



UNIVERSITAT DE
BARCELONA

Diversitat bacteriana a les etapes de potabilització de les aigües de consum


Anna Pinar Méndez



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement- NoComercial – SenseObraDerivada 4.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento - NoComercial – SinObraDerivada 4.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution-NonCommercial-NoDerivs 4.0. Spain License.**



Diversitat bacteriana a les etapes de potabilització de les aigües de consum

Anna Pinar Méndez

Tesi Doctoral, Maig 2022



UNIVERSITAT DE
BARCELONA

Programa de Doctorat de Biotecnologia

Departament de Genètica, Microbiologia i Estadística

Secció de Microbiologia, Virologia i Biotecnologia

Facultat de Biologia

Diversitat bacteriana a les etapes de potabilització de les aigües de consum

Memòria presentada per l'**Anna Pinar Méndez** per a optar al grau de
Doctora per la Universitat de Barcelona.

Tesi realitzada sota la direcció de la Dra. Cristina García Aljaro i el Prof. Dr. Anicet Blanch i Gisbert al Departament de Genètica, Microbiologia i Estadística de la Universitat de Barcelona i al Departament de Microbiologia d'Aigües de Barcelona.

La Directora i tutora ,

Dra. Cristina García Aljaro

El Director,

Prof. Dr. Anicet R. Blanch i Gisbert

La Doctoranda,

Anna Pinar Méndez

Barcelona, maig de 2022

*“The difficult is
what takes a little time.
The impossible is
what takes a little longer.”*

~ Fridtjof Nansen (Norwegian polar explorer). Listener 14 December 1939

A Victor

AGRAÏMENTS

Vaig començar la tesi al 2017 amb una mica d'experiència com a analista en microbiologia d'aigües però completament inexperta en biologia molecular, bioinformàtica, estadística, o taxonomia bacteriana, així que assumia el repte molt motivada per submergir-me en aquestes àrees i començar una nova etapa. Ara, cinc anys després, entrego la meva tesi. Malgrat la muntanya russa que suposa l'etapa predoctoral, aquests últims anys m'han obert els ulls a nous aspectes de la meva vida com a científica i també com a persona, i he viscut experiències inoblidables que confio m'aportaran les eines necessàries per continuar creixent en l'àmbit personal i professional.

La tesi és una muntanya amb alt desnivell, i per ascendir, necessites eines i coneixement. No és una muntanya normal, es variable i el pic s'allunya i s'apropa contínuament, però no puges sola. I és que al llarg del camí t'acompanyen moltes persones de les qui aprens perquè puguis arribar a fer el cim i per això tu estàs aquí, i gràcies a tu, jo també. El cert és que ja estava avisada de la "muntanya tesi" abans de començar (*mea culpa*) però el que no m'havien dit és que havia de portar el *anillo al Monte del Destino* a contrarellotge amb sabates de taló...així que com bé aconsellava el gran Dr. Xavi Fernández Casi, si estàs dubtant en fer una tesi... HUYE! ;)

Per començar volia agrair als meus directors de tesi i tutora, els responsables que avui estigui aquí escrivint aquestes línies, m'heu acompanyat fins al cim i ara ja em queden pocs metres per a la defensa :)

Belén, aquesta era la idea que et rondava pel cap quan vas picar la porta de l'Anicet i la Cristina, i al final s'ha fet realitat. Gràcies per confiar en mi, escoltar-me i el teu suport. Tot això ja va començar abans del doctorat quan em vas donar l'oportunitat d'entrar al gran equip de micro, i fer el TFM i la resta de projectes que han anat sorgint, i durant aquesta etapa m'has ensenyat a volar i alçar el vol. Gràcies pels teus *Belén al rescate*, a ensenyar-me a estimar els colorcitos, a la teva capacitat crítica i rigor científic i als nostres "5 minuts" i prou. Haurem de tancar aquesta etapa com quan vam entregar el TFM!

