L., Scalfone, N.B., Smith, S., Yarasheski, K., Cummings, T.A., Beers, A.R., Knight, R., Angenent, L.T., 2011. Bacterial community structures are unique and resilient in full-scale bioenergy systems. Proc. Natl. Acad. Sci. U. S. A. 108, 4158–63. https://doi.org/10.1073/pnas.1015676108

Wickham, H., 2017. ggplot2 - Elegant Graphics for Data Analysis | Hadley Wickham | Springer. https://doi.org/10.1007/978-0-387-98141-3

Wiegant, W.M., 1986. Thermophilic anaerobic digestion for waste and wastewater treatment.

Xia, Y., Wang, Yubo, Wang, Yi, Chin, F.Y.L., Zhang, T., 2016. Cellular adhesiveness and cellulolytic capacity in Anaerolineae revealed by omics-based genome interpretation. Biotechnol. Biofuels 9, 1–13. https://doi.org/10.1186/s13068-016-0524-z

Yamada, T., Imachi, H., Ohashi, A., Harada, H., Hanada, S., Kamagata, Y., Sekiguchi, Y., 2007a. Bellilinea caldifistulae gen. nov., sp. nov and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum Chloroflexi isolated from methanogenic propionate-degrading consortia. Int. J. Syst. Evol. Microbiol. 57, 2299–2306. https://doi.org/10.1099/ijs.0.65098-0

Yamada, T., Sekiguchi, Y., Hanada, S., Imachi, H., Ohashi, A., Harada, H., Kamagata, Y., 2006. Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the . Int. J. Syst. Evol. Microbiol. 56, 1331–1340. https://doi.org/10.1099/ijs.0.64169-0

Yamada, T., Sekiguchi, Y., Imachi, H., Kamagata, Y., Ohashi, A., Harada, H., 2005a. Diversity, localization,

and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl. Environ. Microbiol. 71, 7493–7503. https://doi.org/10.1128/AEM.71.11.7493-7503.2005

Yamada, T., Sekiguchi, Y., Imachi, H., Kamagata, Y., Ohashi, A., Harada, H., 2005b. Diversity, localization, and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl. Environ. Microbiol. 71, 7493–7503. https://doi.org/10.1128/AEM.71.11.7493-7503.2005

Yamada, T., Yamauchi, T., Shiraishi, K., Hugenholtz, P., Ohashi, A., Harada, H., Kamagata, Y., Nakamura, K., Sekiguchi, Y., 2007b. Characterization of filamentous bacteria, belonging to candidate phylum KSB3, that are associated with bulking in methanogenic granular sludges. ISME J. 1, 246–255. https://doi.org/10.1038/ismej.2007.28

Yuan, Y., Wang, S., Liu, Y., Li, B., Wang, B., Peng, Y., 2015. Long-term effect of pH on short-chain fatty acids accumulation and microbial community in sludge fermentation systems. Bioresour. Technol. 197, 56–63. https://doi.org/10.1016/j.biortech.2015.08.025

Zamorano-López, N., Borrás, L., Seco, A., Aguado, D., 2020. Unveiling microbial structures during raw microalgae digestion and co-digestion with primary sludge to produce biogas using semi-continuous AnMBR systems. Sci. Total Environ. 699, 134365. https://doi.org/10.1016/j.scitotenv.2019.134365

Zheng, D., Angenent, L.T., Raskin, L., 2006. Monitoring granule formation in anaerobic upflow bioreactors using oligonucleotide hybridization probes. Biotechnol. Bioeng. 94, 458–472. https://doi.org/10.1002/bit.20870

Zhu, X., Kougias, P.G., Treu, L., Campanaro, S., Angelidaki, I., 2017. Microbial community changes in methanogenic granules during the transition from mesophilic to thermophilic conditions. Appl. Microbiol. Biotechnol. 101, 1313–1322. https://doi.org/10.1007/s00253-016-8028-0

3.5 Material suplementario

Table S3.1 Studies included in this meta-analysis and operational parameters

	Reactor configuration	Sample name	Substrate	Geographical region	T(°C) avg	COD removal %	Platform	Primers	Amplicon length	Raw data sources	Study
	UASB	U1m0807			27	96.5 ± 0.3				SRR096048	
		U2m0710			34	96.0 ± 0.5				SRR095975	
		U3m0710			29	95.4 ± 1.2				SRR095970	
		U4m0710			33.7	94.1 ± 0.5		27F AGAGTTTGATCCTGGCTCAG		SRR095977	
	IC	I1m0710	Brewery	USA	32	90.2 ± 0.9	Roche 454 FLX	338R TGCTGCCTCCCGTAGGAGT	311	SRR095976	Werner et al. 2011
	EGSB	E1m0710			32.7	88 ± 2				SRR095972	
		E2m0710			31.5	86 ± 2				SRR095974	
		E3m0710			29	88.7 ± 1.0				SRR095971	
		E4m0710			31	94.3 ± 1.1				SRR095973	
	UBF	UBFSS	High-Strength Pharmaceutical	NA	35	60	Illumina MiSeq	8F AGAGTTTGATYMTGGCTCAG 338R TGCTGCCTCCCGTAGGAGT	330	PRJNA331231	Ma et al. 2017
		WF	Pulp and paper		22	68					
		FX			28	78					
		GW	Starch		30	83		357F CCTACGGGAGGCAGCAG			
		QX	Juice	Shaanxi (China)	26	92	Roche 454 GS-FLX	926R CCGTCAATTCMTTTRAGT	569	SRR1765717	Shu et al. 2015
WTAR		SK	Municipal and campus		24	84					
		SW	domestic sewage		25	89					
		G1	Potato	Netherlands	NA	NA	Illumina MiSeq	515F GTGCCAGCMGCCGCGGTAA 806R GGACTACVSGGGTATCTAAT	291	PRJNA327299	Zhu et al. 2016
	UASB	Delf16	Poultry slaughterhouse	Sao Paulo (Brazil)	NA	NA	Ion 318™	577F AYTGGGYDTAAAGNG 924R CCGTCAATTCMTTTRAGT	347	ERS1289856	Delforno et al. 2016
		AL	Vinasse		20	60-78					
		SR	Dairy		20	55-90	Roche 454 FLX	515F GTGCCAGCMGCCGCGGTAA 806R GGACTACVSGGGTATCTAAT	291		
		COA	Dairy	Uruguay	25	29-97				PRJNA453814	Bovio et al. 2018
		COB	Dairy		25	19-97					
		MO	Malting		20	35-80	Ion Torrent PGM	520F AYTGGGYDTAAAGNG 802R TACNNGGGTATCTAATCC	282		
		Spo	Paper mill	Halden (Norway)	NA	NA		515F GTGCCAGCMGCCGCGGTAA 806R GGACTACHVGGGTWTCTAAT	291	PRJNA398347	Sposob et al. 2018
	AnaEG	AnaEG	Cassava, potato, maize starch	Hangzhou (China)	NA	NA	Illumina MiSeq	341F CCTACGGGAGGCAGCAG		SRP111625	Qin et al. 2018
	IC	IC-b	Ethanol processing wastewater	Jilin (China)	33	NA		805R GACTACHVGGGTATCTAATCC	464	SRR7281077	Qin et al. 2019
	AnaEG	AnaEG-b			33					SRR7281078	
		DMSW4	Municipal solid wastes	Madrid (Spain)	35	NA		27F AGAGTTTGATCCTGGCTCAG 518R ATTACCGCGGCTGCTGG	491	SRR1928203	Cardinali-Rezende et al. 2016
	Plug-flow	BSF	Dairy cattle manure suppl. with cheese		37.8					SRR768352	
		GMD	Cattle manure with ice cream waste	Vermont (USA)	38.3	-	Roche 454	AGTTTGATCCTGGCTCAG	492	SRR768352	St-Pierre et al. 2013
	Complete mix	CFF	Cattle manure with oil waste		36.1			519R WTTACCGCGGCTGCTGG		SRR768352	
	CSTR	TC7	Food waste-recycling wastewater	Gwangju (South Korea)	58.5	79		787F ATTAGATACCCNGGT 1492R GNTACCTTGTTACGACTT	705	SRR3281818	Lee et al. 2016
	-	F-C-BP42-F	Primary and biological sewage sludge	-	34.6	96.54	Ion Torrent PGM	515F GTGCCAGCMGCCGCGGTAA	291	SRR3198570	Hao et al. 2016
	-	F-C-BP42-S		-	34.7	96.45		806R GGACTACHVGGGTWTCTAAT		SRR3198571	
	-	GA0	Mixed sludge		38	-		515F GTGCCAGCMGCCGCGGTAA		SRR6173222	
	-	GB0	Thin stillage	Sweden	38	-	Illumina MiSeq	805R GACTACHVGGGTATCTAATCC	290	SRR6173226	Liu et al. 2017
	-	GC0	Agricultural waste		38	-				SRR6173228	
		R1	Maize silage	Mecklenburg-Western Pomerania (Germany)	38	-	Roche 454 GS	27F AGAGTTTGATCCTGGCTCAG 519R WTTACCGCGGCTGCTGG	492	ERR638341	Lucas et al. 2015
		C1		Heilongjiang (China)	32.5	-				ERR579094	
STAR	CSTR	C2		Beijing (China)	35	-				ERR579095	
		C3		Jiangsu (China)	35	-				ERR579096	
		C4	Cattle manure		35	-				ERR579097	
		C5		Chongqing (China)	35	-				ERR579098	
	USR	C6			35	-				ERR579099	
	CSTR	C7		Guangxi (China)	35	-				ERR579100	
		C8			35	-				ERR579101	
	USR	S9		Beijing (China)	36.5	-				ERR579102	
		S10			33	-	Illumina MiSeq	515F GTGCCAGCMGCCGCGGTAA 806R GGACTACHVGGGTWTCTAAT	291	ERR579103	Li et al. 2015
		S11		Jiangsu (China)	35	-				ERR579104	
		S12			35	-				ERR579105	
		S13		Shanghai (China)	35	-				ERR579106	
		S14	Swine manure		35	-				ERR579107	
		S15		Zhejiang (China)	25	-				ERR579108	
		S16			25	-				ERR579109	
	CSTR	S17		Ningxia (China)	35	-				ERR579110	
		S18		Sichuan (China)	36	-				ERR579111	
		S19		Guangdong (China)	30	-				ERR579112	
		S20			30	-				ERR579113	
		A-S1-4		Seoul (South Korea)	35	45.3				SRR3281890	
		B-C2-4	Sewage sludge	Daegu (South Korea)	35	39.7	Roche 454 GS-FLX	787F ATTAGATACCCNGGT 1492R GNTACCTTGTTACGACTT	705	SRR3281863	Shin et al. 2016
		C-S2-4		Incheon (South Korea)	35	30.1				SRR3281959	
		D-S3-2		Asan city (South Korea)	35	24.3				SRR3282014	

NA, Not applicable. T (°C), Temperature. COD, Chemical oxygen demand.

Table S3.2 Relative abundance in percentage of *Chloroflexi* in STARs and WTARs

Target region	Primers	Chloroflexi (%)	Reactor		Reference
V1-V2	27F-338R	**6**	U1m0807	WTAR	Werner et al. 2011
		7	U2m0710		
		5	U3m0710		
		2	U4m0710		
		6	I1m0710		
		6	E1m0710		
		7	E2m0710		
		3	E3m0710		
		5	E4m0710		
	8F-338R	**9**	UBFSS		Ma et al. 2017
V3 -V5	357F-926R	**30**	FX		Shu et al. 2015
		15	GW		
		35	QX		
		16	SK		
		13	SW		
		4	WF		
V4	515F-806R	**31**	G1		Zhu et al. 2016
V4-V5	577F-924R	**24**	Delf16		Delforno et al. 2016
V4	520F-802R	**21**	AL		Bovio et al. 2019
		23	SR		
		33	COA		
		27	COB		
		3	MO		
V4	515F-806R	**6**	Spo		Sposoba et al. 2018
V3-V4	341F-805R	**22**	AnaEG		Qin et al. 2018
		13	IC-b		Qin et al. 2019
		23	AnaEG-b		
V1-V3	27F-518R	**0**	DMSW4	STAR	Cardinali-Rezende et al. 2016
V1-V3	27F-519R	**0**	BSF		St-Pierre et al. 2013
		2	GMD		
		19	CFF		
V5-V9	787F-1492R	**0**	TC7		Lee et al. 2016
V4	515F-806R	**4**	F-C-BP42-F		Hao et al. 2016
		3	F-C-BP42-S		
V3-V4	341F-805R	**0**	GA0		Liu et al. 2017
		0	GB0		
		0	GC0		
V1-V3	27F-519R	**7**	R1		Lucas et al. 2015
V4	515F-806R	**7**	C1		Li et al. 2015
		1	C2		
		12	C3		
		16	C4		
		5	C5		
		4	C6		
		25	C7		
		12	C8		
		0	S9		
		1	S10		
		1	S11		
		1	S12		
		2	S13		
		3	S14		
		14	S15		
		16	S16		
		6	S17		
		1	S18		
		0	S19		
		5	S20		
V5-V9	787F-1492R	**0**	A-S1-4		Shin et al. 2016
		0	B-C2-4		
		2	C-S2-4		
		1	D-S3-2		

Table S3.3. Twenty-five most abundant species of the phylum Chloroflexi

Accession number	Class	Order	Family	Genus	Species
CU922464	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_467	midas_s_1462
AB700375	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_467	midas_s_1532
CR933301	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_156	midas_s_156
CU923351	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_467	midas_s_1625
HQ183891	Anaerolineae	SBR1031	midas_f_1469	midas_g_2178	midas_s_2178
JQ180422	Anaerolineae	Caldilineales	Amarolineaceae	Ca_Sarcinithrix	midas_s_2699
EF515732	Anaerolineae	Anaerolineales	Anaerolineaceae	Anaerolinea	midas_s_3123
CU922316	Anaerolineae	Anaerolineales	Anaerolineaceae	Leptolinea	midas_s_3887
CU922268	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_2702	midas_s_4236
FQ659015	Anaerolineae	Caldilineales	Amarolineaceae	Ca_Sarcinithrix	midas_s_425
CU923743	Anaerolineae	SBR1031	midas_f_1469	midas_g_4419	midas_s_4419
CU918857	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_467	midas_s_467
FJ710678	Anaerolineae	Anaerolineales	Anaerolineaceae	Ca_Villigracilis	midas_s_471
CU923695	Anaerolineae	Anaerolineales	Anaerolineaceae	Anaerolinea	midas_s_5349
CU924221	Anaerolineae	RBG-13-54-9	midas_f_5792	midas_g_5834	midas_s_5834
CU921338	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_467	midas_s_6158
AJ306792	Anaerolineae	Caldilineales	Amarolineaceae	midas_g_667	midas_s_667
EF515680	Anaerolineae	Anaerolineales	Anaerolineaceae	Bellilinea	midas_s_6727
EF515619	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_6140	midas_s_7406
CU927538	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_789	midas_s_789
AB514641	Anaerolineae	midas_o_1	midas_f_813	midas_g_813	midas_s_813
FN563322	Anaerolineae	midas_o_1	midas_f_813	midas_g_824	midas_s_824
GQ182573	Anaerolineae	Anaerolineales	Anaerolineaceae	Leptolinea	midas_s_825
CU923482	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_156	midas_s_876
EF515722	Anaerolineae	Anaerolineales	Anaerolineaceae	midas_g_156	midas_s_956

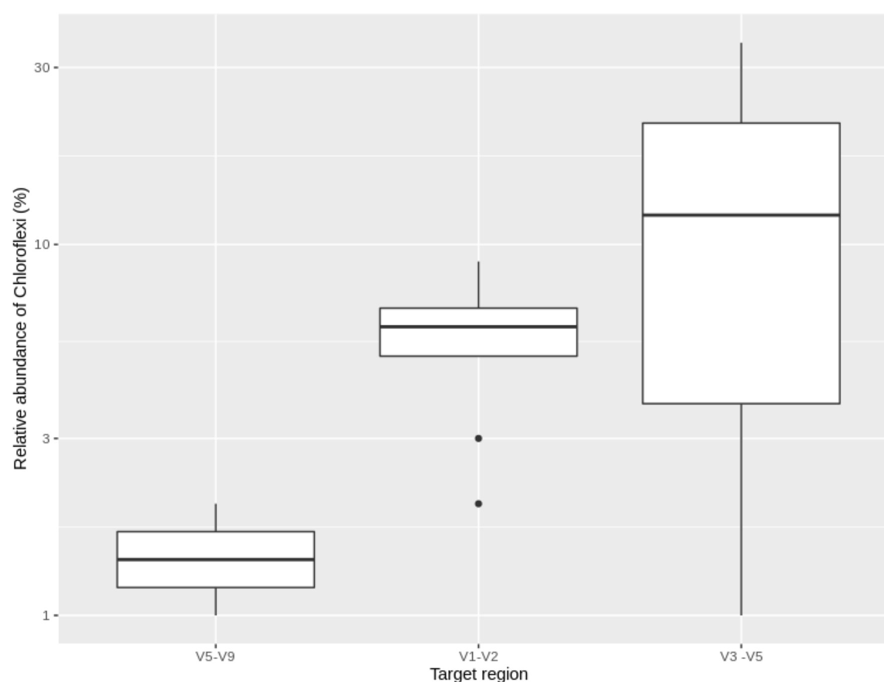

Figure S1 Boxplot showing the relative abundance of Chloroflexi using primers targeting for different 16S rRNA gene region (by average).

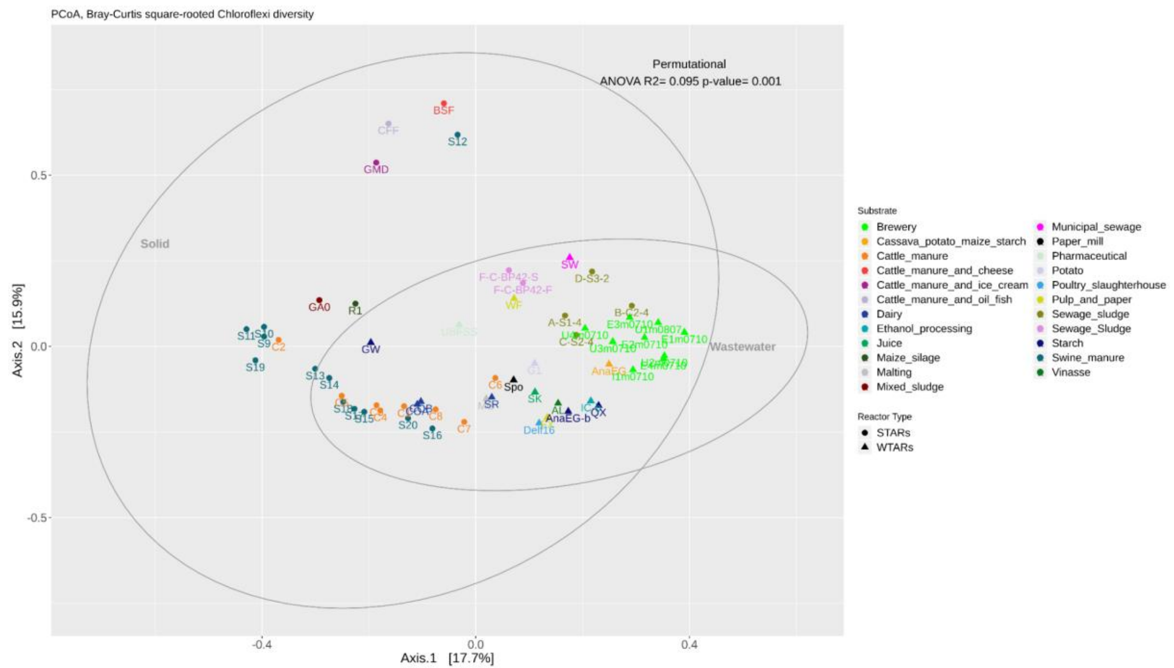

Figure S2 PCoA conducted using distance matrices constructed using Bray Curtis metric, PERMANOVA test was used to determine the differences between the samples by reactor type.