Anicet, he de dir que em feia vertigen començar la tesi perquè no sabia si podria estar a l'alçada del grup ni quin ambient em trobaria, però ràpidament m'heu fet sentir com a casa perquè MARS és un gran equip. Escoltar-te sempre és una classe magistral! I n'he pogut aprendre molt durant aquesta etapa. Gràcies per tenir la porta sempre oberta, per aconseguir que pogués anar a fer l'estada amb l'Owen a Tromsø, per la teva confiança i recolzament durant aquesta etapa i per ensenyar-me a ser pragmàtica quan el temps apreta.

Cristina, mil gràcies per ensenyar-me a fer pràcticament tot el que no sabia abans de començar la tesi, com tota la part molecular, escriure articles, a barallar-nos amb R i amb les seqüències... També per la teva paciència infinita, per escoltar-me, treballar tantíssim i tant bé colze a colze fins a hores intempestives, per les autoritzacions i estar pendent quan portava les mostres a la maleta cap Tromsø per si em detenien a l'aeroport! Quina d'experiències bones i altres més complicades hem passat. Sento que hem crescut i evolucionat molt juntes aquests 5 anys, agafant les dues més confiança en les nostres tasques com a directora i com a doctoranda. Has estat la meva mentora per poder assolir poc a poc les fites de la tesi, així que gràcies per sempre ajudar-me en tot! Moltes gràcies també a la resta de caps del grup. No he pogut treballar directament amb vosaltres, però m'ha encantat compartir estones divertides. Als veterans, per ser tant autèntics i explicar-nos "batallitas" que ens alegren el dia, gràcies **Francisco** i **Joan Maite**, moltes gràcies pels teus consells al llarg de la tesi i pel suport a Viena! Em fa molta il·lusió que puguis formar part del tribunal de la tesi per tancar aquesta etapa.

També moltes gràcies a tots els **MARSIT@S**, l'engranatge perquè el grup funcioni tant bé!! Com deia abans, hi ha tantes històries rares fent tesis que em feia por caure en un grup i no encaixar. Quina sort he tingut de caure en aquest grup, sou una petita família fantàstica. Gràcies per escoltar els meus monòlegs i fer *corrillos* per parlar de mil coses. M'he sentit molt a gust amb vosaltres i el que em sap greu és no haver pogut passar més temps amb vosaltres, però espero recuperar aquest temps ara que ja seré més lliure. **Aiora** tens una capacitat d'organització i coordinació admirables. Gràcies per ajudar-me amb tot, i no només el material, pressupostos i coses rares que ens hem trobat amb la meva partida pressupostària...També per escoltar-me i aconsellar-me!. **Raquel!** Pura energia, buen rollo e increíble tu capacidad de trabajo! Fuiste mi guía para dar los primeros pasos en el grupo. Gracias por tus consejos de la vida y cosmética *zero waste y cruelty-free*, tus abrazitos de ánimo y sobretodo por tus anécdotas tan divertidas de tu día a día, nadie podría explicarlo con más gracia! **Laura** tu vas ser la meva doctora de referència just a l'entrar tot i que ja ens coneixíem del treball amb el RETOS. No sé si ho recordes però em vas donar mil consells com en 5 minuts que encara segueixo processant! Quina crack i quina energia! Gràcies per ajudar-me en els meus inicis! **Miriam**...MIL GRÀCIES per tots els consells, ajuda, suport, i aguantar els meu *podcast* de socorro per *whats* (aprofito per demanar disculpes a tots els que llegiu perquè segur que heu rebut algun dels meus *podcast* per *WhatsApp* de més de 10 minuts...) i "assistència tècnica" al llarg de tota la tesi. Ets una gran persona i gran professional. **Julia**, gracias por tus consejos técnicos y personales al lo largo de la tesis, y por compartir juntas todos los virus que han circulado estos años! Creo que no nos hemos saltado ni uno jeje. **Dani**, compañero de fatigas desde que empezamos juntos el máster! Muchas