CAPÍTULO 4

Genomas ensamblados a partir de metagenomas muestra la redundancia funcional de organismos no cultivables del filo Chloroflexi en reactores de lodos activados, anammox y metanogénicos

Title: Microbial assembled genomes reveal functional redundancy of uncultured Chloroflexi in activated sludge, methanogenic and anammox wastewater treatment reactors

Patricia Bovio-Winkler[1] • Leandro D. Guerrero[3] • Leonardo Erijman[3] • Pía Oyarzua[4] • María Eugenia Suarez Ojeda[4] • Angela Cabezas[2] • Claudia Etchebehere[1]

[1]Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomic, Biological Research Institute "Clemente Estable", Avenida Italia 3318, CP: 11600, Montevideo, Uruguay.

[2]Instituto Tecnológico Regional Centro Sur, Universidad Tecnológica, Francisco Antonio Maciel s/n, CP: 97000, Durazno, Uruguay.

[3]Instituto de Investigaciones en Ingeniería Genética y Biología Molecular "Dr Héctor N. Torres" (INGEBI-CONICET) Vuelta de Obligado 2490 - C1428ADN; Buenos Aires, Argentina

[4]GENOCOV Research Group, Department of Chemical, Biological and Environmental Engineering, Escola d'Enginyeria, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

* Correspondence to: Claudia Etchebehere

Microbial Ecology Laboratory, Department of Microbial Biochemistry and Genomics, Biological Research Institute "Clemente Estable", Av. Italia 3318. CP 11600. Montevideo - Uruguay.

Phone: (598) 24871616. Fax: (598) 24875461

cetchebehere@iibce.edu.uy

Abstract

The phylum Chloroflexi is highly abundant in diverse wastewater treatment bioreactors, including aerobic or anaerobic, carbon or nitrogen removal reactors. Nevertheless, their function is not yet well understood as most species are not present in axenic cultures. It has been suggested that they play relevant roles in these ecosystems, particularly in degrading carbon compounds and on structuring flocs or granules. We here used a metagenomic approach to investigate their diversity and metabolic potential in three completely different bioreactors: a methanogenic full-scale reactor, a full-scale activated sludge reactor and a lab scale anammox reactor. Differential coverage binning approach was used to assemble 17

Chloroflexi genomes from metagenomes. The ANI value indicates that all assembled genomes may represent novel species which emphasizes that this phylum is still understudied in wastewater treatment systems. Even though samples analyzed were collected from completely different bioreactors, the assembled genomes shared several metabolic features: anaerobic metabolism, fermentative pathways and several hydrolytic enzymes. Interestingly, Chloroflexi genomes from the anammox reactor indicate a putative role in Nitrogen conversion. Filamentous morphology, and genes related to adhesiveness and exopolysaccharides production were detected. Our results suggest that Chloroflexi participate in organic matter degradation, nitrogen removal in anammox reactors and biofilm aggregation.

4.1 Introduction

The phylum Chloroflexi, comprises an ecologically and physiologically diverse group of bacteria, which have been detected in high abundance in methanogenic reactors (Petriglieri et al. 2018; Bovio et al., 2020), anammox reactors (Kindaichi et al. 2012; Cao et al. 2016; Wang et al. 2016) and activated sludge systems (Speirs et al. 2017; Andersen et al. 2018). Despite of their high abundance in wastewater treatment systems (WWTS), the basic ecophysiology of this group is still largely unknown because only few pure cultures have been investigated to date (Sekiguchi et al. 2003a; Yamada et al. 2006; Yamada et al. 2007a; Yoon et al. 2010). The study of the phylum Chloroflexi has been approached from different molecular biology techniques, allowing the formulation of some hypotheses about its role in WWTS. The three types of reactors above mentioned share some of these hypotheses. Based on the isolated members, assembled genomes and *in situ* characterization, their role seems to be focused on the hydrolysis of complex organic matter (eg. by expression of membrane associated hydrolytic enzymes), fermentation of carbohydrates and proteins, and degradation of debris from lysed bacterial cells (Kragelund et al. 2007; Nielsen et al. 2009; McIlroy et al. 2016; McIlroy et al. 2017; Andersen et al. 2018; Petriglieri et al. 2018; Nierychlo et al. 2019b). On the other hand, it has been widely speculated that Chlorofleximembers might play an important role during the sludge granulation or flocculation process because as they grow as filaments (and some members have advantageous cellular adhesiveness) they may form the initial framework for small sludge particles, serving as cores or carriers in anammox, anaerobic and aerobic reactors (Gao et al. 2011; Chu et al. 2015; Cao et al. 2016; Xia et al. 2016). On the other hand, several authors have related the overgrowth of Chloroflexi with

bulking episodes in full scale methanogenic reactors and activated sludge, meanwhile in anammox reactors there was only one report in a lab scale reactor (Sekiguchi et al. 2001; Björnsson et al. 2002a; Li et al. 2008; Song et al. 2017; Nierychlo et al. 2020b). Another hypothesis of the prevalence of this group would be based on the syntrophic association with hydrogenotrophic methanogenic Archaea, supported by the fact that all Chloroflexi isolated members from methanogenic reactors required co-cultivation with methanogenic archaea for their efficient growth (e.g.: Methanospirillum hungatei DSM864 or Methanothermobacter thermautotrophicus DSM1053) (Sekiguchi et al. 2003b; Yamada et al. 2005; Yamada et al. 2006; Yamada et al. 2007b; Sun et al. 2016). From the isolated members and assembled genomes, we know that some of the fermentation products of Chloroflexi are acetate and hydrogen (Sekiguchi et al. 2003b; Yamada et al. 2006; Yamada et al. 2007b). Thus, these products could be consumed by acetotrophic and hydrogenotrophic methanogens favoring the Chloroflexi growth. Also, a co-localization was reported between *Ca.* Brevefilum fermentans (Anaerolineae class, Chloroflexi phylum) and filamentous Archaea Methanosaeta, suggesting potential undetermined synergistic relationships (McIlroy et al. 2017). Finally, we reported a positive correlation between Chloroflexi and Euryarchaeota (methanogens) in anaerobic reactors (Bovio-Winkler et al. 2021).

In the anaerobic ammonium oxidation (anammox) process, ammonium oxidation and nitrite reduction are coupled to form nitrogen gas under anoxic conditions (Strous et al. 1998). The main hypothesis about the role of Chloroflexi in anammox reactors would seem to be focused on their capability of scavenge organic matter derived from anammox bacterial cell debris, survive from soluble microbial products (SMP) and/or extracellular polymeric substances (EPS) of autotrophs (Kindaichi et al. 2012; Chu et al. 2015). This hypothesis is supported because this occurs even when anammox reactors were fed with synthetic wastewater containing NH_4^+ as the sole electron donor and no organic carbon compounds (Kindaichi et al. 2004; Strous et al. 2006; Cho et al. 2011; Kindaichi et al. 2012). Recent metagenome based works also suggested that members of Chloroflexi could facilitate a nitrite loop with anammox bacteria or support complete denitrification due to the gene expression of the nitric oxide reductase (norZ) and the nitrite reductases (nirK, nirS), thus enhancing the overall nitrogen removal performance in bioreactors (Lawson et al. 2017; Zhao et al. 2018). On the other hand, some Chlorofleximembers encoded the function of biosynthesizing exopolysaccharides for anammox consortium aggregation, based on the partial nucleotide sugars produced by anammox bacteria (Zhao et al. 2018).

The study of Chloroflexi ecophysiology has implications for plant operation and our general understanding of the ecology of WWTS, in particular to prevent and control bulking problems. We hypothesized that even though anaerobic, aerobic and anammox reactors harbor different genera and species of Chloroflexi, their function is redundant. Therefore, different genera or species perform the same metabolic function. Hence, the aim of this study was to apply genome-centric metagenomics to obtain representative genomes of non yet cultured members of Chloroflexi from different wastewater treatment systems and thus reveal their phylogenomic diversity and metabolic potential in each type of system. To achieve this, 16S rRNA gene amplicon sequencing was used to determine the distribution, diversity, and abundance of Chloroflexi in anammox, aerobic and anaerobic wastewater treatment systems. The morphology and abundance were also studied by specific FISH. Their metabolic potential and phylogeny were analyzed from 17 near-complete genomes classified within Chloroflexi, 16 MAGs within Anaerolineae class and 1 MAG within Dehalococcoidia assembled through differential coverage binning. The combination of different molecular approaches and the study of samples from three different reactors allowed us to obtain information on the specific role of Chloroflexi in the different wastewater treatment systems.

4.2 Material and methods

4.2.1 Sampling and DNA extraction

Samples from the following four reactors were analyzed: 1) a full-scale UASB methanogenic reactor installed in Uruguay treating effluent from a brewery-malt processing industry (MO) (previously studied by (Borzacconi et al. 2008; Bovio et al. 2019), 2) a full-scale aerobic reactor treating health-care waste treatment wastewater (RH) installed in Uruguay, 3) a lab scale UASB Anammox reactor (AM) operated in Barcelona, Spain fed with synthetic wastewater (previously studied by (Juan-Díaz et al. 2021)) and 4) a lab scale SBR Anammox reactor (S) operated in Barcelona (previously studied by (Isanta et al. 2015; Reino and Carrera 2017)) fed with real municipal wastewater coming from a WWTP located in Catalonia, NE Spain (Table 4.1).

Table 4.1. Characteristics of the studied reactors

Reactors	Configuration	Samples name	Days of samples	Scale	Substrate	Condition
Activated sludge	floccular biomass	RH4 RH5 RH6	0 300 390	Full-scale	health-care waste treatment wastewater	stable conditions for more than 6 years
Methanogenic	UASB/granular biomass	MO1 MO2 MO3	0 210 300		brewery-malt processing industry	stable conditions for more than 12 years
Anammox	SBR/granular biomass	AM	-	Lab scale	synthetic wastewater	stable conditions for more than 6 years
	UASB/granular biomass	S1 S3 S5	81		real urban wastewater	81 days of operation (seeded with sludge from AM)

In order to apply the differential coverage binning approach to assemble genomes (Albertsen et al. 2013) three samples taken at different times were taken from the methanogenic and aerobic reactors, another three samples were taken at different heights from the S anammox reactor (S1, S3 and S5) and also one sample was taken from AM reactor. All reactors were operated under mesophilic conditions. Biomass was stored at -20 °C for sequencing workflows and fixed according to the protocol of (Amann et al. 1995) for Fluorescence *in situ* hybridization (FISH) and stored at -20 ºC. For DNA extraction, sludge samples were thawed and centrifuged (5 min, 10,000 g). Approximately 0.35 g of wet pellet was used for DNA extraction with the ZR Soil Microbe DNA MiniPrepTM kit (Zymo Research) according to the manufacturer's instructions. The quality of the extracted genomic DNA was determined by 1% agarose gel electrophoresis (Nucleic Acid Stain, GoodViewTM, Beijing) and stored at -20 °C until further use.

4.2.2 Community profiling using 16S rRNA gene amplicon sequencing

The taxonomic composition of the communities was studied by amplicon sequencing of the 16S rRNA gene. For that, ten samples were selected for sequencing on an Ion Torrent PGM (Life Technologies) in the platform at the Biological Research Institute 'Clemente Estable' (Montevideo, Uruguay), with primers which covered the region V4: 520F (5`-AYTGGGYDTAAAGNG-3`) and 802R (5`-TACNNGGGTATCTAATCC-3`), as previously described (Claesson et al. 2009) as was previously described (Callejas et al., 2019). The raw data analysis was performed using quantitative insights into microbial ecology' pipeline (QIIME2 2020.11 release) (Bolyen et al., 2019). Multiplexed single-end sequencing reads (1,628,931 in total) were imported into the QIIME2. The 'divisive amplicon denoising algorithm' DADA2 (Callahan et al. 2016) plugin in QIIME2 was used to 'denoise' sequencing reads. This step filters out noise and corrects errors in marginal sequences,

removes chimeric sequences and singletons and finally dereplicates the resulting sequences, resulting in high resolution amplicon sequence variants (ASVs) for downstream analysis. The consensus sequences for the ASVs were classified with a classify-sklearn classifier trained against the the pre-trained classifier MiDAS 3.7 database (Nierychlo et al. 2020a).

All the statistical analysis was performed in R version 3.5.1 (R Core Team 2020) with R Studio environment Version 1.3.1093. The biom file from QIIME2 was imported and analyzed through phyloseq-modified workflow (McMurdie and Holmes 2013). Sequences were rarefied to the lowest number of sequences per sample (n= 40,016 sequences). In order to determine the taxonomic composition of Chloroflexi in each sample, a specific analysis of the sequencing data was performed using only the sequences affiliated to the phylum Chloroflexi. We calculated the relative abundance of the phylum Chloroflexi at different taxonomic levels using the total of sequences of Chloroflexi in each sample (normalization). Barplots and heatmaps were generated using ampvis2 (v.2.6.5) (Andersen et al. 2018a).

4.2.3 Phylogenetic analysis

A phylogenetic tree was constructed to evaluate the phylogenetic position of the ASVs classified within the phylum Chloroflexi. It included ASVs with a relative abundance greater than 5% in at least one sample (sequences from amplicon sequencing), 16S rRNA sequences from assembled genomes (retrieved from the metagenomes as will be explained) and sequences from cultured representatives and MAGs retrieved from the database (NCBI). Four 16S rRNA gene sequences from different genera within the phylum Thermotoga served as an outgroup for rooting the tree. The phylogenetic tree was performed using the maximum likelihood method with the generalized time reversible (GTR) (Nei and Kumar 2000) substitution model and GAMMA distribution model using 1,000 bootstraps in MEGA X (Kumar et al. 2018). The resultant phylogenetic tree was visualized in iTOL (Letunic and Bork 2019).

4.2.4 Fluorescence in situ hybridization (FISH)

FISH analysis was performed in all samples using probes targeting the phylum Chloroflexi GNSB941 and CFX1223R labeled with Cy3 (Björnsson et al. 2002b; Gich et al. 2002). The analysis was performed to determine the morphology of members affiliated to the Chloroflexi phylum detected in these samples. Fixation of the samples and FISH was performed according to the protocol of (Amann et al. 1995). Hybridization was performed according to

(Björnsson et al. 2002b), with the conditions established by these authors. For cellular staining 4', 6' -diamin-2-phenylindole (DAPI) was used. The pure culture of *Sphaerobacter thermophilus* strain DSM 20745 was used as positive control, and *Escherichia coli* strain DH5α was used as negative control. The glass slides with the samples were embedded in Citifluor AF1 (Citifluor, Canterbury, United Kingdom). Images of fluorescent cells were recorded using a confocal laser scanning microscope (LSM 800, Carl Zeiss, Oberkochen, Germany). For detection of DAPI and Cy3 -labeled cells, the 405 nm and 543 nm lasers were used, respectively.

4.2.5 Metagenome sequencing

Ten metagenomes were shotgun sequenced by Illumina HiSeq 4000 platform (Macrogen, Seoul, Korea) using Library Kit TruSeq Nano DNA Kit (350bp). The yield was approximately 50 Gb (~5Gb/sample) of raw short-read sequences per library (2 × 100 bp).

4.2.6 Genome assembly, binning and genome annotation

The global quality of raw and trimmed reads was checked using FastQC (v0.11.8) (Andrews 2010). The reads were trimmed to remove adapters and bases below a quality score of 25 using Trimmomatic (v0.39) (Bolger et al. 2014). The trimmed reads from each reactor (MO, RH or AMX, separately) were pooled and assembled using MEGAHIT (v1.1.4-2) (Li et al. 2016) with minimum k-mer length 43, maximum k-mer 75, with steps of four. The contigs obtained shorter than 1000 bp were removed.

Quality filtered reads for each metagenome were mapped to the co-assembly of short reads (>1 kb) using Bowtie2 (v2.3.4.1) (Langmead and Salzberg 2012) with default parameters. The sequence alignment map (SAM) files were converted to a binary format (BAM) that were then imported into the MetaBAT 2 (v2.12.1) (Kang et al. 2015) to determine genome bins based on differential coverage (Albertsen et al. 2013). The completeness levels and contamination of the bins were assessed using CheckM (v1.0.13) (Parks et al. 2015). The automatic taxonomic classification of CheckM was then extracted to determine the nature of each bin considered for further refinement and validation. The bins classified as bacteria with an estimated completeness >80% and contamination <10% were reclassified to determine if they belonged to the Chloroflexi phylum. This was carried out using the open reading frames (ORFs) identified by the Prodigal program (Hyatt et al. 2010) in metagenomic mode (-p meta). ORFs were translated to amino acid sequences and subsequently searched for using

HMMER 3.0 (http://hmmer. Janelia.org/) against a set of 107 hidden Markov models of essential single-copy genes (Johnson et al. 2010) using default settings and trusted cutoff enabled. Protein sequences coding for essential single copy genes were searched against NCBI non-redundant database (retrieved in 2018) using BLASTP and an e-value cutoff of 10^{-6}. The final prediction was given as the organism with the highest number of best hits across all genes belonging to Chloroflexi. These extracted bins were mapped with Bowtie2 and reassembled using SPAdes 3.10.0 (Bankevich et al. 2012). All resulting contigs of > 1500 bp were clustered using ESOM tools (emergent self-organizing map) (Ultsch and Mörchen 2005), on the basis of its tetranucleotide frequency to help extract contigs of interest in each genome. Statistics of the reassembled bin were obtained using QUAST v5.0.2 and CheckM lineage wf (v1.0.13) with default settings. Genes encoding ribosomal RNAs (rRNA) were predicted using the Bacterial ribosomal RNA predictor (Barrnap, https://github.com/tseemann/barrnap).

4.2.7 Metabolic Analysis

Protein coding sequences (CDS) were determined with Prokka (1.14.5) (Seemann 2014). Predicted amino acid sequences (amino acid sequences in FASTA format produced by PROKKA) were annotated with Prokka, KOALA (KEGG Orthology And Links Annotation) for K number assignment of KEGG Genes (Kanehisa et al., 2016) and Metaerg (against the Pfam and TIGRfam HMM models) (Dong and Strous 2019).

4.2.7 Phylogenomic analysis

We performed taxonomic annotation on recovered genomes sequences using Genome Taxonomy Database GTDB-Tk (version v1.3.0 and GTDB-Tk reference data version r95) (Parks et al. 2018). Genome phylogenetic analyses were performed with the GTDB-Tk tool using the de novo workflow with a set of 120 single copy marker proteins and the genome taxonomy database (GTDB). Newick format tree file was uploaded onto iTOL, a web based tool for annotating and editing trees (Letunic and Bork 2019).

4.3 Results and discussion

In this work, organisms from the Chloroflexi phylum were studied in three different wastewater treatment systems: a methanogenic reactor (anaerobic for C-removal), an activated sludge reactor (aerobic for C-removal) and an anammox system (anoxic for N-

removal). Our question is whether the organisms classified within this phylum are the same in all three systems and if they have the same function. To elucidate this, we applied a series of molecular biology tools to samples taken from the three reactors.

4.3.1 Abundance, diversity and taxonomy of Chloroflexi members in the different wastewater systems

According to the 16S rRNA gene sequencing, the phylum Chloroflexi presented variation in their relative abundance, with higher values observed for the anammox reactor (26.9 - 33.5 % for AMX, 4.9 -22 % for RH and 0.8 - 10.7 % for MO) (Figure 4.1a). Then, we calculated the relative abundance of the phylum Chloroflexi at different taxonomic levels using the total of sequences of Chloroflexi in each sample (normalization). The results confirmed that Anaerolineae class widely dominated the Chloroflexi community in all reactors with a relative abundance ranging from 93.8 to 100% (Figure 4.1b).

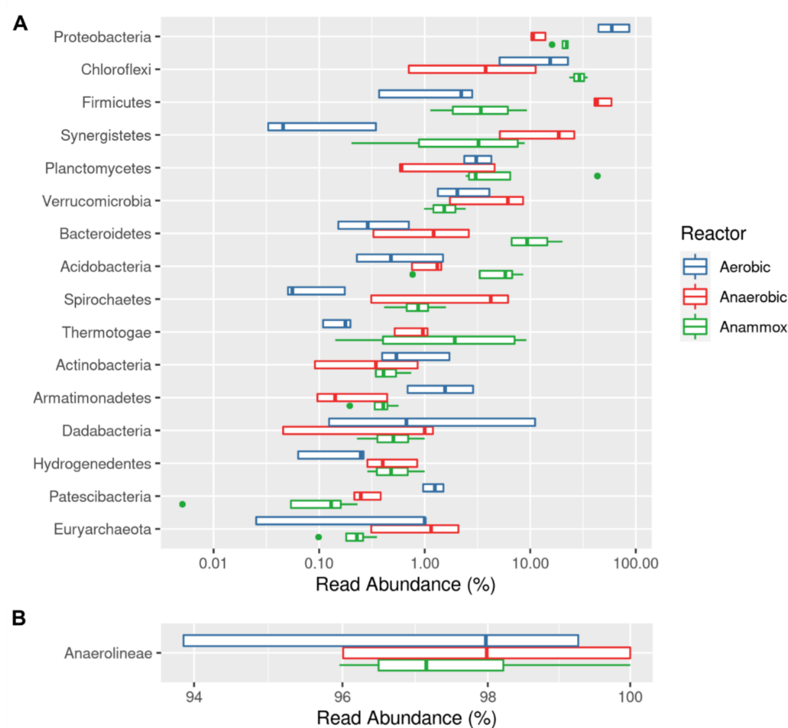

Figure 4.1. Taxonomic composition of the microbial communities in the samples taken from the three systems according to the 16S rRNA gene amplicon sequence. A) Relative abundance of the different phyla B) Relative abundance of the Anaerolineae class within the phylum Chloroflexi (only reads classified to the Anaerolineae class were observed in the Chloroflexi phylum).

The predominance of Anaerolineae class was in accordance with results previously reported in aerobic, anaerobic and anammox reactors (Mielczarek et al. 2012; Gonzalez-Gil et al. 2015; Cao et al. 2016; Pereira et al. 2017; Bovio et al. 2019; Campanaro et al. 2020; Bovio-Winkler et al. 2021).

MiDAS database was used to classify the sequences at the different taxonomic levels, this database was specifically designed for activated sludge systems and was also expanded to other wastewater treatment systems. The advantage of using MiDAS is that no-described microorganisms (with no culture representatives), frequently detected in activated sludge systems, were named with a number.

According to this classification at the genus level the genus *Candidatus* Villigracilis (from the family Anaerolineaceae) were observed in the three reactors but with higher abundance in the anaerobic reactors (MO and AMX), while the genus MiDAS g-1668 (classified within Caldilineaceae family) was also present in the three reactors but predominate in the aerobic reactor (Figure 4.2a).

A

Anaerolineae	Aerobic			Anaerobic			Anammox			
	RH4	RH5	RH6	MO1	MO2	MO3	AM	S1	S3	S5
Anaerolineaceae; Ca_Villigracilis	5	3.2	3.6	17.6	24.1	3.7	73.4	12.3	20.9	19.6
midas_f_8354; midas_g_8354	0.3	1	0.4	0	0.6	0	0	48.7	46.8	57.8
Caldilineaceae; midas_g_1668	70.9	13.3	46.4	3.2	2	3.8	1.1	2.7	1.2	0.2
midas_f_813; midas_g_5336	0	0	0	15.8	25.4	29.2	0	0.1	1.1	0.4
midas_f_119; midas_g_119	13.1	4.3	26.3	0	0.2	0	0	1.6	1.5	0
Caldilineaceae; midas_g_3043	0.4	29.8	0	0	3	0	2.5	2.7	1.7	0.6
Caldilineaceae; midas_g_344	0	0	0	0	0	0	0	5.6	10.9	12
Caldilineaceae; midas_g_169	0.1	26.8	0	0	0.2	1	0	0	0	0
Amarolineaceae; midas_g_667	0	0	0	7.5	6.9	7.8	0	0	0.4	0
Anaerolineaceae; Flexilinea	0	0	0	13.3	2.1	6.7	0	0	0.1	0.1
Anaerolineaceae; Leptolinea	0	0.2	0	4.7	2	8.2	0	0	0.1	0.1
midas_f_813; midas_g_7034	0	0.1	0	0	0.1	0	11.1	0.5	0.8	1.7
Anaerolineaceae; midas_g_156	0	0.4	0.3	0	1.4	1.7	0	6.6	0.9	1.1
Anaerolineaceae; midas_g_467	0	0.3	0	3.9	2.8	4.8	0	0.2	0.2	0.2
Anaerolineaceae; Bellilinea	0	0.1	0	3.6	1.3	5.1	0	0.3	0	0

B

	Aerobic			Anaerobic			Anammox			
	RH4	RH5	RH6	MO1	MO2	MO3	AM	S1	S3	S5
midas_g_8354; ASV5347	0	0	0	0	0	0	0	46	45.2	0
Ca_Villigracilis; ASV5807	0	0.4	0	0	0	0	46.7	0	0	10.1
midas_g_8354; ASV5346	0	0.2	0	0	0	0	0	0	0	55.1
midas_g_119; ASV5367	0	3.6	26.3	0	0	0	0	0	0	0
midas_g_3043; ASV5551	0	29.8	0	0	0	0	0	0	0	0
midas_g_5336; ASV5472	0	0	0	14.7	0	14	0	0	0	0
midas_g_1668; ASV5597	28.6	0	0	0	0	0	0	0	0	0
midas_g_1668; ASV5607	26.6	0	0	0	0	0	0	0.5	0.6	0
midas_g_1668; ASV5598	0	0	23.8	0	0	0	0	0	0	0
Ca_Villigracilis; ASV5834	0	0.2	0	0	0	0	20.6	0	0	0
Ca_Villigracilis; ASV5804	0	0	0	0	20.8	0	0	0	0	0
midas_g_5336; ASV5471	0	0	0	0	19	0	0	0	0.8	0
midas_g_1668; ASV5608	0	10	9.3	0	0	0	0	0	0	0.2
Ca_Villigracilis; ASV5806	0	0	0	0	0	0	0	6.9	12.2	0
midas_g_1668; ASV5600	0	3.3	13.3	0	0	0	0	0	0	0
midas_g_344; ASV5666	0	0	0	0	0	0	0	5.5	10.8	0
midas_g_119; ASV5368	13.1	0	0	0	0	0	0	1.6	1.5	0
midas_g_1668; ASV5599	15.7	0	0	0	0	0	0	0	0	0
midas_g_5336; ASV5467	0	0	0	1.1	0	14.1	0	0	0	0
Flexilinea; ASV5889	0	0	0	12.2	0	3	0	0	0	0
midas_g_7034; ASV5674	0	0.1	0	0	0	0	11.1	0	0	1.7
midas_g_169; ASV5581	0	11.9	0	0	0	0	0	0	0	0
midas_g_344; ASV5667	0	0	0	0	0	0	0	0	0	11.9
midas_g_169; ASV5582	0	11.8	0	0	0	0	0	0	0	0
midas_g_667; ASV5478	0	0	0	7.5	0	3.9	0	0	0	0

Figure 4.2. Heatmaps showing the relative abundance of members of the phylum Chloroflexi classified at A) genus level, B) ASV level according to the 16S rRNA genus analysis

Several genera classified in MiDAS database but with no culture representative presented high abundance within the Anaerolineae class. But these genera were specific for each kind of reactor, for example, the genus midas_g_8354 was specifically abundant in the anammox reactor (AMX), while the genus midas_g_5336 was specifically abundant in the anaerobic reactor (MO). From the 15 genera with higher abundance within Chloroflexi, only four: *Ca.* Villigracilis, *Flexilinea*, *Bellilinea* and *Leptolinea,* have been metabolically characterized. Therefore, it is of utmost importance to isolate these microorganisms in pure culture or to assemble genomes from metagenomes in order to get a deeper insight into their role in these systems. *Ca.* Villigracilis has been reported as an abundant filament widely distributed in activated sludge systems, probably contributing to the matrix supporting floc formation (Nierychlo et al. 2019b). It was also described as the first facultatively anaerobic member belonging to the Anaerolineaceae family so far (Nierychlo et al. 2019b). Despite that, *Ca.* Villigracilis presented a lower relative abundance in the samples taken from the aerobic reactor (with floccular biomass) than in the samples taken from the anaerobic reactors (with granular biomass). *Flexilinea flocculi* has been described as strictly anaerobic chemoorganotroph involved in carbohydrates fermentation (Sun et al. 2016). The presence of *Flexilinea, Bellilinea* and *Leptolinea* only in the methanogenic reactor could be explained by its need for co-cultivation with a hydrogenotrophic methanogen, which enhanced its growth in pure culture (Sun et al. 2016).

In order to determine if the genera observed were represented by one or more species in the different reactors, the reads were grouped in amplicon sequence variants (ASVs, sequences with only 1 nucleotide difference) and then classified at species level. The results showed that the above-mentioned genera were represented by few dominant ASVs (Figure 4.2b) which were not shared within the reactors.

4.3.2 Morphology determined by FISH

As the filamentous morphology was described in most of the Chloroflexi members and this could be an important characteristic of this group of microorganisms in wastewater treatment systems with granules or flocules, the morphology of the Chloroflexi cells was study by FISH in samples from the three reactors. Using specific probes for the Chloroflexi phylum it can be

observed different filamentous morphology in the samples from reactor MO and RH (Figure 4.3).

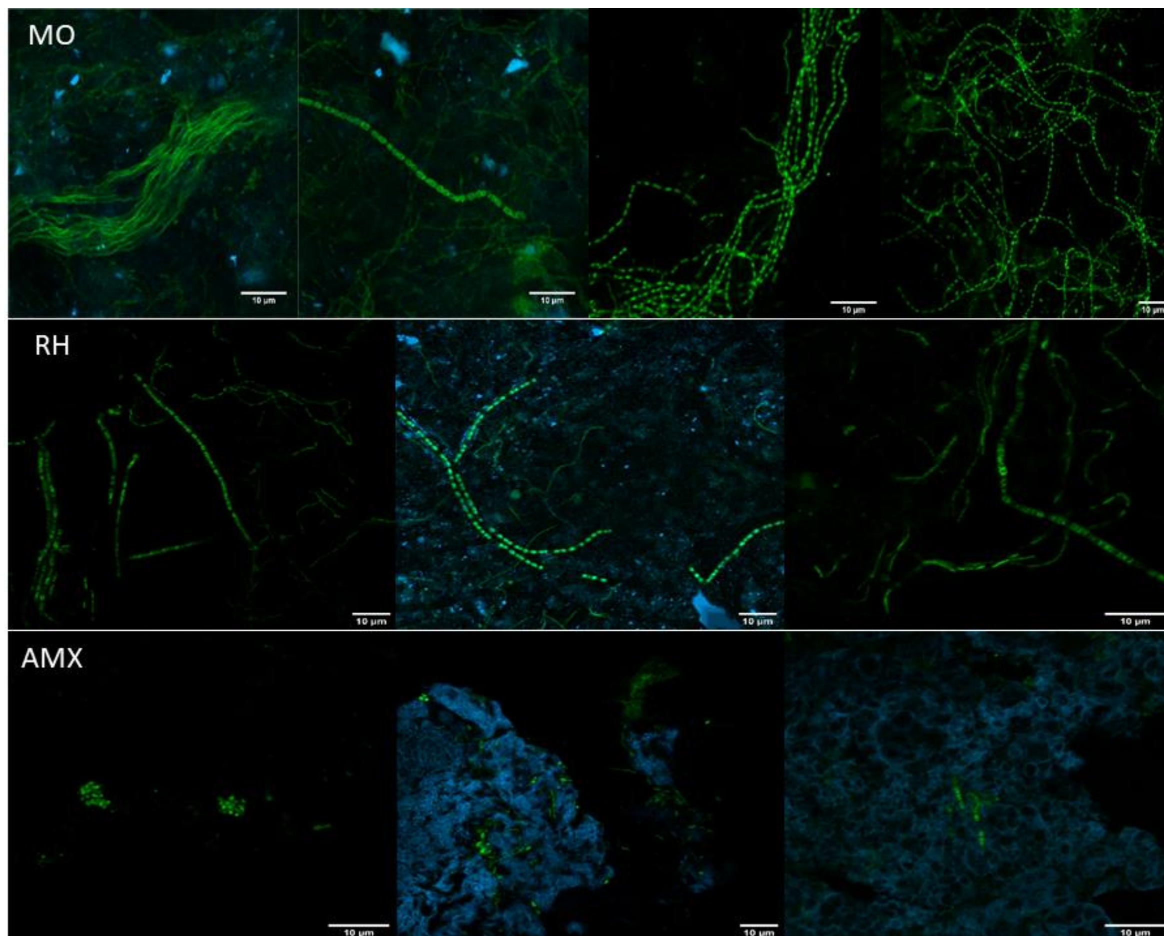

Figure 4.3. Fluorescence *in situ* hybridization (FISH) micrographs: (A) MO1, MO2, and MO3, (B) RH4, RH5 and RH6, (C) AMX. 100X Scale bar = 10 μm.

For the samples taken from the anammox (AMX) the filamentous morphology could not be observed and the signal was very low. Interference of polysaccharides in the FISH technique in anammox granules from the same reactor was previously reported (personal communication). Then, the confirmation of filamentous morphology in the anammox samples has to be further investigated.

4.3.3 General genomic features of Chloroflexi metagenomic assembled genomes

We used primarily differential coverage binning (Albertsen et al. 2013) to obtain metagenomic assembled genomes (MAGs) from the three different systems. A total of 75, 48 and 56 successfully assembled genomes were retrieved from the samples taken from MO,

AMX and RH reactors, respectively, with >50 completeness and <10% contamination. For further analyses, we selected 9 MAGs belonging to Chloroflexi from the samples taken from AMX, 4 from RH samples and 4 from MO samples, with >85.59 completeness and <7.45% contamination after the reassembly and polishing (Table 4.2). The minimum information about a MAG (MIMAG) standard was introduced to unify the field's reporting standards, stating that high-quality (HQ) draft MAGs for bacteria and archaea must be >90% complete, <5% contaminated, and, importantly, include the full-length 16S, 23S, and 5S rRNA genes, and >18 tRNA genes (Bowers et al. 2017). Under the MIMAG standards, we recovered 6 MAGs of phylum Chloroflexi (AMX9, 47, 55, 56, 57, MO118, RH21, 38, 43).

Table 4.2. Summary statistics for the 17 MAGs of Chloroflexi recovered in this study. Gray color indicates genomes with full-length 16S rRNA gene. Orange color indicates high-quality draft MAGs according to Bowers et al. 2017.

Reactor	Genome	Genome size (bp)	CDS[a]	Completeness[b]/Contamination	Contigs	N50 (bp)	16S/23S/5S[d]	tRNA[a]	%GC	Coverage	Taxonomic annotation[c]
Activated sludge	RH 21	5,657,355	5097	85.59/0	655	12275	1/1(492nt) /1	45	56.34	8.2	c_Anaerolineae; o_SBR1031
	RH 38	5,163,972	4600	87.29/1.73	1167	5326	1/1(342nt) /2	32	64.57	8.6	c_Anaerolineae; o_Caldilineales
	RH 43	12,055,906	9727	95.45/2.91	196	88920	1/1/1	62	55.76	16.6	c_Anaerolineae; o_Caldilineales
	RH 52	4,922,116	3510	96.79/1.1	201	35320	1(62nt) /1/1	57	73.22	20.4	c_Anaerolineae; o_UCB3
Methanogenic	MO 16	3,073,917	2871	86.36/3.36	387	11990	1/0/0	35	58.98	10	c_Anaerolineae; o_Anaerolineales
	MO 53	4,940,107	4005	87.27/0.91	227	44487	1(337nt) /1(370nt) /1	45	58.31	13.5	c_Anaerolineae; o_B4-G1
	MO 66	4,384,884	3803	92.73/3.36	478	14973	0/0/0	44	64.91	14.7	c_Anaerolineae; o_UBA3071
	MO 118	3,352,740	2734	90.18/2.73	86	51993	1/1(337nt) /1	38	41.05	13.2	c_Anaerolineae; o_Anaerolineales
Anammox	AMX 9	4,128,696	3450	98.18/0	54	178058	1/1(617nt) /1	48	62.69	32.7	c_Anaerolineae; o_SBR1031
	AMX 14	3,390,571	3072	92.73/0.91	124	43811	1(329nt) /0/0	41	56.87	25.1	c_Anaerolineae; o_Anaerolineales
	AMX 15	3,646,596	3399	90.91/0.18	38	293047	1/2/1	42	54.08	160.3	c_Anaerolineae; o_Anaerolineales
	AMX 39	2,675,910	2460	90.91/0.91	18	265768	1/0/0	48	52.31	19.8	c_Anaerolineae; o_Anaerolineales
	AMX 47	2,814,255	2926	91.21/0	319	11859	2/1/1	42	66.7	10.4	c__Dehalococcoidia; o__UBA2991
	AMX 55	2,990,515	2772	92.73/1.27	61	69781	1/1(317nt) /1	40	60.81	15.1	c_Anaerolineae; o_Anaerolineales
	AMX 56	4,499,265	3993	92.42/1.09	96	86147	4/2/2	58	63.48	54.1	c_Anaerolineae; o_Caldilineales
	AMX 57	4,700,319	4113	92.73/0	236	29630	1/1(461nt) /1	46	60.87	13.6	c_Anaerolineae; o_SBR1031
	AMX 68	4,329,913	4110	92.73/0.91	70	171406	1(195nt) /1(219nt) /1	46	53.34	15.8	c_Anaerolineae; o_Anaerolineales

[a] As predicted using Prokka (see Methods). The rRNA count refers to the occurrence of complete rRNA operons (5S, 16S, 23S).