gracias por todo! Eres un crack con un sentido del humor brutal! Gracias por todos tus consejos, por existir cuando estabamos en el curso de Python xD, ayudarme a encontrar mi casa de Hogwarts y por mantenerme tranquila en momentos de caos. **Paula**, hem coincidit poquet però compartim passió por lo polar! Gràcies per avisar-nos dels les sessions *café con hielo* tant interessants. Molts ànims en aquesta recta final! **Pedro** a ti también te queda bien poquito, creo que vamos a depositar a la par, así que te deseo lo mejor, que esto ya casi lo tenemos! Gracias por los buenos ratos en Viena. **Eli** moltes gràcies per la teva ajuda amb les qPCR quan em sortien coses rares i per les anècdotes de les teves experiències a l'Antàrtida. **Pablo y Lorena**, con vosotros igual hemos coincidido más en los cafés que en el lab por el tipo de trabajo que he hecho, pero me encanta vuestro sentido del humor y vuestros piques asturianos! Así que gracias por vuestros consejos y Lorena, gracias por el apoyo en Viena cuando estaba nervi por la presentación. Y no podía olvidarme de **Javi Méndez**! Muchísimas gracias por tu ayuda con estadística y con R!. Por preguntarme como llevo la tesis y por las charlas de micro tan interesantes donde salen siempre ideas guays! Y al resto de Marsitos con los que no he podido coincidir mucho o nada, estaréis muy bien en este grupo, está lleno de buena gente y con mucha experiencia que os ayudarán muchísimo, así que mucha suerte en vuestras tesis, TFM y TFG!

Una altra persona molt especial a qui vull agrair l'ajuda (i l'amistat) amb les meves primeres passes de molecular a la uni és la **Maite Falcó**. Ets puro nervio, una curranta brutal, amb molt criteri científic, i molt resolutive, sempre disposada a ajudar i infinitament bona explicant conceptes. Gràcies! També volia agrair a la resta de persones estupendes d'altres grups o que han passat per la uni i ja han marxat però amb qui he coincidit al llarg de la tesi: **Xavi Fernández, Robert, Sandra Martinez, Judith, Laia, Maryuri**,... I a totes aquelles persones que quan vinc a la uni i saludo *random* em saluden amb el cap o amb un somriure.

A tot el grup de virus entèrics, gràcies pel vostre suport i ajuda durant els pitjors moments de la pandèmia. Per la barbaritat de coses que he après de vosaltres a una velocitat canviant vertiginosa, pels bons moments i d'altres complicats que hem viscut, gràcies **Adán, Gemma, Edu, Albert i Rosa**.

I per suposat a tot l'equip de secretaria per ajudar sempre amb tots els tràmits administratius, gràcies Mónica, Bea, Manolo...!

Owen Wangensteen y Sandra Garcés, vosotros no podeis faltar aquí. Owen, muchísimas gracias por darme la oportunidad de hacer la estancia en Noruega, acogerme en la UiT y enseñarme todo lo que sé de metabarcoding :) Gracias por tu paciencia y dedicación, ha sido maravillosa la experiencia en el lab y en Tromsø. Así que aprovecho para dar las gracias a todo el grupo: Thanks

Kim Præbel for letting me join your group for a bit more than two months, I really appreciate the opportunity. It was a great experience on a professional and a personal level. It was also amazing to hear from you and your team your stories about the Arctic expeditions up towards the North Pole! Thanks **Enrique, Julie, Michael, Shripathi, Kristel, Kamilla, Nathan, Paulina**,...it was a pleasure to meet you all! I per suposat també **Sandra**, gràcies per les converses científiques, consells de tesi i per les excursions i sopars a Tromsø i a Barna. Y **Sara** gracias por nuestras maratones con Owen en la uni, también las excursiones de madrugada aprovechando el *midnight sun*, y tus consejos de metabarcoding con la tesis. Mucho ánimo con la recta final que ya te queda bien poco para depositar!

Also I would like to thank to **Sophie Cortois, Jean François Loret, Claire Bertelli, Gilbert Greub** and **Gertjan Medema**, for your valuable advices at the very beginning of the thesis during RESIST'EAU'ME project.