[b] Genome quality estimates from CheckM (see Methods).

[c] Taxonomic assignments from GTDB-Tk (see Methods).

4.3.4 Phylogenomic analysis of Chloroflexi MAG

Using a high number of conserved genes, a better phylogeny can be constructed. A complete phylogenomic tree based on concatenated alignments of 120 single copy marker genes was constructed using 105 reference Chloroflexi genomes retrieved from NCBI (February 2021) and the 17 Chloroflexi genomes obtained in our work (Figure 4.4a). On the other hand, also a phylogenetic tree based on the 16S rRNA gene was constructed using the most abundant

ASVs belonging to phylum Chloroflexi (5% in at least one sample) and the 16S rRNA sequences retrieved from 10 MAGs (Figure 4.4b).

Figure 4.4. Phylogenetic analysis of Chloroflexi based on the metagenomics results and the 16S rRNA gene sequence analysis: A) Phylogenomic tree constructed with the 17 genomes from this work and reference genomes retrieved from the NCBI, B) Phylogenetic tree constructed with the most abundant ASVs retrieved from the 16s rRNA gene amplicon sequences (with abundance higher than 5% in at least one sample) and the sequences of the 16S rRNA gene from 10 MAGs from this work, and reference sequences of the phylum Chloroflexi retrieved from NCBI. The tree was reconstructed using the ML method and the GTR model. ML bootstrap values greater than or equal to 70% are shown at each node. Bar, 0.01 substitutions per site. The ASV are colored by reactor: from RH are in blue color, from MO are in red color, from AMX are in green color. Four species from four genus of phylum Three sequences from Thermotoga phylum were used outgroup for rooting trees.

The results obtained with the phylogenomic tree constructed based on the assembled MAGs were congruent with the 16S rRNA gene phylogenetic analysis based both in the amplicon sequences results and the 16S rRNA gene retrieved from the MAGs . Sixteen genomes were positioned into Anaerolineae class while one MAGs were positioned within Dehalococcoidia

class (classified by Gtdb-Tk). Genomes were distributed within the following orders (classified by Gtdb-Tk): Anaerolineales (MO118, MO16, AMX55, AMX14, AMX68, AMX15 and AMX39), SBR1031 (AMX9, AMX57 and RH21), B4-G1 (MO53), Caldilineales (AMX56, RH43, RH38), UBA3071 (MO66), UCB3 (RH52) and UBA2991 (AMX47) (Table 4.2). Because of the lack of genomes available in the databases, it was not possible to construct both phylogenetic trees (based on 16S rRNA gene sequences and based on genomes sequences) with the same members of Chloroflexi. For example, there were no available genomes for *Ca.* Villigracilis and *Ca.* Sarcinatrix instead found the 16S rRNA gene sequences. On the other hand, although frequently used for phylogenetic identification, 16S rRNA gene sequences are notoriously difficult to assemble from metagenomes (Blazejak and Schippers 2010), thus, 16S rRNA genes were present only in 10 of the 17 genomes assembled. Despite this, the phylogenetic placement of 10 MAGs in the phylogenetic tree based on the 16S rRNA gene was consistent with the phylogenetic tree based on 120 marker genes (Figure 4.4b). It is quite remarkable that these recovered genomes appear in several taxonomic orders, attesting to the broad diversity of Chloroflexi members inhabiting wastewater treatment systems.

In order to determine if MAGs obtained in this work represented new species, the average nucleotide identity (ANI) value was calculated between MAGs, and between MAGs and the Chloroflexi genomes retrieved from NCBI used for phylogenetic tree construction. Between MAGs, ANI values ranged from 74.4 to 79.8 (Table 4.3). Meanwhile, the ANI values between MAGs and the reference genomes ranged between 75.5 and 83.1. Using the defined ANI species threshold of 95% (Konstantinidis and Tiedje 2005), these results indicate that all genomes represent novel species with no cultured or sequenced representatives, emphasizing the lack of knowledge on the Chloroflexi community in these types of ecosystems. In addition, this analysis showed that the five genomes from anammox reactor (AMX55, AMX15, AMX39, AMX14 and AMX68) which formed a cluster without closely related genomes in the phylogenomic tree (Figure 4.4a), clustered with *Ca.* Villigracilis in the phylogenetic tree based on the 16S rRNA gene (Figure 4.4b). Our results were in accordance with several studies which reported Chloroflexi MAGs retrieved from anammox, anaerobic, and aerobic reactors used to treat waste waters were affiliated mainly to Anaerolineae class (Speth et al. 2016; Xia et al. 2016; Lawson et al. 2017; McIlroy et al. 2017; Zhao et al. 2018).

4.3.5 General metabolic pathways and genes according to the MAGs

According to the previous information, different species and genera of Chloroflexi members predominate in the different systems, then, the question is whether they have the same function or not. To infer the metabolic potential of the 17 Chloroflexi MAGs retrieved in our work, we annotated genes within each of the MAGs using a variety of protein databases . Because most of these MAGs were estimated to be between 85% and 93% complete, it is likely that they contain genes for additional pathways or transporters, not identified in this study.

4.3.5.1 The potential for anaerobic respiration

Annotation of the 17 genomes confirmed the potential of a strict anaerobic metabolism due to the incompleteness of at least one of the following pathways: tri-carboxylic acid (TCA) cycle, NADH:quinone oxidoreductase, Cytochrome c oxidase and F-type ATPase (incomplete respiration machinery) (Figure 4.5, 4.6). Also, genes for superoxide dismutase in all MAGs, catalases in five MAGs (AMX47, RH52, RH43, MO118 and MO53) and glutathione peroxidase in MO118 were annotated, indicating protection against oxidative stress (Figure 4.5, 4.6).

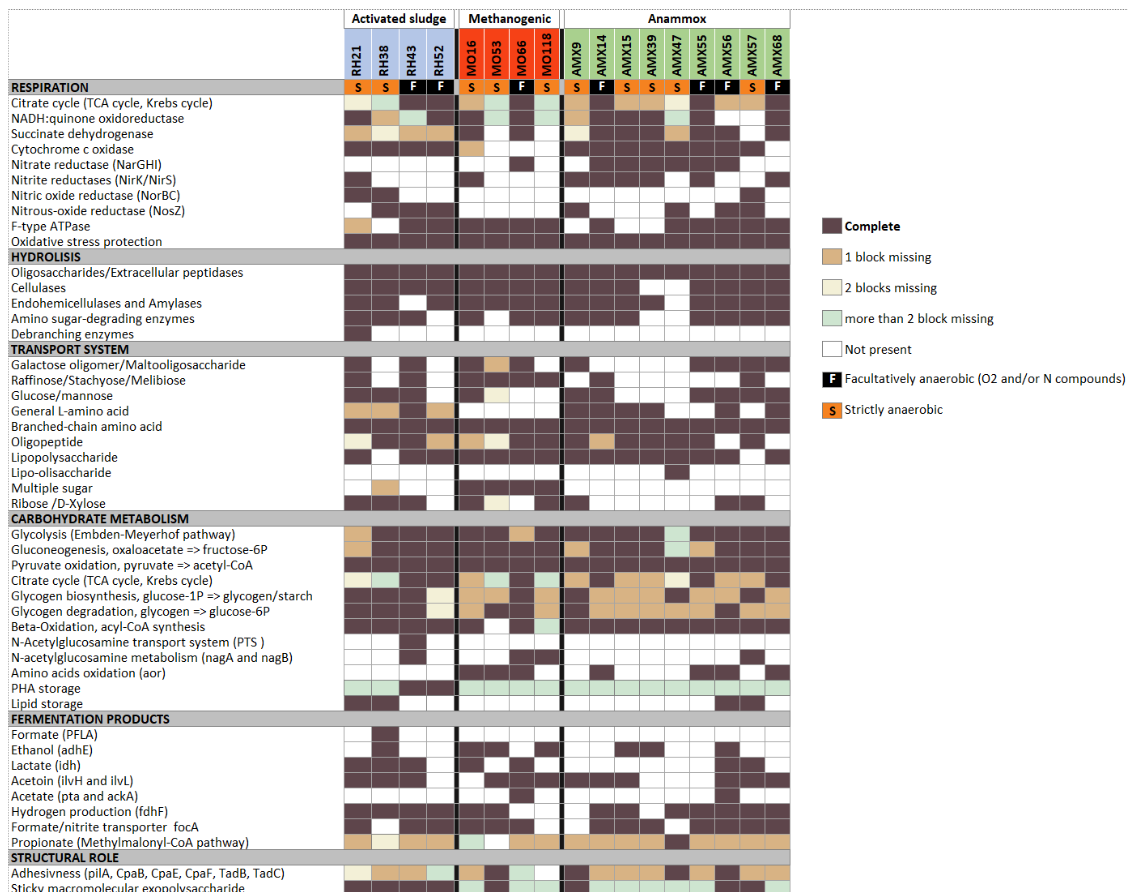

Figure 4.5. Heatmap showing the presence or absence of each metabolic pathway in each genome. Pathway absences are colored with gray. The presence of the different pathways is colored with different colors.

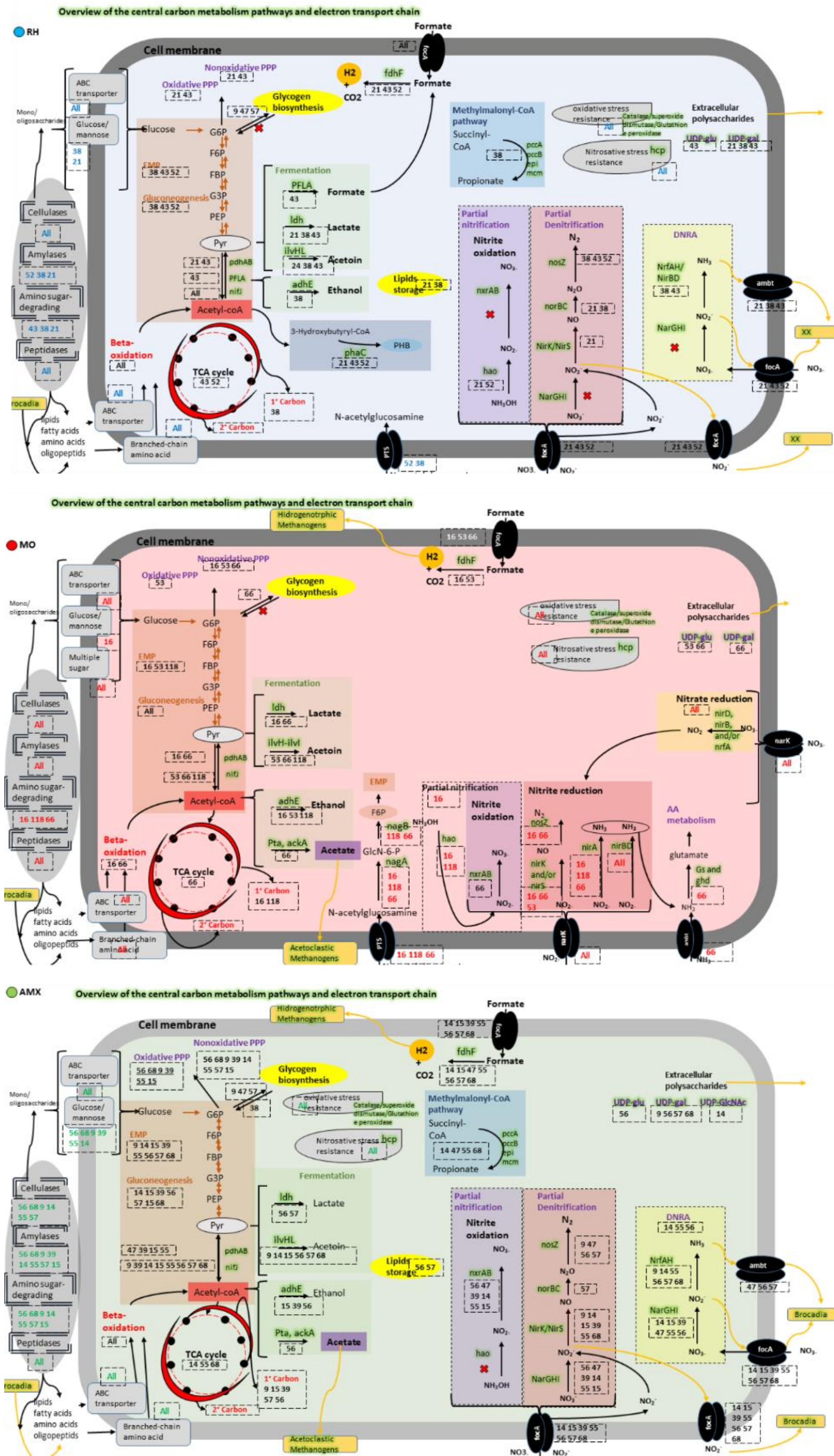

Figure 4.6. Composite metabolic overview of the 17 Chloroflexi genomes from RH, MO and AMX based on identified genes and pathways.

Some of these genes were also found in *Ca. Brevefilum fermentans* from an anaerobic digester (McIlroy et al. 2017). Only five genomes from xxx of the described Anaerolineae species (*Ardenticatena maritima*, *Ca.* Promineofilum breve, *Ca.* Amarolinea aalborgensis, *Caldilinea aerophila* and *Ca. Villigracilis*) have the complete pathway genes to utilize oxygen as an electron acceptor (Kawaichi et al. 2013; McIlroy et al. 2017; Andersen et al. 2018b; Nierychlo et al. 2019). However, aerobic uptake of substrates in full-scale activated sludge systems is a common feature of members of Anaerolineae class (Kragelund et al. 2007; Kragelund et al. 2011; McIlroy et al. 2016). On the other hand, genes for the reduction of different nitrogen compounds were found in all MAGs retrieved in this work, except in MO118, indicating that nitrogenous compounds could be used as electron acceptor in the anaerobic respiration. In summary, all these features suggested that the genomes assambled in this work become from microorganisms having an anaerobic lifestyle consistent with the anammox and anaerobic reactors.

These results are not consistent with the aerobic environment from an activated sludge reactor (RH). he genomes retrieved from RH samples did not show the complete pathways to use O_2 as the final electron acceptor. Interestingly, the assembled genomes RH43 and RH38, which phylogenetically belonged to the facultative anaerobic Caldilineales order, also lack the potential of aerobic respiration. It has been suggested that the presence of superoxide dismutase and part of respiration machinery provides mechanisms for oxygen tolerance under aerobic conditions experienced in the bioreactor (Lawson et al. 2017; Ward et al. 2018). Furthermore, these microorganisms may could form aggregates often buried deeply in floc material, which would promote micro-anaerobic conditions. Recently, *Ca.* Villigracilis was characterized physiologically *in situ* using MAR-FISH in samples from an activated sludge reactor. It was described as the first facultative anaerobic chemoorganotrophs member belonging to the Anaerolineaceae family within Chloroflexi (Nierychlo et al. 2019b). MAGs which were phylogenetically related to*Ca.* Villigracilis (AMX55, AMX14, AMX68, AMX15 and AMX39) had an incomplete respiratory machinery. Despite the *in situ* characterization of *Ca.* Villigracilis, its genome has not been recovered yet, so, it can not be used in this comparison, highlighting the importance of studies based on metagenomics.

4.3.5.2 Metabolic pathways involved in hydrolysis of polymers

In order to test the hypothesis that Chloroflexi occupy a niche capable of recycling microbial derived dissolved organic matter, and act as hydrolytic bacteria, we search for genes encoding these functions. All MAGs showed genes for the transport and catabolism of carbohydrates, peptides and amino acids to central carbon intermediates. Annotation of genes for glycosyl hydrolases such as cellulases, endohemicellulases, amylases, amino sugar-degrading enzymes, oligosaccharide-degrading enzymes, and also extracellular peptidases indicated that most MAGs had the potential of hydrolyzing cellulose, starch, protein and/or peptides (Figure 4.5, 4.6). According to the bibliography, *in situ* studies reveal high level of surface associated hydrolytic enzymes indicates their involvement in the breakdown of complex organics (Xia et al. 2007).

4.3.5.3 Central carbon metabolism

Fermentative organisms can play important roles in the system via removal of organic matter, and providing growth substrates for other community members. Multiple transporters for different organic compounds were found in the same genome, which gives species alternative routes for the incorporation and recycling of dissolved organic matter in the reactors. Genes for branched chain amino acid, lipopolysaccharide, and single and/or multiple sugar transporters were annotated in eleven genomes (Figure 4.5, 4.6). Seven MAGs also contained genes for transporters of glucose/mannose, polar amino acid, general L-amino acid and raffinose/stachyose/melibiose. These findings could indicate that all of the assembled genomes scavenge macronutrients having access to a diverse substrate pool including sugars, amino acids, proteins and fatty acids, indicating a heterotroph lifestyle. Despite this, substrate and transporters specificity is particularly difficult to deduce from genome annotation data, as bacteria are much more specialized *in situ* than what their complete metabolic potential suggests (Gelfand and Rodionov 2008; Kindaichi et al. 2013).

Genes encoding Embden-Meyerhof-Parnas (EMP) pathway for glycolysis were annotated in all MAGs except in AMX47, RH21 andMO66. These last two genomes missed only one gene of the glycolysis pathway, probably due to the incomplete genome recovery. Meanwhile AMX47 lacked genes of at least four steps of glycolysis, thus probably this microorganism did not perform glycolysis (see below). Genes for pyruvate oxidation to acetyl-CoA were annotated in all MAGs (Figure 4.5, 4.6). Therefore, the simple and complex substrates

(previously processed by their hydrolytic enzymes) imported by the transporters can then be fed into the glycolysis pathway and then fermented.

The TCA cycle was annotated in AMX68, AMX14, AMX55, RH52, RH43 and MO66, suggesting that these MAGS represent microorganisms with the capacity to use the pyruvate through fermentation or through TCA cycle. The genomic evidence suggests that terminal oxidation via the TCA cycle is not a major energy source for most of the microorganisms represented by the MAGs.

The pentose phosphate pathway (PPP) was found complete in AMX56, AMX68, AMX15, RH43, RH21 and MO53 genomes while only the non-oxidative branch of the PPP was present in AMX57, RH52, MO16 and MO66 genomes. In the latter genomes, other enzymes could be important for NADPH as a reducing equivalent and intermediates of nucleotide and histidine generation in these organisms.

As mentioned, AMX47 genome had an incomplete EMP indicating that the microorganism represented by this genome has not the capacity to grow on sugars, however, beta-oxidation pathway was annotated in this genome as an alternative to EMP. Interestingly, beta-oxidation was annotated in all MAGs except in MO53 and MO118. Recent studies in activated sludge, in addition to the fermentation of sugars by the abundant Chloroflexi, reported assimilation of short and long chain fatty acids, amino acids and glycerol by the less abundant members of Chloroflexi (Nierychlo et al. 2019), but beta-oxidation was not mentioned. Beta-oxidation was previously identified in genomes of Dehalococcoidia and Anaerolineae from different natural aquatic ecosystems (Wasmund et al. 2014; Dombrowski et al. 2017; Hug et al. 2013a). The presence of transporters for branched-chain amino acid and lipids in addition of annotated beta-oxidation pathway suggests that the metabolism of various fatty acids and branched-chain amino acids probably indicates the utilization of decaying bacterial cell materials. These results may suggest that beta-oxidation pathway represents an important metabolic route to obtain carbon and reducing equivalents.

Interestingly, genes for N-acetylglucosamine metabolism (nagA, nagB) were annotated in AMX57, RH43, MO118 and MO66 genomes although phosphotransferase system (transporter) was only annotated in RH43. It could suggest that this genome had a specific role in the breakdown of cellular material. Some Chloroflexi members appear to retrieve N-acetylglucosamine from other lysed cells in the environment, as seen with several micro-autoradiography or FISH studies that confirm their role as scavenger (Kindaichi et al. 2004; Okabe et al. 2005; Kragelund et al. 2007; Zang et al. 2008; Miura and Okabe 2008; Kragelund et al. 2011; Campbell et al. 2014; McGonigle et al. 2020). Anaerolinea oral taxon

439 which was previously reported as scavenger (presence of PTS and nagAB genes) clustered with MO118 in the phylogenomic tree (Campbell et al. 2014). It has been suggested that coexisting Chloroflexi in anammox reactors also scavenge organic matter derived from anammox bacterial cells (Kindaichi et al. 2012), therefore, we suppose that it occurs through the glycolysis and beta-oxidation pathways (degradation of fatty acids and branched-chain amino acids) because only one genomes from the anammox reactors (AMX57) had genes for N-acetylglucosamine degradation. These results confirm the previous hypothesis that Chloroflexi has an important beneficial role in feeding from the lysed bacterial cell debris.

4.3.5.4 Ability to generate energy by fermentation

The annotation of pyruvate dehydrogenase (pdhABCD), pyruvate formate-lyase (PFLA) and/or pyruvate-flavodoxin oxidoreductase (nifj) indicated the potential to convert pyruvate to acetyl-CoA in all MAGs (Figure 4.5, 4.6). The potential fermentation products annotated in the genomes were acetate (Phosphate acetyltransferase-acetate kinase pathway (pta, ackA)), ethanol (bifunctional enzyme aldehyde/alcohol dehydrogenase (adhE)), lactate (L-lactate dehydrogenase, idh), acetoin (Acetolactate synthase, ilvH and ilvI), formate (pyruvate formate-lyase) and/or propionate via methylmalonyl-CoA pathway (pccA, pccB, epi and mcm). Our results were in accordance with the information from isolates and MAGs from Anaerolineae class retrieved from wastewater treatment systems. These organisms were identified as facultative anaerobic chemoorganotrophs involved in sugar, amino acid or protein fermentation with different products (Sekiguchi et al. 2003a; Yamada et al. 2006; Yamada et al. 2007b; Sun et al. 2016).

The annotated methylmalonyl-CoA pathway in AMX14, AMX55, AMX47, AMX68 and RH38 genomes suggest a role in either amino acid degradation with propionate formation as a fermentative end product (Hug et al. 2013a; McIlroy et al. 2017; Wasmund et al. 2014). Also, an aldehyde:ferredoxin oxidoreductase (aor) was annotated in these MAGs which may function to oxidize aldehydes derived from amino acid oxidation (Heider et al., 1995). The annotation of methylmalonyl-CoA pathway and AORs which also were annotated in *Ca.* Brevefilum fermentans (McIlroy et al. 2017), indicates that propionate could be produced as a metabolic by-product from the fermentation of amino acids. These findings confirm the fermentative lifestyle in Chloroflexi populations. 13 of the 17 MAGs contain the gene that encodes formate dehydrogenase H (fdhF) that facilitates the decomposition of formic acid to H_2 and CO_2 under anaerobic conditions in the absence of exogenous electron acceptors. In

addition, the formate transporter (FocA) was annotated in most of these genomes. Thus, these genomes have the information to oxidate formate with hydrogen production. Synergistic interaction of Anaerolineae members with the hydrogenotrophic methanogenic archaea, such as M. thermautotrophicus, were previously reported in anammox reactors (Zhao et al. 2019). This could be a common scenario in anaerobic reactors and anammox reactors where hydrogenotrophic methanogenic archaea are present.

4.3.5.5 Wood–Ljungdahl pathway

The Wood–Ljungdahl pathway (WLP) consists of two branches, the methyl-branch and the carbonyl-branch. Interestingly, AMX55, AMX56 and AMX68 contained most of the genes encoding for the key enzymes required for CO_2 fixation via WLP (Figure S1). For example, AMX56 only lacked formate dehydrogenase (NADP+) beta subunit (fdhB). Considering that anammox reactors were fed with very low concentration of organic carbon, a mixotrophic growth is expected in absence of organic compounds. The ability to use CO_2 as an alternative electron donor was also observed in Chloroflexi genomes from other anammox bioreactors (Ali et al. 2020). On the other hand, the annotated WLP in genomes of Dehalococcoidia members suggested the use of this pathway to completely oxidize the organic compounds processed via beta-oxidation pathways (Hug et al. 2013; Wasmund et al. 2014) or to obtain carbon by autotrophy (Wasmund et al. 2014; Zhuang et al. 2014). WLP was common among members of the anaerobic digestion, and Chloroflexi populations were the second phyla containing this pathway (Campanaro et al. 2020).

4.3.5.6 Potential for storage of polymers and lipids

Complete pathway for glycogen biosynthesis was annotated in 9 MAGs (Figure 4.5, 4.6). These results suggest glycogen accumulation might be a common feature to cope with the changing pools of C and N in the bioreactors and to facilitate their survival when either one is temporally limited. In spite of this, glycogen degradation was only complete in RH38 genome, indicating that this genome has the information for heterotrophic growth using glycogen storage. Glycogen biosynthesis and degradation was previously reported in *Ca.* Amarolinea aalborgensis and *Ca.* Promineofilum breve (McIlroy et al. 2016; Andersen et al. 2018b). The presence of an annotated polyhydroxyalkanoate synthase (PhaC) indicates the capability for polyhydroxyalkanoates (PHA) storage in the microorganisms represented by

threeMAGs. Their potential for PHA storage was suggested to favour these organisms with the intermittent availability of carbon in these systems (Onetto et al. 2019). On the other hand, the annotation of a putative acyl-CoA:DAG acyltransferases (AtfA), catalyzing the final step in the synthesis of triacylglycerols, indicates the potential for lipid storage (Kalscheuer and Steinbüchel, 2003) in AMX56, AMX57, RH38 and RH21 MAGs. This capability was also reported for *Ca.* Promineofilum breve (McIlroy et al. 2016).

4.3.5.7 Nitrogen metabolism

Removal of ammonium and other nitrogen species from wastewater is essential to prevent pollution and ecosystem disturbances (Schmidt et al. 2003; McIlroy et al. 2016). Both nitrifiers and denitrifiers are required for efficient removal, by oxidizing ammonia and/or nitrite to nitrate, and then reducing nitrate to nitrogen gas. Different genes related to the nitrogen cycle such as the dissimilatory nitrate reduction to ammonia (DNRA), denitrification and nitrification were annotated in the genomes.

In general, several genes for each of these pathways were annotated in MAGs retrieved from anammox reactor (Figure S2). Nitrate appeared to be widely used as a potential electron acceptor by many of the AMX genomes (6 MAGs). Meanwhile, nitrate reduction was not annotated in any genome retrieved from the aerobic reactor (RH), and nitrite reduction to nitric oxide, nitric oxide reduction to nitrous oxide and nitrous oxide reduction to N_2 were not annotated in any genome retrieved from MO reactor, thus, genes for denitrification were completely absent in the recovered genomes from MO reactor.

DNRA is part of the nitrogen removal processes in anammox reactors and it has been annotated in several Chloroflexi genomes (McIlroy et al. 2016; Lawson et al. 2017; Andersen et al. 2018b). Our results suggest that Chloroflexi MAGs retrieved from the anammox reactor samples could have a specific role in the nitrogen cycle. Interestingly, we found that AMX14, AMX55 and AMX56 genomes have the capacity to perform DNRA, as genes encoding both dissimilatory nitrate (NarGHI) and nitrite reductases (NrfAH) were found Thus, these microorganisms could facilitate a nitrite loop with anammox bacteria, enhancing overall nitrogen removal performance. On the other hand, RH52, RH21 and MO118 genomes did not present neither NarGHI, nor NrfAH. The rest of the MAGs encoded one of the two genes, thus they have the capacity to perform nitrate reduction to nitrite (AMX47, AMX39, AMX15 and MO66) or nitrite reduction to ammonia (AMX9, AMX57, AMX68, RH43, RH38 and MO16).

Genes for nitrate reduction to nitrite were not annotated in genomes from RH reactors and they were annotated only in one MAG from MO reactor (MO66). These results suggested that most of the genomes have the capacity to use nitrite as electron acceptor collaborating with nitrogen removal in WWTS. Regarding the nitrate reduction in aerobic reactors, our results suggest that this is not the role of Anaerolineae members. It was suggested that the role of *Ca.* Amarolinea which had the potential to perform DNRA was to contribute to the accumulation of the latter inside the aerobic reactors at the expense of efficient nitrogen removal (Andersen et al. 2018b). Despite the annotation of DNRA in *Ca.* Amarolinea, the *in situ* characterization revealed a capability to take up substrates under anoxic conditions in presence or absence of nitrate/nitrite, suggesting fermentative metabolisms without a confirmed uptake of nitrate/nitrite (Nierychlo et al. 2019). Also putative nitrite reductase (NrfAH) have been found in *Caldilinea aerophila* and *Anaerolinea thermophile*, but their ability to utilize nitrite as an electron acceptor to support anaerobic growth is yet to be assessed (Sekiguchi et al., 2003). It would be interesting to focus efforts on the characterization of the isolates of the Anaerolineae class obtained so far, in order to determine the use of nitrogenous compounds as electron acceptors.
Except MO53 and MO118, all MAGs encode the ability to respire nitrate (NarGHI), nitrite (nirK or nirS), nitric oxide (NorBC) and/or nitrous oxide (NosZ) via partial denitrification. AMX68, AMX9, AMX39, AMX15 and RH21 harbour the genes to perform nitrite reduction to nitric oxide via NirK, and_AMX68, AMX14, AMX55 and MO16 via NirS. These results were in accordance with the fact that we this function was more frequently reported ingenomes from anammox reactors than in the ones retrieved from aerobic and anaerobic reactors (Speth et al. 2016; Lawson et al. 2017; Zhao et al. 2018). The key enzyme for denitrification is nitric oxide reductase (NorBC), which reduced nitric oxide to nitrous oxide. NorBC was annotated in AMX57 , RH38 and RH21 genomes. This gene was missing in *Ca.* Primineofilum breve and *Ca.* Brevefilum fermentans, although nitrite reductase (NrfA) and putative nitric oxide reductase (norV) respectively has been shown to reduce nitric oxide to ammonia, a reaction that was suggested to be associated with protection from nitrosative stress (Nitric oxide is the primary source of this stress) (McIlroy et al. 2016; Speth et al. 2016; McIlroy et al. 2017). *Ardenticatena maritima* had a complete set of genes for denitrification (Kawaichi et al. 2015), meanwhile NorBC was reported to be expressed in a Anaerolineae genome from an anammox reactor (Speth et al. 2016). A recent comparative study determined that Chloroflexi was the most abundant denitrifying group in a partial nitritation anammox reactor compared to an activated sludge reactor (Orschler et al. 2021).

(Kragelund et al. 2007) reported a lack of denitrifying capability of some Chloroflexi isolates retrieved from activated sludge samples (lack of the napA/narG, nirK/nirS, norBi, and nosZ genes). These results were in accordance with our results showing more genes related with denitrification in the MAGs retrieved from anammox reactors samples than in the MAGS retrieved from aerobic and anaerobic reactors. Finally, the reduction of nitrous oxide to nitrogen (NosZ) was annotated in AMX56, AMX47, AMX9, AMX57, RH52, RH43 and RH38. Together, these functions could facilitate a nitrite loop with anammox bacteria or support complete denitrification, thus enhancing overall nitrogen removal performance in the bioreactor. Recently, complete denitrification was reported in a MAG retrieved from an anammox reactor classified within Chloroflexi (Lawson et al. 2017). The exact role of heterotrophic Chloroflexi in anammox systems treating synthetic wastewater has not yet been determined, though a few clues have been uncovered. A recent metagenomic study revealed that most of the heterotrophic organisms in anammox granules encode the ability to respire nitrate via partial denitrification, possibly completing a nitrite loop with anammox and nitrite oxidizing bacteria (NOB) by reducing nitrate back to nitrite (Speth et al. 2016).
Regarding partial nitrification, the MAGs retrieved from the aerobic reactor (RH52 and RH21) have genes encoding the oxidation of hydroxylamine to nitrite via hydroxylamine dehydrogenase (Hao). Meanwhile, AMX56, AMX47, AMX39, AMX14, AMX55, AMX15 and MO66 genomes encoded a nitrite oxidoreductase (NxrAB) for the nitrite oxidation to nitrate. Finally, MO118 was the only MAGs, which did not present the potential for reduction or oxidation of nitrogenous compounds.

4.3.5.8 Filamentous morphology, adhesion capability and exopolysaccharides production

In order to test the possible role of Chloroflexi in the granules and flocs formation, three important capabilities were specifically searched in the genomes: the filamentous morphology, the adhesiveness and the production of polysaccharides.
As known from other studies, after chromosome replication and subsequent segregation of the sister chromosomes, cell division in *E. coli* is initiated with the assembly of a circumferential scaffold on the cytoplasmic membrane, the Z-ring (Goehring and Beckwith 2005). The formation of the Z-ring in *E.coli* is dependent of FtsA and ZipA genes which connect FtsZ to the cytoplasmic face of the inner membrane (Pichoff and Lutkenhaus 2002). For example, cable bacteria lack zipA, like bacteria from many other phyla (Margolin 2000). Although, either ZipA or FtsA genes are essential for formation and stabilization of the Z-

ring in *E. coli* (Pichoff and Lutkenhaus 2002), the presence of ftsA adjacent to ftsZ in cable bacteria seems to fulfill this role. Interestingly, all MAGs lacked zipA gene and filamentous morphology was observed by FISH. Therefore, the absence of the ZipA gene could prevent cell division in these Chloroflexi members.

The adhesiveness of five assembled genomes belonged to Anaerolineae class was previously addressed by (Xia et al. 2016). In this study, the expression of the tight adherence protein on the Tfp indicated its function for cellular attachment, which was further testified to be more likely related to cell aggregation other than cellulose surface adhesion. However, this Tfp structure was found to not contribute to syntrophic methanogenesis. In this work, a Tfp cluster in the isolated A. thermophilia UNI-1 was also observed, which contains a series of genes encoding pilus assembly proteins (pilA, CpaB, CpaE, CpaF) and two consecutive Tad proteins (TadB, TadC). In the same way, a complete set of genes (pilA, CpaB, CpaE, CpaF, TadB, TadC) were annotated in AMX56, AMX47, AMX9 and MO53. The rest of anammox MAGs and MO16 lack only the pilus assembly protein CpaB. As has been already noted, pili are often involved in facilitating adhesion and colonization in a wide variety of scenarios.

It has been suggested that some Chloroflexi-affiliated bacteria encoded the function of biosynthesizing exopolysaccharides such as UDP-N-acetyl-D-glucosaminuronate (UDP-GlcNAcA), GDP-mannose (GDP-Man), and GDP-D-rhamnose (GDP-Rha) for anammox consortium aggregation, based on the partial nucleotide sugars produced for example by anammox bacteria (Zhao et al. 2018). These authors suggested that Anaerolineae members could affect the nitrogen removal performance through affecting the aggregation because of the positive correlation relationship of this group with the abundance of EPS formation genes (Zhao et al. 2019). We found genes for the UDP-GlcNAc synthesis in AMX14 and UDP-galactose synthesis in AMX56, AMX68, AMX9, AMX57, RH43, RH38, RH21 and MO66 genomes. Therefore, Chloroflexi could be favoring cell aggregation and consequently the formation of cores or carriers, which help to form the initial framework of small sludge particles. .

4.3.6 Ecological significance and concluding remarks

This study provides a first insight into the Chloroflexi members diversity and role in three different wastewater treatment bioreactors as activated sludge, methanogenic and anammox reactors. Despite the phylogenetically distant Chloroflexi members recovered in our study, there was functional redundancy in all reactors. Functional redundancy enables functional

resistance of an ecosystem to environmental perturbations due to the presence of multiple species that can perform the same metabolic function such that the loss of one species due to perturbation will be substituted by another species in the community.

The Anaerolineae class largely dominated the Chloroflexi phylum community and exhibited filamentous morphology in most reactors. Each type of reactor presented differents dominant genera in which few species (represented by ASVs) were predominant. Most of these ASVs sequences were not phylogenetically closer positionated with sequences from cultured representatives. These results were reflected in the 17 genomes recovered through the differential coverage binning method, where the ANI values were indicative of new species. The MAGs obtained were affiliated with the orders Anaerolineales (7 genomes), SBR1031 (3 genomes), Caldilineales (3 genomes), UCB3 (1 genome), B4-G1 (1 genome), UBA3071 (1 genome) and UBA2991 (1 genome). It is quite remarkable that these recovered genomes appear in several taxonomic orders, attesting to the broad diversity of Chloroflexi members inhabiting wastewater treatment systems. It allowed the inference of potential metabolic diversity and ecological role of many species within a phylum with few cultured and assembled genome members.