La major part de la tesi l'he passat al **laboratori d'aigües de Barcelona** i per suposat no podien faltar en aquestes línies. Per començar, als caps del laboratori. **Miquel** moltes gràcies pel teu suport, confiança i esforços perquè pogués formar part avui de la família lab. AB, i també per preocupar-te sempre com portava la tesi! :) **Jordi Martín**, ets tota una referència! ets un llibre obert i sempre ens deixes bocabadats amb el teu coneixement. Gràcies per ser tant bona persona i estar sempre disposat a ajudar-nos. Gràcies pels teus consells tècnics i personals, pels grans descobriments musicals que ens has anat fent durant la pandèmia, per cuidar-nos durant els moments de confinament total i per les anècdotes. Et desitjo moltíssima sort en la teva nova etapa que segur que gaudiràs molt! Et trobarem a faltar. **Lluís**, gràcies per preocupar-te perquè pogués acabar la tesi i animar-me per fer-ho. Ha estat un cop fort, però et recordarem amb un somriure...

I per suposat a les meves companyes i companys del lab, la segona família! En especial la **família micro**, que és meravellosa, som una autèntica pinya (una mica *gruyère* de vegades jiji), però l'admiració, el suport i estima que ens tenim ens fa imparables! Sempre unides i colze a colze. Gràcies pel vostre suport i per donar-me un cop de mà amb les mostres de la tesi quan he anat desbordada, per aguantar-me les hores i hores de "batallites" amb la tesi i altres mil coses, per la teràpia de "corrillo" on arreglem el món i repassem l'actualitat, pels moments de festa i : **Ivan** per ser tant divertit, enèrgic, bona gent i preparar les millors decoracions de nadal i calendaris d'advent! Que visqui el nadal! **Eva** per ser tant maca sempre i sobretot per ensenyar-me amb molta paciència quan vaig entrar, **Marta** per tenir aquest sentit de l'humor i preocupar-te per mi, **Carles** per transmetre calma quan les coses es compliquen i pels teus consells tècnics. I com

no, al equipo *Whopper con queso* i *SARS Team*, **Gemma i Ane**...noies no sé descriure el que sento per vosaltres! Hem passat moments durs i altres moments boníssims, però sempre juntes i això no esperava trobar-ho a la feina sincerament, i aquí esteu. MILESKER pel suport sempre, consells i tot, sempre aprenc molt de vosaltres. Sou molt cracks! Ara que tindrè més temps lliure, se'ns han acumulat plans a tope... Y en este super grupo de micro no podia faltar **Conxiti**, encara que estiguis a residuals, i **MJo** aunque físicament estés en Alicante, siempre seréis de micro! Conxiti ets una de les persones més bones que he tingut el plaer de conèixer, i ets molt forta, molt! Per cuidar-nos amb les teves receptes de *pancakes* i detallets varis, i per ser tant maca, gràcies i força! Y MJo, te echo muchísimo de menos y ha sido genial todo este tiempo poder trabajar contigo y aprender juntas, nuestros bailoteos, nuestros cálculos de conversiones y esquemas de procesos varios en papel que lo entendemos mejor, nuestras búsquedas del famoso maletín de bioterrorismo...nos veremos pronto! un besazo para ti y Ricardo, Martina y Laia!