On the one hand, we focused on determining their metabolism (carbohydrate metabolism and nitrogen cycle) in each type of system. e also focused on determining the presence of genes related to filamentous growth, adhesiveness and exopolysaccharides production in order to reveal their structural role in granules or floccules. The genomic annotation of the 17 genomes suggested: (1) anaerobic respiration with a versatile metabolism related to the hydrolysis and fermentation of complex and simple organic compounds, (2) an important role in the nitrogen cycle mainly in anammox reactors and (3) a crucial structural role in flocs and granules in all reactors.

In addition to the known role of Chlorofleximembers as hydrolytic or fermentative bacteria, our results suggest that these bacteria occupy a niche in recycling microbial dissolved organic matter, attested by the diversity of hydrolases and membrane transporters present in the genomes. This capacity gives to the Chlorofleximembers the ability to survive when the C or N source are limited in the reactors. All genomes presented the potential for degradation by glycolysis and/or beta-oxidation, with the subsequent oxidation of pyruvate by fermentation. Some of the genomes presented, in addition to the fermentative potential, the ability of using the TCA cycle. Among the fermentative products potentially generated by the Chloroflexi represented by 17 of the assembled genomes, acetate and H2 may provide growth substrates for other community members such as hydrogenotrophic or acetoclastic archaea in anaerobic

as well as anammox reactors. Interestingly, some genomes also contained genes encoding for the key enzymes required for CO2 fixation via the WLP that would allow a mixotrophic growth in the mother reactor which was fed with synthetic water.

According to the results from this work, Chloroflexi plays an important role in nitrogen cycle mainly in anammox reactors. In these reactors, all AMX genomes presented the capacity to reduce nitrate and/or nitrate to ammonia (DNRA) for direct use by anammox bacteria. Partial denitrification in AMX genomes could facilitate a nitrite loop with anammox bacteria or support complete denitrification, thus enhancing overall nitrogen removal performance in the bioreactor. The nitrite oxidation was common in genomes from anammox reactors. In the aerobic reactor, the assambled genomes did not reduce nitrate, but seem to contribute to the accumulation of ammonia by nitrite reduction, improving the nitrogen removal. In this system, Anaerolineae members also may perform partial denitrification. Finally, in the MO genomes fewer genes related to the nitrogen cycle were annotated.

Chloroflexi detected by FISH from samples taken from RH and MO reactors showed filamentous morphology. It was interesting to discover that all of them lack an important gene related to the formation of the z ring necessary for cell division. In addition to the filamentous morphology, two other important factors in the formation of granules and floccules were evaluated, such as adhesiveness and EPS synthesis. Genes for both processes were annotated in most of the assembled genomes. Therefore, filamentation, adhesiveness and EPS production could represent a selective advantage for Chloroflexiin WWTS, probably having an important role in the formation of the initial framework for small sludge particles, serving as cores or carriers in all reactors.

While MAGs are required to determine the diversity and potential function of the phylum Chloroflexi, further experiments are necessary to confirm the identified metabolic functions. One of the possible options would be to focus efforts on the sequencing of metatranscriptomics together with metagenomes. This would allow to confirm which of the potential metabolic functions determined by the annotation of the MAGs, are being expressed under certain operational conditions. On the other hand, the study using these molecular tools in bulking episodes would expand the knowledge about the role of Chloroflexi in these events. This information is necessary for a better management of operational parameters, to reduce and control bulking events. This knowledge makes it possible to determine which groups are functionally important in these systems and how their disappearance or overgrowth would affect the process.

The difficulty of isolating organisms of the phylum Chloroflexi are probably related to their slow growth. Competition with other fast growing fermenting microorganisms is probably one of the greatest barriers in obtaining good primary enrichments. Recently, new tools have been developed that facilitate the isolation of slow-growing microorganisms such as Droplet-based high-throughput cultivation (Watterson et al. 2020). Therefore, it is expected in the coming years to have a greater quantity of organisms of the phylum Chloroflexi isolated in pure culture, in order to reveal and confirm what they are doing in the wastewater treatment systems.

4.4 References

Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 31:533–538 . doi: 10.1038/nbt.2579

Ali M, Shaw DR, Albertsen M, Saikaly PE (2020) Comparative Genome-Centric Analysis of Freshwater and Marine ANAMMOX Cultures Suggests Functional Redundancy in Nitrogen Removal Processes. Front Microbiol 11:1–16 . doi: 10.3389/fmicb.2020.01637

Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169 . doi: 10.1016/j.jip.2007.09.009

Andersen KS, Kirkegaard RH, Karst SM, Albertsen M (2018a) ampvis2: An R package to analyse and visualise 16S rRNA amplicon data. bioRxiv 299537

Andersen MH, McIlroy SJ, Nierychlo M, Nielsen PH, Albertsen M (2018b) Genomic insights into *Candidatus* Amarolinea aalborgensis gen. nov., sp. nov., associated with settleability problems in wastewater treatment plants. Syst Appl Microbiol 1–8 . doi: 10.1016/j.syapm.2018.08.001

Andrews (2010) FastQC: a quality control tool for high throughput sequence data. Babraham Inst

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477 . doi: 10.1089/cmb.2012.0021

Björnsson L, Hugenholtz P, Tyson GW, Blackall LL (2002a) Filamentous Chloroflexi (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. Microbiology 148:2309–2318 . doi: 10.1099/00221287-148-8-2309

Björnsson L, Hugenholtz P, Tyson GW, Blackall LL (2002b) Filamentous Chloroflexi (green non-sulfur bacteria) are abundant in wastewater treatment processes with biological nutrient removal. Microbiology 148:2309–2318

Blazejak A, Schippers A (2010) High abundance of JS-1- and Chloroflexi-related Bacteria in deeply buried marine sediments revealed by quantitative, real-time PCR. FEMS Microbiol Ecol 72:198–207 . doi: 10.1111/j.1574-6941.2010.00838.x

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120 . doi: 10.1093/bioinformatics/btu170

Borzacconi L, López I, Passeggi M, Etchebehere C, Barcia R (2008) Sludge deterioration in a full scale UASB reactor after a pH drop working under low loading conditions. Water Sci Technol 57:797–802 . doi: 10.2166/wst.2008.071

Bovio-winkler P, Cabezas A, Etchebehere C (2021) Database Mining to Unravel the Ecology of the Phylum Chloroflexi in Methanogenic Full Scale Bioreactors. 11: . doi: 10.3389/fmicb.2020.603234

Bovio P, Cabezas A, Etchebehere C (2019) Preliminary analysis of Chloroflexi populations in full-scale UASB methanogenic reactors. J Appl Microbiol 126:667–683 . doi: 10.1111/jam.14115

Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloe-Fadrosh EA, Tringe SG, Ivanova NN, Copeland A, Clum A, Becraft ED, Malmstrom RR, Birren B, Podar M, Bork P, Weinstock GM, Garrity GM, Dodsworth JA, Yooseph S, Sutton G, Glöckner FO, Gilbert JA, Nelson WC, Hallam SJ, Jungbluth SP, Ettema TJG, Tighe S, Konstantinidis KT, Liu WT, Baker BJ, Rattei T, Eisen JA, Hedlund B, McMahon KD, Fierer N, Knight R, Finn R, Cochrane G, Karsch-Mizrachi I, Tyson GW, Rinke C, Lapidus A, Meyer F, Yilmaz P, Parks DH, Eren AM, Schriml L, Banfield JF, Hugenholtz P, Woyke T (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. 35:725–731

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583 . doi: 10.1038/nmeth.3869

Campanaro S, Treu L, Rodriguez-R LM, Kovalovszki A, Ziels RM, Maus I, Zhu X, Kougias PG, Basile A, Luo G, Schlüter A, Konstantinidis KT, Angelidaki I (2020) New insights from the biogas microbiome by comprehensive genome-resolved metagenomics of nearly 1600 species originating from multiple anaerobic digesters. Biotechnol Biofuels 13:1–18 . doi: 10.1186/s13068-020-01679-y

Campbell AG, Schwientek P, Vishnivetskaya T, Woyke T, Levy S, Beall CJ, Griffen A, Leys E, Podar M (2014) Diversity and genomic insights into the uncultured Chloroflexi from the human microbiota. Environ Microbiol 16:2635–2643 . doi: 10.1111/1462-2920.12461

Cao S, Du R, Li B, Ren N, Peng Y (2016) High-throughput profiling of microbial community structures in an ANAMMOX-UASB reactor treating high-strength wastewater. Appl Microbiol Biotechnol 100:6457–6467 . doi: 10.1007/s00253-016-7427-6

Cho S, Fujii N, Lee T, Okabe S (2011) Development of a simultaneous partial nitrification and anaerobic ammonia oxidation process in a single reactor. Bioresour Technol 102:652–659 . doi: 10.1016/j.biortech.2010.08.031

Chu Z rui, Wang K, Li X kun, Zhu M ting, Yang L, Zhang J (2015) Microbial characterization of aggregates within a one-stage nitritation-anammox system using high-throughput amplicon sequencing. Chem Eng J 262:41–48 . doi: 10.1016/j.cej.2014.09.067

Claesson MJ, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS One 4: . doi: 10.1371/journal.pone.0006669

Dombrowski N, Seitz KW, Teske AP, Baker BJ (2017) Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. Microbiome 5:106 . doi: 10.1186/s40168-017-0322-2

Dong X, Strous M (2019) An Integrated Pipeline for Annotation and Visualization of Metagenomic Contigs. Front Genet 10: . doi: 10.3389/fgene.2019.00999

Gao D, Liu L, Liang H, Wu WM (2011) Aerobic granular sludge: Characterization, mechanism of granulation and application to wastewater treatment. Crit. Rev. Biotechnol. 31:137–152

Gelfand MS, Rodionov DA (2008) Comparative genomics and functional annotation of bacterial transporters. Phys. Life Rev. 5:22–49

Gich F, Garcia-Gil J, Overmann J (2002) Previously unknown and phylogenetically diverse members of the green nonsulfur bacteria are indigenous to freshwater lakes. Arch Microbiol 177:1–10 . doi: 10.1007/s00203-001-0354-6

Goehring NW, Beckwith J (2005) Diverse paths to midcell: Assembly of the bacterial cell division machinery. Curr. Biol. 15:R514–R526

Gonzalez-Gil G, Sougrat R, Behzad AR, Lens PNL, Saikaly PE (2015) Microbial Community Composition and Ultrastructure of Granules from a Full-Scale Anammox Reactor. Microb Ecol 70:118–131 . doi: 10.1007/s00248-014-0546-7

Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR, Williams KH, Tringe SG, Banfield JF (2013) Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. Microbiome 1:22 . doi: 10.1186/2049-2618-1-22

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010) Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11: . doi: 10.1186/1471-2105-11-119

Isanta E, Bezerra T, Fernández I, Suárez-Ojeda ME, Pérez J, Carrera J (2015) Microbial community shifts on an anammox reactor after a temperature shock using 454-pyrosequencing analysis. Bioresour Technol 181:207–213 . doi: 10.1016/j.biortech.2015.01.064

Johnson LS, Eddy SR, Portugaly E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics 11:431 . doi: 10.1186/1471-2105-11-431

Juan-Díaz X, Pérez J, Carrera J (2021) Effective dampening of temperature effects in an anammox reactor treating real mainstream wastewater. J Water Process Eng 40:101853 . doi: 10.1016/j.jwpe.2020.101853

Kang DD, Froula J, Egan R, Wang Z (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 2015:1–15 . doi: 10.7717/peerj.1165

Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y (2013) Ardenticatena maritima gen. nov., sp. nov., a ferric iron- and nitrate-reducing bacterium of the phylum "Chloroflexi" isolated from an iron-rich coastal hydrothermal field, and description of Ardenticatenia classe s nov. Int J Syst Evol Microbiol 63:2992–3002 . doi: 10.1099/ijs.0.046532-0

Kawaichi S, Yoshida T, Sako Y, Nakamuraa R (2015) Draft Genome Sequence of a Heterotrophic Facultative Anaerobic Thermophilic Bacterium, Ardenticatena maritima Strain 110S. 3:1–2 . doi: 10.1128/genomeA.01145-15.Copyright

Kindaichi T, Ito T, Okabe S (2004) Ecophysiological Interaction between Nitrifying Bacteria and Heterotrophic Bacteria in Autotrophic Nitrifying Biofilms as Determined by Microautoradiography-Fluorescence *In situ* Hybridization. Appl Environ Microbiol 70:1641–1650 . doi: 10.1128/AEM.70.3.1641-1650.2004

Kindaichi T, Nierychlo M, Kragelund C, Nielsen JL, Nielsen PH (2013) High and stable substrate specificities of microorganisms in enhanced biological phosphorus removal plants. Environ Microbiol 15:1821–1831 . doi: 10.1111/1462-2920.12074

Kindaichi T, Yuri S, Ozaki N, Ohashi A (2012) Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor. Water Sci Technol 66:2556–2561 . doi: 10.2166/wst.2012.479

Konstantinidis KT, Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. J Bacteriol 187:6258–6264 . doi: 10.1128/JB.187.18.6258-6264.2005

Kragelund C, Levantesi C, Borger A, Thelen K, Eikelboom D, Tandoi V, Kong Y, Van Der Waarde J, Krooneman J, Rossetti S, Thomsen TR, Nielsen PH (2007) Identity, abundance and ecophysiology of filamentous Chloroflexi species present in activated sludge treatment plants. FEMS Microbiol Ecol 59:671–682 . doi: 10.1111/j.1574-6941.2006.00251.x

Kragelund C, Thomsen TR, Mielczarek AT, Nielsen PH (2011) Eikelboom's morphotype 0803 in activated sludge belongs to the genus Caldilinea in the phylum Chloroflexi. FEMS Microbiol Ecol 76:451–462 . doi: 10.1111/j.1574-6941.2011.01065.x

Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35:1547–1549 . doi: 10.1093/molbev/msy096

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359 . doi: 10.1038/nmeth.1923

Lawson CE, Wu S, Bhattacharjee AS, Hamilton JJ, McMahon KD, Goel R, Noguera DR (2017) Metabolic network analysis reveals microbial community interactions in anammox granules. Nat Commun 8:1–12 . doi: 10.1038/ncomms15416

Letunic I, Bork P (2019) Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res 47:W256–W259 . doi: 10.1093/nar/gkz239

Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW (2016) MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102:3–11 . doi: 10.1016/j.ymeth.2016.02.020

Li J, Hu B, Zheng P, Qaisar M, Mei L (2008) Filamentous granular sludge bulking in a laboratory scale UASB reactor. Bioresour Technol 99:3431–3438 . doi: 10.1016/j.biortech.2007.08.005

Margolin W (2000) Themes and variations in prokaryotic cell division. FEMS Microbiol. Rev. 24:531–548

McGonigle JM, Lang SQ, Brazelton WJ (2020) Genomic Evidence for Formate Metabolism by Chloroflexi as the Key to Unlocking Deep Carbon in Lost City Microbial Ecosystems. Appl Environ Microbiol 1–12 . doi: 10.1128/aem.02583-19

McIlroy SJ, Karst SM, Nierychlo M, Dueholm MS, Albertsen M, Kirkegaard RH, Seviour RJ, Nielsen PH (2016) Genomic and *in situ* investigations of the novel uncultured Chloroflexi associated with 0092 morphotype filamentous bulking in activated sludge. Isme J (In press):1–12 . doi: 10.1038/ismej.2016.14

McIlroy SJ, Kirkegaard RH, Dueholm MS, Fernando E, Karst SM, Albertsen M, Nielsen PH (2017) Culture-independent analyses reveal novel anaerolineaceae as abundant primary fermenters in anaerobic digesters treating waste activated sludge. Front Microbiol 8: . doi: 10.3389/fmicb.2017.01134

Mcilroy SJ, Starnawska A, Starnawski P, Saunders AM, Nierychlo M, Nielsen PH, Nielsen JL (2016) Identification of active denitrifiers in full-scale nutrient removal wastewater treatment systems. Environ Microbiol 18:50–64 . doi: 10.1111/1462-2920.12614

McMurdie PJ, Holmes S (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8:e61217 . doi: 10.1371/journal.pone.0061217

Mielczarek AT, Kragelund C, Eriksen PS, Nielsen PH (2012) Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. Water Res 46:3781–3795 . doi: 10.1016/j.watres.2012.04.009

Miura Y, Okabe S (2008) Quantification of cell specific uptake activity of microbial products by uncultured Chloroflexi by microautoradiography combined with fluorescence *in situ* hybridization. Environ Sci Technol 42:7380–7386 . doi: 10.1021/es800566e

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press

Nielsen PH, Kragelund C, Seviour RJ, Nielsen JL (2009) Identity and ecophysiology of filamentous bacteria in activated sludge. FEMS Microbiol. Rev. 33:969–998

Nierychlo M, Andersen KS, Xu Y, Green N, Jiang C, Albertsen M, Dueholm MS, Nielsen PH (2020a) MiDAS 3: An ecosystem-specific reference database, taxonomy and knowledge platform for activated sludge and anaerobic digesters reveals species-level microbiome composition of activated sludge. Water Res 115955 . doi: 10.1016/j.watres.2020.115955

Nierychlo M, McIlroy SJ, Kucheryavskiy S, Jiang C, Ziegler AS, Kondrotaite Z, Stokholm-Bjerregaard M, Nielsen PH (2020b) *Candidatus* Amarolinea and *Candidatus* Microthrix Are Mainly Responsible for Filamentous Bulking in Danish Municipal Wastewater Treatment Plants. Front Microbiol 11:1–17 . doi: 10.3389/fmicb.2020.01214

Nierychlo M, Miłobędzka A, Petriglieri F, McIlroy B, Nielsen PH, McIlroy SJ (2019) The morphology and metabolic potential of the Chloroflexi in full-scale activated sludge wastewater treatment plants. FEMS Microbiol Ecol 95:1–11 . doi: 10.1093/femsec/fiy228

Okabe S, Kindaichi T, Ito T (2005) Fate of 14C-labeled microbial products derived from nitrifying bacteria in autotrophic nitrifying biofilms. Appl Environ Microbiol 71:3987–3994 . doi: 10.1128/AEM.71.7.3987-3994.2005

Onetto CA, Grbin PR, McIlroy SJ, Eales KL (2019) Genomic insights into the metabolism of “*Candidatus* Defluviicoccus seviourii”, a member of Defluviicoccus cluster III abundant in industrial activated sludge. FEMS Microbiol Ecol 95:1–12 . doi: 10.1093/femsec/fiy231

Orschler L, Agrawal S, Lackner S (2021) Targeted metagenomics reveals extensive diversity of the denitrifying community in partial nitritation anammox and activated sludge systems. Biotechnol Bioeng 118:433–441 . doi: 10.1002/bit.27581

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 36:996 . doi: 10.1038/nbt.4229

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 25:1043–1055 . doi: 10.1101/gr.186072.114

Pereira AD, Cabezas A, Etchebehere C, Chernicharo CA de L, de Araújo JC (2017) Microbial communities in anammox reactors: a review. Environ Technol Rev 6:74–93 . doi: 10.1080/21622515.2017.1304457

Petriglieri F, Nierychlo M, Nielsen PH, McIlroy SJ (2018) *In situ* visualisation of the abundant Chloroflexi populations in full-scale anaerobic digesters and the fate of immigrating species. PLoS One 13:1–14 . doi: 10.1371/journal.pone.0206255

Pichoff S, Lutkenhaus J (2002) Unique and overlapping roles for ZipA and FtsA in septal ring assembly in Escherichia coli. EMBO J 21:685–693 . doi: 10.1093/emboj/21.4.685

R Core Team (2020) R Core Team. R A Lang. Environ. Stat. Comput. R Found. Stat. Comput. Vienna, Austria.