I al lab està ple de bona gent, que sempre s'ha preocupat per veure com anaven els avenços de la tesi com la **Núria**, la meva mama del lab i companya *runner* (m'has de posar les piles!), gràcies per tot! Ets tant enèrgica i positiva que es contagia i per això et fas estimar molt. **Verònica Gràcia** gràcies per ser tant maca, preocupar-te de tothom i ajudar-me amb temes administratius moltes vegades. **Meritxell**, moltes gràcies pel teu suport tècnic i per sempre preguntar-me com porto la tesi, ets un sol! **Carmen** lo que nos reimos y hablamos no está escrito! Eres una bellíssima persona y explicas maravillosamente bien! Gracias por tu ayuda técnica y personal siempre que te he necesitado. **Rosa**, moltes gràcies pels cosells amb la tesi, ja estic gairebé al final! **Xavi Pardo**, gràcies per tota l'ajuda durant aquests anys, des de quan estava a presa de mostres i em perdia literalment per la muntanya buscant els dipòsits, a ajudar-me amb els mostreigs, i per presentar-me a la Belén per poder fer el TFM! Què et sembla? Això en part és culpa teva ;P. **Raúl**, gracias por ser tan divertido y buena gente, darme consejos de *runner*, nutrición, escucharme siempre, y venir ayudarme cuando se me han hecho las tantas trabajando en el lab. **Jorge**, gracias por ser tan auténtico y decir las cosas tal cual, por escucharme y animarme en toda esta etapa. **Aleix i Dani**, gràcies a vosaltres també per ajudar-me a descarregar extractors o fer fotos de plaques, o guardar-les quan algun dia les havia deixat fora jeje. Al compi **Pérez**! Qué etapa más guay la de recogida de muestras! He aprendido muchísimo de un maestro, gracias por tu paciencia! A **Alejandra** por estar siempre tan dispuesta a ayudar y ser tan maja! A **Raquel Cardeñoso** por saludarme siempre con un "pero bueno, ¿que ven mis ojos?" y preocuparte por como llevo mi tesis. A **Haki** y **Juani** por cuidarnos siempre tanto y ser tan buenas con nosotros. Gracias a vosotras y a todas las compañeras que nos ayudais tantísimo en el lab con todo el material. I no em vull deixar a ningú, així que gràcies a tot el laboratori de Collblanc!

I també als company@s de l'ETAP, que tant m'han ajudat en aquesta etapa com la **Marga, Pau, David, Sergio i Marta** (tot i que esteu a Collblanc), GRÀCIES! També al **Daniel Garrote** per la fotografia de la posta del sol a l'ETAP durant el confinament.

També, com no, als impulsors d'Innovació dels doctorands industrials a AB i al grup, **Ana Reyes** i **Ramón López**, gràcies! I a la **Núria Roigé** amb qui vaig començar com a doctoranda a la casa.

Als **Cetaquos** (i ex-Cetaquos) pels bons moments dins i fora del lab, en especial al equipo *Bichotes* **David** i **Sonia**, gràcies per totes les converses megaprofundes científiques i siderals, i molta sort en els vostres projectes! Ens retrobem aviat. Així com també a la **Clàudia, Miquel Jorba, Ruben, Kata, Susana, Albert,...**

I com no podia ser d'una altra manera en aquesta tesi, estic MOLT agraïda a la família *workshopera* MRAMA, en especial al **Josep Yuste** i la **Marta Capellas**, perquè per la vostra passió, dedicació, i perquè no podeu ser més professionals i bona gent a la vegada, em vàreu obrir els ulls a la microbiologia, just acabar la carrera en el moment crític quan et plantejes cap a on seguir, i tota la resta ha anat prenent sentit sobre la marxa. Sou un referent per a mi i sempre us estaré molt agraïda de tot el que he après i m'ha servit per endavant. Aprofito també per agrair als dos grans persones i habituals del workshop, el **Jon** i el **Pascal**, per tot el que he après de vosaltres. I no puc parlar del workshop sense recordar al **Dr. Fung**, a qui vaig tenir la sort de conèixer any rere any al workshop, i que va ser una persona impressionant!

Per suposat, he de donar les gràcies als meus amics, un suport indispensable SEMPRE. **Mireia**, amiga de fatigues bio, de "merendoles", de converses davant del cotxe de "ya me voy" i que se'ns donin sempre les tantes de la matinada, workshopera habitual i de suport sempre en totes les etapes. GRÀCIES pel suport i per ser com ets! Gràcies també als meus amics i amigues **Santi, Albeta, Agnès, Lili i Pini** perquè vam formar una petita família a la uni a Girona i passats tots aquests anys, és preciós que seguim units. Gràcies pel vostre suport, consells i ànims per la tesi, i entendre les meves absències dels plans dels últims mesos, per estar sempre que us necessito i per estirar per veure'ns sempre que podem. Us he trobat molt a faltar però aviat ens veiem! I no podíeu faltar vosaltres, gràcies **Nina, Òscar, Rosa i Duane** per aguantar les "chapas" escoltant-me parlar del que anava fent a la tesi, els articles, i que sempre tingueu tant interès en preguntar i preocupar-vos per com va la tesi i com està l'aigua i les caques de la gent, així com per ser els millors i més guapos Ironman supporters del món mundial! **Kelsey**, muchas gracias por estar siempre, porque cuando vuelves a Barcelona siempre buscas un hueco para vernos y es como si no pasara el tiempo.

Para terminar, he dejado la parte invisible que también “sufre” la tesis y sin embargo es esencial para llegar hasta aquí: mi familia. Gracias a mi **yayita Paquita**, por ser tan buena, escucharme siempre con admiración, por cantar tan bien y por quererme tanto. Sé que estarías muy contenta de ver que he podido terminar la tesis... un abrazo muy fuerte! Gracias a mis abuelos **Alicia y Manolo**, por todo el cariño y apoyo, y por entender que he estado ausente demasiado estos últimos meses. También a todos mis tíos y primos que he echado mucho de menos estos largos meses. Gracias a mis suegros **Reme y Pedro** por los ánimos y apoyo, y a mi cuñada **Lau** y mi cuñado **Marc**. Pero sobre todo gracias a mi hermana **Lidia** y a mis padres **Tere y Pedro** por soportar “mis cosas de la tesis” que han hecho que muchas veces pasen varias semanas sin vernos por ir demasiado liada. Gracias por ayudarme a llegar hasta aquí, ayudarme a ser quien soy y por animarme y apoyame en todo el proceso, os quiero!

Y para terminar, esta vez si Victor, no voy a olvidarme de ponerte en los agradecimientos. Ahora si, éstas son las últimas palabras que escribo de la tesis, nuestra tesis, y ahora estás dormido (como es lógico porque estoy escribiendo de madrugada) pero llevas desde que nos conocimos apoyándome y animándome en todo siempre, y este reto ha sido largo, muy largo, pero tú también has subido la montaña conmigo y, sin dudarlo, me has acompañado hasta la meta. Formamos un gran equipo, y esta tesis es tan tuya como mía. Gracias por ser tan extremadamente buena persona, confiar siempre en mi, por aguantar esta montaña rusa y por crecer cada día juntos. Por todos los viajes, los conciertos, la furgo, por la locura de los últimos meses con hipotecas, traslados...nos merecemos unas buenas vacaciones y por fin, estan más cerca que nunca, Aotearoa! Te quiero Victor, gracias por existir :)

ÍNDIX

AGRAÏMENTS	9
ÍNDIX.....	16
LLISTAT D'ABREVIATURES	18
ABSTRACT	22
RESUM.....	23
INTRODUCCIÓ	25
1. Aigua, accés a l'aigua potable i microbiologia	27
2. Estrès hídric i escalfament global a la regió mediterrània	28
3. One Water : canvis en les percepcions de la qualitat, el valor i la utilitat de les aigües.....	29
4. Tractament de l'aigua potable a Barcelona	30
4.1. L'abastament de l'aigua a Barcelona: Antecedents de la captació	30
4.2. Particularitats del riu Llobregat al punt de captació de l'ETAP de Sant Joan Despí	30
4.3. Potabilització i distribució. La ETAP de Sant Joan Despí en l'actualitat	34
5. Gestió integral de l'aigua	38
5.1. L'ús de recursos hídrics alternatius a Barcelona.....	38
5.2. Gestió i demanda hídrica a Barcelona.....	39
6. Context legislatiu de la Qualitat de l'aigua de consum. Directiva EU 2020/2184	40
6.1. Control microbiològic de la qualitat de l'aigua: indicadors de la contaminació fecal	41
7. Comunitats bacterianes a l'aigua.....	43
7.1. Mètodes de concentració de microorganismes.....	49
7.2. Mètodes de detecció i identificació de microorganismes.....	52
7.2.1. Espectrometria de masses MALDI-TOF MS per a la identificació de bacteris	54
7.2.2. Caracterització de comunitats bacterianes totals per seqüenciació massiva del gen 16S ARNr (metabarcoding).....	60
7.3. El paper dels microorganismes als sistemes d'aigua potable	64
OBJECTIUS.....	67
INFORMES DE LES PUBLICACIONS	71