Reino C, Carrera J (2017) Low-strength wastewater treatment in an anammox UASB reactor: Effect of the liquid upflow velocity. Chem Eng J 313:217–225 . doi: 10.1016/j.cej.2016.12.051

Schmidt I, Sliekers O, Schmid M, Bock E, Fuerst J, Kuenen JG, Jetten MSM, Strous M (2003) New concepts of microbial treatment processes for the nitrogen removal in wastewater. FEMS Microbiol Rev 27:481–492 . doi: 10.1016/S0168-6445(03)00039-1

Seemann T (2014) Genome analysis Prokka: rapid prokaryotic genome annotation. 30:2068–2069 . doi: 10.1093/bioinformatics/btu153

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001) *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl Environ Microbiol 67:5740–5749 . doi: 10.1128/AEM.67.12.5740-5749.2001

Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H, Kamagata Y (2003a) Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. Int J Syst Evol Microbiol 53:1843–1851 . doi: 10.1099/ijs.0.02699-0

Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H, Kamagata Y (2003b) Anaerolinea thermophila gen. nov., sp. nov. and Caldilinea aerophila gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. Int J Syst Evol Microbiol 53:1843–1851 . doi: 10.1099/ijs.0.02699-0

Song YX, Liao Q, Yu C, Xiao R, Tang CJ, Chai LY, Duan CS (2017) Physicochemical and microbial properties of settled and floating anammox granules in upflow reactor. Biochem Eng J 123:75–85 . doi: 10.1016/j.bej.2017.04.002

Speirs LBM, Dyson ZA, Tucci J, Seviour RJ (2017) Eikelboom filamentous morphotypes 0675 and 0041 embrace members of the Chloroflexi: resolving their phylogeny, and design of fluorescence *in situ* hybridisation probes for their identification. FEMS Microbiol Ecol 93:1–13 . doi: 10.1093/femsec/fix115

Speth DR, In'T Zandt MH, Guerrero-Cruz S, Dutilh BE, Jetten MSM (2016) Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. Nat Commun 7: . doi: 10.1038/ncomms11172

Strous M, Heijnen JJ, Kuenen JG, Jetten MSM (1998) The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. Appl Microbiol Biotechnol 50:589–596 . doi: 10.1007/s002530051340

Strous M, Pelletier E, Mangenot S, Rattei T, Lehner A, Taylor MW, Horn M, Daims H, Bartol-Mavel D, Wincker P, Barbe V, Fonknechten N, Vallenet D, Segurens B, Schenowitz-Truong C, Médigue C, Collingro A, Snel B, Dutilh BE, Op Den Camp HJM, Van Der Drift C, Cirpus I, Van De Pas-Schoonen KT, Harhangi HR, Van Niftrik L, Schmid M, Keltjens J, Van De Vossenberg J, Kartal B, Meier H, Frishman D, Huynen MA, Mewes HW, Weissenbach J, Jetten MSM, Wagner M, Le Paslier D (2006) Deciphering the evolution and metabolism of an anammox bacterium from a community genome. Nature 440:790–794 . doi: 10.1038/nature04647

Sun L, Toyonaga M, Ohashi A, Matsuura N, Tourlousse DM, Meng X, Tamaki H, Hanada S, Cruz R, Yamaguchi T, Sekiguchi Y (2016) Isolation and characterization of Flexilinea flocculi gen . nov ., sp . nov ., a filamentous , anaerobic bacterium belonging to the class Anaerolineae in the phylum Chloroflexi. Int J Syst Evol Microbiol 66:988–996 . doi: 10.1099/ijsem.0.000822

Ultsch A, Mörchen F (2005) ESOM-Maps: tools for clustering, visualization, and classification with Emergent SOM. Tech Rep Dept Math Comput Sci Univ Marburg, Ger 1–7

Wang S, Guo J, Lian J, Ngo HH, Guo W, Liu Y, Song Y (2016) Rapid start-up of the anammox process by denitrifying granular sludge and the mechanism of the anammox electron transport chain. Biochem Eng J 115:101–107 . doi: 10.1016/j.bej.2016.09.001

Ward LM, Hemp J, Shih PM, McGlynn SE, Fischer WW (2018) Evolution of phototrophy in the Chloroflexi phylum driven by horizontal gene transfer. Front Microbiol 9:1–16 . doi: 10.3389/fmicb.2018.00260

Wasmund K, Schreiber L, Lloyd KG, Petersen DG, Schramm A, Stepanauskas R, Jørgensen BB, Adrian L (2014) Genome sequencing of a single cell of the widely distributed marine subsurface Dehalococcoidia, phylum Chloroflexi. ISME J 8:383–397 . doi: 10.1038/ismej.2013.143

Watterson WJ, Tanyeri M, Watson AR, Cham CM, Shan Y, Chang EB, Eren AM, Tay S (2020) Droplet-based high-throughput cultivation for accurate screening of antibiotic resistant gut microbes. Elife 9:1–22 . doi: 10.7554/eLife.56998

Xia Y, Kong Y, Nielsen PH (2007) *In situ* detection of protein-hydrolysing microorganisms in activated sludge. FEMS Microbiol Ecol 60:156–165 . doi: 10.1111/j.1574-6941.2007.00279.x

Xia Y, Wang Y, Wang Y, Chin FYL, Zhang T (2016) Cellular adhesiveness and cellulolytic capacity in Anaerolineae revealed by omics-based genome interpretation. Biotechnol Biofuels 9:1–13 . doi: 10.1186/s13068-016-0524-z

Yamada T, Imachi H, Ohashi A, Harada H, Hanada S, Kamagata Y, Sekiguchi Y (2007a) Bellilinea caldifistulae gen. nov., sp. nov and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum Chloroflexi isolated from methanogenic propionate-degrading consortia. Int J Syst Evol Microbiol 57:2299–2306 . doi: 10.1099/ijs.0.65098-0

Yamada T, Imachi H, Ohashi A, Harada H, Hanada S, Kamagata Y, Sekiguchi Y (2007b) Bellilinea caldifistulae gen. nov., sp. nov and Longilinea arvoryzae gen. nov., sp. nov., strictly anaerobic, filamentous bacteria of the phylum Chloroflexi isolated from methanogenic propionate-degrading consortia. Int J Syst Evol Microbiol 57:2299–2306 . doi: 10.1099/ijs.0.65098-0

Yamada T, Sekiguchi Y (2009) Cultivation of uncultured Chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi "subphylum i" with natural and biotechnological relevance. Microbes Environ 24:205–216 . doi: 10.1264/jsme2.ME09151S

Yamada T, Sekiguchi Y, Hanada S, Imachi H, Ohashi A, Harada H, Kamagata Y (2006) Anaerolinea thermolimosa sp. nov., Levilinea saccharolytica gen. nov., sp. nov. and Leptolinea tardivitalis gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes Anaerolineae classis nov. and Caldilineae classis nov. in the . Int J Syst Evol Microbiol 56:1331–1340 . doi: 10.1099/ijs.0.64169-0

Yamada T, Sekiguchi Y, Imachi H, Kamagata Y, Ohashi A, Harada H (2005) Diversity, localization, and physiological properties of filamentous microbes belonging to Chloroflexi subphylum I in mesophilic and thermophilic methanogenic sludge granules. Appl Environ Microbiol 71:7493–7503 . doi: 10.1128/AEM.71.11.7493-7503.2005

Yoon DN, Park SJ, Kim SJ, Jeon CO, Chae JC, Rhee SK (2010) Isolation, characterization, and abundance of filamentous members of Caldilineae in activated sludge. J Microbiol 48:275–283 . doi: 10.1007/s12275-010-9366-8

Zang K, Kurisu F, Kasuga I, Furumai H, Yagi O (2008) Analysis of the phylogenetic diversity of estrone-degrading bacteria in activated sewage sludge using microautoradiography-fluorescence *in situ* hybridization. Syst Appl Microbiol 31:206–214 . doi: 10.1016/j.syapm.2008.03.005

Zhao Y, Jiang B, Tang X, Liu S (2019) Metagenomic insights into functional traits variation and coupling effects on the anammox community during reactor start-up. Sci Total Environ 687:50–60 . doi: 10.1016/j.scitotenv.2019.05.491

Zhao Y, Liu S, Jiang B, Feng Y, Zhu T, Tao H, Tang X, Liu S (2018) Genome-Centered Metagenomics Analysis Reveals the Symbiotic Organisms Possessing Ability to Cross-Feed with Anammox Bacteria in Anammox Consortia

Zhuang WQ, Yi S, Bill M, Brisson VL, Feng X, Men Y, Conrad ME, Tang YJ, Alvarez-Cohen L (2014) Incomplete Wood-Ljungdahl pathway facilitates one-carbon metabolism in organohalide-respiring Dehalococcoides mccartyi. Proc Natl Acad Sci U S A 111:6419–6424 . doi: 10.1073/pnas.1321542111

Pati, A., LaButti, K., Pukall, R., Nolan, M., Del Rio, T. G., Tice, H., ... & Ivanova, N. 2010. Complete genome sequence of Sphaerobacter thermophilus type strain (S 6022 T). Standards in genomic sciences, 2(1), 49-56.

Takami H, Taniguchi T, Moriya Y, Kuwahara T, Kanehisa M, Goto S. Evaluation method for the potential functionome harbored in the genome and metagenome. BMC Genomics 2012; 13: 699. doi: 10. 1186/1471-2164-13-699 PMID: 23234305.

Takami H. New method for comparative functional genomics and metagenomics using KEGG module. In: Nelson K, editor. Encyclopedia of metagenomics. Berlin, Heidelberg: Springer-Verlag. 2014; 525–539.

Nikitin F, Rance B, Itoh M, Kanehisa M, Lisacek F. Using protein motif combinations to update KEGG pathway maps and orthologue tables. Genome Inform. 2004; 15: 266–275. PMID: 15706512.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and environment. Nucleic Acids Res. 2008; 36: D480–D484. PMID: 18077471.

Gonzalez-Martinez, A. et al. Microbial community analysis of a full-scale DEMON bioreactor. Bioprocess Biosyst. Eng. 38, 499–508 (2015)

Gonzalez-Gil, G., Sougrat, R., Behzad, A. R., Lens, P. N. L. & Saikaly, P. E. Microbial community composition and ultrastructure of granules from a full-scale anammox reactor. Microb. Ecol. 70, 118–131 (2014).

Li, X. R. et al. The bacterial diversity in an anaerobic ammonium-oxidizing (anammox) reactor community. Syst. Appl. Microbiol. 32, 278–289 ð2009Þ

Zere et al. 2015 Genomic Targets and Features of BarA-UvrY (-SirA) Signal Transduction Systems

Pernestig et al. 2003. The Escherichia coli BarA-UvrY Two-Component System Is Needed for Efficient Switching between Glycolytic and Gluconeogenic Carbon Sources

4.5 Material suplementario

Supplementary

Figure S1. Annotated genes related to rhe WLP pathway

A

		AMX56	AMX47	AMX68	AMX9	AMX39	AMX14	AMX55	AMX57	AMX15	RH52	RH43	RH38	RH21	MO16	MO118	MO66	MO53
	Transporters																	
Formate/nitrite transporter	focA COG2116 K06212	■		■		■	■	■	■	■	■	■		■	■		■	■
Nitrate/nitrite antiporter	narK, nrtP, nrt K02575																■	
Ammonium transporter	amtb K03320	■	■						■			■	■	■			■	■
Ammonium transporter	nrgA	■	■									■					■	
	Dissimilatory nitrate reduction DNRA																	
Nitrate to Nitrite	NarGHI K00370 + K00371 + K00374	■	■			■	■	■		■							■	
	NapAB K02567 + K02568																	
Nitrite to Ammonia	NirBD K00362 + K00363											■						■
	NrfAH K03385 + K15876	■		■	■		■	■	■			■	■		■			
	Denitrification																	
Nitrate to Nitrite	NarGHI K00370 + K00371 + K00374	■	■			■	■	■		■							■	
	NapAB K02567 + K02568																	
Nitrite to Nitric oxide	NirK K00368			■	■	■				■				■				
	NirS K15864			■			■	■							■			
Nitric oxide to Nitrous oxide	NorBC K04561 +K02305								■				■	■				
Nitrous oxide to Nitrogen	NosZ K00376	■	■		■				■		■	■	■					
	Nitrification																	
Ammonia to hydroxylamine	AmoCAB																	
Hydroxylamine to Nitrite	Hao K10535										■			■				
Nitrite to Nitrate	NxrAB K00370 + K00371	■	■			■	■	■		■							■	

B

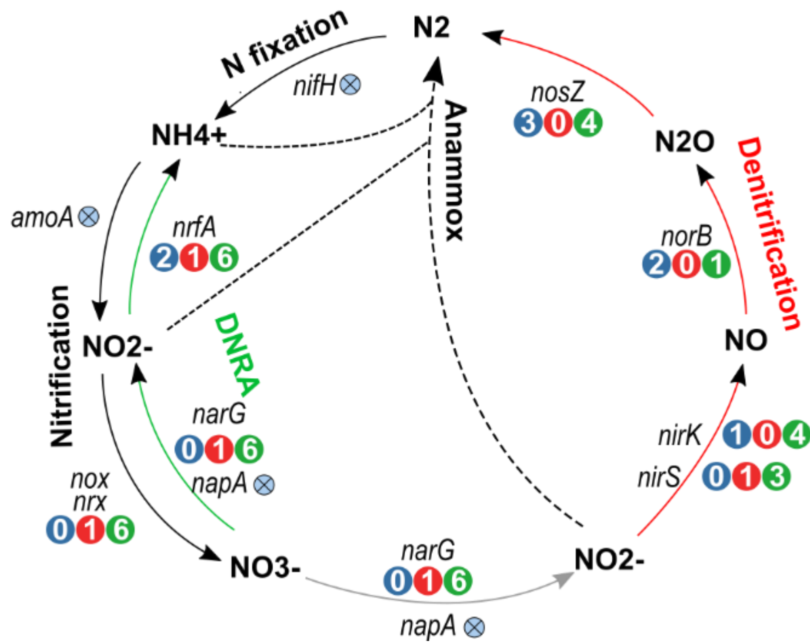

Figure S2. Annotated genes related to the nitrogen cycle

CAPÍTULO 5

DISCUSIÓN GENERAL, CONCLUSIONES Y PERSPECTIVAS

5.1 ¿Por qué es importante estudiar organismos del filo Chloroflexi sistemas de tratamiento de aguas residuales?

A pesar de la gran cantidad de investigaciones sobre la estructura y función de las comunidades microbianas en sistemas de tratamiento de aguas residuales, se siguen operando como cajas negras. Esto implica que si bien se conoce la composición de las comunidades microbianas complejas que llevan adelante los procesos de degradación de materia orgánica y otros contaminantes, no se conoce quiénes llevan adelante cada proceso, en qué condiciones, de qué manera y cómo interactúan Si bien parecen ser sistemas robustos y tolerantes a variaciones en parámetros operacionales como pH, concentración de materia orgánica, temperatura, etc., frecuentemente ocurren problemas operacionales relacionados a la microbiología de los procesos. En particular, en sistemas aerobios de lodos activados, el bulking debido al sobrecrecimiento de bacterias filamentosas es un problema frecuentemente detectado. El interés en el estudio del filo Chloroflexi en diferentes sistemas de tratamiento de aguas residuales y de residuos sólidos radica en que es un grupo de bacterias filamentosas que presenta una alta abundancia, pero se sabe muy poco sobre su rol funcional. Dentro de este filo, la clase Anaerolineae es la predominante, y presenta una gran diversidad revelada mediante distintas técnicas de biología molecular. El aislamiento en cultivo puro de algunas especies de la clase Anaerolineae ha permitido conocer las condiciones óptimas de crecimiento de estos organismos. Es así que para las 6 especies de la clase Anaerolineae aisladas de reactores anaerobios, se ha podido determinar que su rol se centraría en la fermentación de compuestos orgánicos, con una potencial sintrofía con metanógenos, que aún no ha sido demostrada.

Si bien el aislamiento es una herramienta muy poderosa para determinar condiciones óptimas de crecimiento y prever condiciones de sobrecrecimiento, no se debe olvidar que en un biorreactor están interaccionando comunidades complejas de microorganismos. Esto implica la presencia de productos metabólicos secundarios provenientes de otros microorganismos y otras interacciones microbianas, que en cultivo puro no pueden ser evaluadas. Por lo tanto, la respuesta funcional real de estos aislamientos a partir de determinaciones en cultivo puro, podría no ser tan clara en los biorreactores, y por ende limitar nuestro entendimiento. De la misma manera ocurriría con su función estructural en gránulos y flóculos.

A través de las nuevas técnicas de biología molecular, como la metagenómica, es posible la recuperación de genomas casi completos y con muy baja contaminación. A partir de estos genomas es posible inferir en el potencial metabólico, pero con la limitante de que no

sabemos realmente qué vías metabólicas estarían expresando *in situ*. Es por esto que se deberían aplicar además otras técnicas para determinar qué genes están expresando, como por ejemplo metatranscriptómica, FISH-MAR o SIP (Stable Isotope Probing), para entender qué función cumplen *in situ* estos microorganismos.

La importancia del estudio de la diversidad, composición taxonómica del filo Chloroflexi y la función cumplen en los biorreactores, radica en poder mantener la estabilidad operacional de estos sistemas, con una buena eficiencia de remoción de la materia orgánica y contaminantes. El conocimiento del rol metabólico de Chloroflexi así como de otros organismos poco conocidos permitiría mantener las condiciones operacionales que mejoren el proceso. En particular, como se ha demostrado que los organismos del filo Chloroflexi tienen un rol en el desarrollo estructural de los gránulos o flóculos y en el problema de bulking, el conocer las condiciones en las cuales se puede evitar el sobrecrecimiento de estas bacterias filamentosas, podría ayudar a controlar este gran problema que causa grandes pérdidas económicas y contaminación ambiental.