1. Llistat de publicacions	73
2. Informe sobre el Factor d'impacte	74
3. Informe de participació en les publicacions	75
PUBLICACIONS	77
1. CAPITOL 1.....	78
2. CAPITOL 2.....	108
3. CAPITOL 3.....	153
DISCUSSIÓ GENERAL.....	201
CONCLUSIONS.....	215
BIBLIOGRAFIA	221
ANNEXOS.....	237
1. Altres publicacions	239

LLISTAT D'ABREVIATURES

°C	Unitats de temperatura, graus centígrads
× <i>g</i>	Unitats de força centrífuga
μL	Unitats de volum, microlitre
μs/cm	Unitats de conductivitat, microSiemens per centímetre
ADN (ó DNA)	Àcid desoxiribonucleic (de l'anglès, <i>Deoxyribonucleic acid</i>)
ARN (ó RNA)	Àcid ribonucleic (de l'anglès, <i>Ribonucleic acid</i>)
ARNr 16S (ó 16S rRNA)	Gen que codifica l'ARN ribosomal procariota 16S
ASV	Variant de seqüència d'amplicons (de l'anglès, <i>Amplicon Sequence Variants</i>)
BDAL	Bruker Daltonics
bp	Parell de bases (de l'anglès, <i>base pair</i>)
CECT	Col·lecció Espanyola de Cultius Tipus
CF	Filtració per carbó (de l'anglès, <i>Carbon Filtration</i>)
CFU	Unitats formadores de colònies (de l'anglès, <i>Colony-Forming Unit</i>)
CP	<i>Clostridium perfringens</i>
GC-FAME	Cromatografia de gasos d'àcids grassos metil ester (de l'anglès, <i>Gas Chromatographic Fatty Acid Methyl Esters</i>)
Da	Unitats de massa atòmica, Daltons
DEC	Decantació (de l'anglès, <i>Decantation</i>)
DEUF	Mètode d'ultrafiltració sense sortida (de l'anglès, <i>Dead-end Ultrafiltration</i>)
DW	Aigua potable (de l'anglès, <i>Drinking Water</i>)
DWL	Llibreria d'aigua potable (de l'anglès, <i>Drinking Water Library</i>)
<i>E. coli</i> (ó EC)	<i>Escherichia coli</i>

EE	Error esperat (de l'anglès, <i>Expected Error</i>)
ETAP (o DWTP)	Estació de Tractament d'Aigua Potable (en l'anglès, <i>Drinking Water Treatment Plant</i>)
et al.	I altres (autors) (del llatí, <i>et alia</i>)
GW	Aigua subterrània de pous (de l'anglès, <i>Groundwater</i>)
h	Unitats de temps, hores
HCCA	Àcid alfa-ciano-4-hidroxicinamic
hm ³	Unitats de volum, hectòmetre cúbic
HPC	Enumeració de bacteris aerobis heterotròfics en placa (de l'anglès, <i>Heterotrophic Plate Counts</i>)
HT	Temperatura alta (de l'anglès, <i>High Temperature</i>)
HTS	Seqüenciació d'alt rendiment (de l'anglès, <i>High Troughput Sequencing</i>)
IE	Enterococs intestinals (de l'anglès, <i>Intestinal Enterococci</i>)
ISO	Organització internacional per a l'estandardització (de l'anglès, <i>International Organization for Standardization</i>)
Kb	Unitats de longitud, Quilo-bases
km ³	Unitats de volum, quilòmetre cúbic
KDa	Unitats de massa, Quilo-daltons
L	Unitats de volum, litre
log/s	Logaritme/s
LT	Temperatura baixa (de l'anglès, <i>Low Temperature</i>)
M	Unitats de massa, Molar
m ³	Unitats de volum, metre cúbic
m ³ /s	Unitats de flux, metre cúbic per segon
MALDI-TOF MS	Espectrometria de masses (MS) de desorció/ionització mitjançant làser assistit per matriu (MALDI) i detector acoblat del temps de vol dels ions