5.2 Los organismos del filo Chloroflexi son diferentes pero tienen la misma función en los diferentes tipos de reactores, ¿que puede implicar esto?

Como lo indican diversos estudios en sistemas de tratamiento de aguas residuales, los biorreactores con una baja diversidad microbiana tienen también una menor redundancia funcional (Wittebolle et al. 2009). Por lo tanto, los biorreactores con una menor redundancia funcional son más propensos a fallas en el proceso frente a perturbaciones, como pueden ser cambios en la carga orgánica, temperatura, pH, etc. Dicho de otro modo, biorreactores con una mayor diversidad, albergan una redundancia funcional mayor, lo que le confiere una mayor robustez al sistema. Resulta entonces imprescindible conocer qué grupos microbianos son funcionalmente importantes para optimizar las condiciones operacionales que los favorecen. La redundancia funcional permite la resistencia funcional de cualquier ecosistema a las perturbaciones ambientales debido a la presencia de múltiples especies que pueden realizar la misma función metabólica, de tal manera que la pérdida de una especie debido a la perturbación será sustituida por otra especie de la comunidad. En este sentido, mediante el ensamblado de genomas a partir de metagenomas en los reactores aerobios, anaerobios y anammox, recuperamos una gran diversidad de especies de la clase Anaerolineae en cada

reactor. Al determinar su potencial metabólico, encontramos que cumplían el mismo rol funcional. Cada una de estas especies presentó el potencial de hidrolizar distintos compuestos complejos, metabolizarlos vía glicólisis o beta-oxidación y posteriormente fermentarlos, produciendo distintos compuestos, entre ellos hidrógeno y acetato. Por lo tanto el rol principal de estos organismos sería la hidrólisis y fermentación de materia orgánica posiblemente proveniente de restos celulares. Esta función de degradar la materia orgánica de restos celulares explicaría porqué están en todos los biorreactores con alta densidad celular. En este trabajo también se observó que su participación en el ciclo del nitrógeno fue particularmente más importante en los reactores anammox, este hecho también se explica con cierta especialización de estos organismos según el nicho. En cada uno de los sistemas, la desaparición de una determinada especie de Anaerolineae podría ser sustituida funcionalmente por otra de la misma clase. Como se discutió anteriormente, no alcanza con saber qué están haciendo individualmente los diferentes organismos en estos sistemas sino cómo interaccionan con otros microorganismos y cual es su redundancia funcional. En el caso de que las perturbaciones no afectasen a estos microorganismos en particular, hay que tener en cuenta que la afectación de otras especies que están interrelacionadas, podría afectar indirectamente a los microorganismos de interés. Por lo tanto, sabiendo además como se interrelacionan los microorganismos en el sistema, podremos saber cómo afectan las perturbaciones directa o indirectamente.

5.3 La gran mayoría de los organismos del filo Chloroflexi descripto en sistemas de tratamiento de aguas residuales tiene morfología filamentosa, ¿Esto tiene una importancia especial en estos sistemas?

Los reactores de gránulos tienen gran aplicación en sistemas de tratamiento de aguas residuales. En general este tipo de reactores está más desarrollado para sistemas anaerobios metanogénicos y sistemas de remoción de nitrógeno tipo anammox. La gran ventaja de estos sistemas donde la biomasa se agrupa formando gránulos es que permite retener la biomasa dentro del reactor y mejora la interacción entre los microorganismos. El estudio de la estructura de los gránulos en sistemas metanogénicos ha mostrado una distribución espacial de distintos grupos de microorganismos (Sekiguchi et al. 1999) mostrando que los microorganismos que hibridan con la sonda de Chloroflexi están en la superficie de los gránulos. Este hallazgo concuerda con el hecho de que son microorganismos hidrolíticos y por lo tanto deberían estar en la parte expuesta a los compuestos a hidrolizar. Por otro lado, se

ha sugerido que pueden tener un rol importante en la formación de gránulos. Aunque el proceso de formación de los gránulos se ha estudiado intensamente, la literatura suele ser contradictoria sobre los mecanismos responsables (Hulshoff Pol et al. 2004b). Probablemente, varios factores contribuyan a la granulación y estén involucrados múltiples mecanismos (Ahn et al. 2009). Una de las hipótesis más interesantes es la "teoría de los spaghetti" (Wiegant and de Man 1986) donde la formación de gránulos se explicaría por la presencia de Archaeas metanogénicas acetoclásticas filamentosas. Posteriormente, los organismos del filo Chloroflexi también fueron propuestos como candidato a participar en la formación de la matriz inicial de gránulos y flóculos, y en especial la clase Anaerolineae (Sekiguchi et al. 1999; Sekiguchi et al. 2001b; Yamada and Sekiguchi 2009; Nielsen et al. 2009). Pero no solo se relaciona al filo Chloroflexi en la formación de gránulos y flóculos, sino que también con episodios de bulking en sistemas aerobios y anaerobios (Yamada and Sekiguchi 2009; Nierychlo et al. 2019b). De hecho, *Anaerolinea thermophila* fue aislada de gránulos de los cuales existían proyecciones de bacterias filamentosas (Yamada and Sekiguchi 2009). Es por esto, que la clase Anaerolineae de la cual la totalidad de sus miembros detectados tiene morfología filamentosa, tiene especial importancia tanto en reactores que contienen biomasa granular como en flóculos.

En resumen, de acuerdo a lo expuesto, los organismos del filo Chloroflexi tienen un rol fundamental en la conformación de los gránulos y flóculos y en mantener su estructura. Existe cierto equilibrio entre la proporción de estos organismos en la comunidad y la buena sedimentación de la biomasa, abundancias muy bajas pueden llevar a gránulos o flóculos poco estables, y abundancias muy altas pueden llevar a problemas de sedimentación de la biomasa (episodios de bulking filamentoso). Como se puede regular la abundancia de estos microorganismos es una de las preguntas que queda aún sin responder.

5.4 ¿Por qué es tan difícil aislar organismos del filo Chloroflexi?

El filo Chloroflexi es considerado uno de los filos que componen la "dark matter", en el sentido de que existen muy pocos organismos de este filo que han podido ser aislados en cultivo puro (reference). Pero, ¿por qué si son tan abundantes en sistemas de tratamiento de aguas residuales es tan difícil aislarlos?, esta pregunta aún no tiene respuesta. Una posible explicación es que dentro del filo Chloroflexi, todos los miembros de la clase Anaerolineae se caracterizan por su lenta tasa de crecimiento. Los aislamientos obtenidos de muestras provenientes de reactores anaerobios presentan tasas de duplicación que van de las 5 horas a

los 19 días. Considerando además, que tienen un metabolismo fermentador de compuestos orgánicos complejos y simples, existe una gran competencia con fermentadores de rápido crecimiento por estos sustratos (Sekiguchi et al. 2001). Esto dificulta el enriquecimiento de estos organismos y por lo tanto su posterior aislamiento. En sedimentos marinos, se ha propuesto que esta propiedad de crecimiento lento presentada por una especie de la clase Anaerolineae (*Aggregatilinea lenta*) podría reflejar la adaptación de la cepa al bajo flujo de energía/nutrientes en su hábitat natural (Nakahara et al. 2019). Por otro lado, en una comunidad compleja, las tasas de crecimiento pueden ser muy diferentes en comparación con las tasas de crecimiento determinadas para especies aisladas en condiciones de laboratorio, posiblemente debido a las asociaciones cooperativas/sintróficas con otros microorganismos, o las dificultades para identificar el apropiado medio de crecimiento (Campanaro et al. 2020).
Se han diseñado nuevas herramientas para el cálculo de tiempos de duplicación de especies a partir de sus genomas (iREP). Esto resulta interesante ya que la obtención de genomas ensamblados a partir de metagenomas está en aumento exponencial. Trabajos recientes utilizando estas herramientas han sugerido que las tasas de duplicación dependen de las propiedades metabólicas de los genomas ensamblados. Por ejemplo, los valores de iRep calculados para genomas que codifican diferentes vías metabólicas puso de manifiesto que aquellos genomas implicados en la degradación de polisacáridos tienen valores de iRep bastante bajos (crecen más rápido) y genomas implicados en la degradación de celulosa o en la fijación del carbono tienen iREP más altos (crecen más lento) (Campanaro et al.2020). La disponibilidad de valores de iRep para un gran número de especies, y su asociación con los roles funcionales de cada microorganismo puede proporcionar una estimación de la dinámica de crecimiento de las especies involucradas en cada paso por ejemplo de la digestión anaerobia. A partir de estos modelos matemáticos se obtiene información de toda la comunidad, y el análisis no queda limitado sólo a aquellos organismos cultivados.
Otra hipótesis es que estos microorganismos necesitan para crecer en cultivo algún compuesto generado por otro microorganismo o que necesitan de la presencia de otro microorganismo para poder crecer. La sinergia de organismos del filo Chloroflexi con archaeas metanogénicas ha sido propuesta en varios trabajos (referencia).

5.5 ¿Cuál es la limitación de cada técnica y cuál de las técnicas de estudio utilizadas sería la que más aporta?

A lo largo de este trabajo de Tesis se utilizaron diferentes técnicas de biología molecular para estudiar la estructura y función de las comunidades microbianas. Estas técnicas han ido avanzando a lo largo del tiempo. La secuenciación masiva de amplicones (productos de PCR) revolucionó la microbiología generando datos de secuenciación masivamente, pudiendo analizar rápidamente la diversidad, taxonomía, abundancia y estructura de las comunidades ante diversos escenarios y perturbaciones en el ecosistema. Determinando así la presencia de especies no cultivadas hasta entonces, las cuales se estima representan un 99% de las comunidades totales. A pesar de esto, esta técnica tiene limitaciones propias de los errores producidos por la amplificación por PCR (Liu et al. 2020). Sumado a esto, considerando que el gen de ARNr 16S tiene una longitud de 1500 pares de bases y que los primers están dirigidos a determinadas regiones que varían entre 200 y 400pb, en la mayoría de los casos sólo es posible alcanzar la clasificación taxonómica a nivel de género. Además, la diversidad detectada puede cambiar sustancialmente dependiendo de la región del gen que se amplifique, lo cual puede sesgar los resultados, dificultando la comparación entre estudios. Otro aspecto no menor, es que en los casos donde se cuenta con muy poca biomasa inicial para hacer la extracción del ADN genómico, la secuenciación de amplicones puede generar falsos positivos. A pesar de esto, es una de las técnicas más utilizadas para contestar la interrogante de qué microorganismos están presentes en las muestras en determinadas condiciones, y cómo responden a ciertas perturbaciones en el ecosistema.

Con la aparición de la metagenómica, es posible secuenciar todo el ADN genómico de la muestra, por lo tanto se puede responder la pregunta de qué microorganismos están presentes, sus potenciales funciones metabólicas, y cómo varían ante diversos escenarios. Esta técnica no requiere amplificación por PCR, evitando los sesgos o errores asociados a la amplificación por PCR, así como también los sesgos de diversidad relacionados a la selección de una región de un gen en particular. Además de esto, es posible recuperar genomas de organismos no cultivados casi completos mediante el ensamblado de las secuencias, y así poder determinar qué función cumple cierto microorganismo de interés. Por otro lado, gracias al ensamblado de secuencias se puede obtener el gen de ARNr 16S de longitud completa. Por lo tanto, con esta técnica se puede alcanzar una asignación taxonómica a nivel de especie. A pesar de que los costos de la secuenciación de un metagenoma van disminuyendo paulatinamente, aún sigue siendo una técnica costosa en comparación con la secuenciación de amplicones. La

complejidad del análisis de los datos de un metagenoma, hace necesario un mayor conocimiento bioinformático, así como también un mayor requerimiento de características informáticas (Liu et al. 2020).

La utilización de cada técnica está bien definida. La secuenciación del gen del ARNr 16S permite un análisis rápido de la diversidad en las muestras, y el estudio de la dinámica de las poblaciones frente a distintas condiciones, con costos relativamente bajos. La metagenómica abarca estos aspectos, aunque es necesario contar con el conocimiento del proceso de análisis y es más costoso. Una de las ventajas más importantes de la metagenómica es la posibilidad de inferir en las potenciales funciones metabólicas de la comunidad. Por lo tanto, la elección del método dependerá de nuestra pregunta a responder.

5.6 Conclusiones

Mediante diversas técnicas de biología molecular, se determinó que organismos del filo Chloroflexi forman parte del grupo central de microorganismos en todos los reactores estudiados (reactores metanogénicos, reactores de lodos activados y anammox).

En los reactores metanogénicos se pudo comprobar que la gran mayoría pertenece a la clase Anaerolineae, que el inóculo, el tiempo de operación del reactor y el sustrato influyen en la diversidad de estos microorganismos.

Dentro de este grupo, predominó la clase Anaerolineae presentando una gran diversidad de especies en cada uno de los sistemas, que fueron específicas de cada reactor.

A partir del ensamblado de genomas se obtuvieron 17 especies nuevas lo cual permite ampliar el conocimiento filogenético y funcional de este grupo de microorganismos difícil de aislar en cultivo puro. Mediante la anotación funcional de los genomas, se determinó que el filo Chloroflexi cumpliría la función de hidrólisis y fermentación de compuestos complejos y simples. Probablemente teniendo un rol importante en la degradación de restos celulares, favoreciendo la integridad de gránulos y flóculos. Por otro lado, en esta tesis se demostró que distintas especies de la clase Anaerolineae dentro de un mismo reactor y entre reactores, conferían a esto una redundancia funcional de gran importancia para la estabilidad de estos sistemas. A su vez, se detectó en la mayoría de los sistemas estudiados que presentaban morfología filamentosa, y presentaban el potencial de tener propiedades de adhesión. Estas propiedades les confiere un rol preponderante en la formación de la estructura de gránulos y flóculos, y a su vez ayuda a su persistencia en los reactores. Particularmente en reactores anammox, se determinó que el filo Chloroflexi colaboraría con los procesos de

desnitrificación, DNRA y nitrificación, favoreciendo la remoción de nitrógeno en estos sistemas.

De acuerdo con los resultados de esta tesis y la bibliografía, los organismos del filo Chloroflexi tienen un rol sumamente importante en sistemas de tratamiento de aguas residuales y residuos sólidos, en primer lugar influyendo en la eliminación de la materia orgánica, y en segundo lugar, influyendo directamente en la estructura funcional de estos sistemas, como son gránulos y flóculos.

5.7 Perspectivas

A pesar de los avances en el tema logrados en este trabajo de Tesis, quedan aún muchas interrogantes sin responder, a continuación, se listan algunas ideas como para poder seguir avanzando en el entendimiento de este grupo de microorganismos.

- Ampliar los esfuerzos direccionados al aislamiento de organismos del filo Chloroflexi en sistemas de tratamiento, para poder profundizar en el conocimiento sobre su fisiología y posible rol en estos sistemas

- Abordar el estudio del filo Chloroflexi en sistemas de tratamiento a partir de la metatranscriptómica para determinar qué genes expresan estos organismos *in situ.*

- Desarrollar experimentos de reactores escala laboratorio aplicando diferentes condiciones operacionales, por ejemplo, variación en la carga orgánica, temperatura, pH, concentración de nitrógeno, para determinar los posibles causantes del sobrecrecimiento de organismos del filo Chloroflexi y generar planes de control para evitar episodios de bulking.

- Generar redes de co-ocurrencia para estudiar la interacción entre distintos microorganismos y miembros del filo Chloroflexi. Este tipo de análisis calcula la posible relación entre diferentes microorganismos o taxas de acuerdo a su abundancia a través de las muestras.

- A partir de los estudios funcionales de organismos Chloroflexi involucrando la operación de reactores escala laboratorio, la metagenómica y metatranscriptómica,

generar manuales con estrategias para el control del sobrecrecimiento de bacterias filamentosas del filo Chloroflexi

5.8 Referencias

Ahn J, McIlroy S, Schroeder S, Seviour R (2009) Biomass granulation in an aerobic:anaerobic-enhanced biological phosphorus removal process in a sequencing batch reactor with varying pH. J Ind Microbiol Biotechnol 36:885–893 . doi: 10.1007/s10295-009-0566-3

Hulshoff Pol LW, De Castro Lopes SI, Lettinga G, Lens PNL (2004) Anaerobic sludge granulation. Water Res 38:1376–1389 . doi: 10.1016/j.watres.2003.12.002

Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X, Bai Y (2020) A practical guide to amplicon and metagenomic analysis of microbiome data. Protein Cell. doi: 10.1007/s13238-020-00724-8

Nielsen PH, Kragelund C, Seviour RJ, Nielsen JL (2009) Identity and ecophysiology of filamentous bacteria in activated sludge. FEMS Microbiol. Rev. 33:969–998

Nierychlo M, Miłobędzka A, Petriglieri F, McIlroy B, Nielsen PH, McIlroy SJ (2019) The morphology and metabolic potential of the Chloroflexi in full-scale activated sludge wastewater treatment plants. FEMS Microbiol Ecol 95: . doi: 10.1093/femsec/fiy228

Sekiguchi Y, Kamagata Y, Nakamura K, Ohashi A, Harada H (1999) Fluorescence *in situ* hybridization using 16S rRNA-targeted oligonucleotides reveals localization of methanogens and selected uncultured bacteria in mesophilic and thermophilic sludge granules. Appl Environ Microbiol 65:1280–1288

Sekiguchi Y, Takahashi H, Kamagata Y, Ohashi A, Harada H (2001) *In situ* Detection, Isolation, and Physiological Properties of a Thin Filamentous Microorganism Abundant in Methanogenic Granular Sludges: A Novel Isolate Affiliated with a Clone Cluster, the Green Non-Sulfur Bacteria, Subdivision I. Appl Environ Microbiol 67:5740–5749 . doi: 10.1128/AEM.67.12.5740-5749.2001

Wiegant WM, de Man AWA (1986) Granulation of biomass in thermophilic upflow anaerobic sludge blanket reactors treating acidified wastewaters. Biotechnol Bioeng 28:718–727 . doi: 10.1002/bit.260280511

Wittebolle L, Marzorati M, Clement L, Balloi A, Daffonchio D, Heylen K, De Vos P, Verstraete W, Boon N (2009) Initial community evenness favours functionality under selective stress. Nature 458:623–626 . doi: 10.1038/nature07840

Yamada T, Sekiguchi Y (2009) Cultivation of uncultured Chloroflexi subphyla: significance and ecophysiology of formerly uncultured Chloroflexi “subphylum i” with natural and biotechnological relevance. Microbes Environ 24:205–216. doi: 10.1264/jsme2.ME09151S