	(TOF) (de l'anglès, <i>Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry</i>)
mg	Unitats de massa, mil·ligram
min	Unitats de temps, minuts
MIX	Cambrà de mescla (de l'anglès, <i>Mixing chamber</i>)
mL	Unitats de volum, mil·lilitre
mm	Unitats de longitud, mil·límetre
MSP	Perfil d'espectre principal (de l'anglès, <i>Main Spectrum Profile</i>)
m/z	Relació massa/càrrega
NA	No analitzat
ND	No detectat
NGS	Seqüenciació d'última generació (de l'anglès, <i>Next-Generation Sequencing</i>)
NMDS	Anàlisi d'escalament multidimensional no mètric (de l'anglès, <i>Non-Metric Multidimensional Scaling</i>)
NMP	Nombre més probable
N	Nucleòtid aleatori (N)
NTU	Unitats de terbolesa (de l'anglès, <i>Nephelometric Turbidity Unit</i>)
OMS (ó WHO)	Organització Mundial de la Salut
ONU	Organització de les Nacions Unides
OZ	Ozonització (de l'anglès, <i>Ozonization</i>)
p	Valor de probabilitat de significació estadística, p-valor
PCR	Reacció en Cadena de la Polimerasa (de l'anglès, <i>Polymerase Chain Reaction</i>)
PERMANOVA	Anàlisi de la variància multivariant permutacional (de l'anglès, <i>Permutational Multivariate Analysis of Variance</i>)

pH	Potencial d'hidrogen
pM	Unitats de massa, picomolar
ppm	Parts per milió
RO	Osmosi inversa (de l'anglès, <i>Reverse Osmosis</i>)
qPCR	PCR quantitativa (de l'anglès, <i>quantitative Polymerase Chain Reaction</i>)
RW	Aigua superficial de riu (de l'anglès, <i>River water</i>)
s	Unitats de temps, segon
SF	Filtració per sorra (de l'anglès, <i>Sand Filtration</i>)
SIMPER	Percentatge de similaritat (de l'anglès, <i>Similarity percentatges</i>)
SOMCPH	Bacteriòfags somàtics d' <i>E.coli</i> (de l'anglès, <i>Somatic Coliphages</i>)
sp.	Espècie
spp.	Diverses espècies
TC	Coliforms totals (de l'anglès, <i>Total Coliforms</i>)
TOC	Carboni orgànic total (de l'anglès, <i>Total Organic Carbon</i>)
UPGMA	Agrupament jeràrquic per parells no ponderat amb mitjana aritmètica (de l'anglès <i>Unweighted Pair Group Method with Arithmetic mean</i>)
WPCA (ISO)	Medi agar ISO per l'enumeració de bacteris heterotròfics de l'aigua (de l'anglès, <i>Water Plate Count Agar (ISO)</i>)

ABSTRACT

In recent years, human population growth and the effects of global warming are reducing the capacity and quality of the available freshwater resources. The challenges we face call for strategies that allow using alternative water resources together with suitable drinking water treatment technologies and water quality monitoring tools that guarantee water safety.

Drinking water treatment technologies can currently meet this challenge by providing an effective barrier to contaminants and pathogenic microorganisms. Although monitoring microbial indicators of fecal pollution according to the drinking water Directive (EU 2020/2184) guarantee water safety, indicator microorganisms represent a minor fraction of the water microbiome. Therefore, complementary methods are needed to detect variations of the whole bacterial communities that could influence the quality or safety of water.

The aim of this thesis work is the characterization of bacterial communities in water treatment stages of the drinking water treatment plant (DWTP) of Sant Joan Despí (Barcelona) by means of two approaches, one culture-dependent (proteomics) and the other culture-independent (massive sequencing or metabarcoding). The research allowed us to develop a proper method to identify culturable bacteria from water by MALDI-TOF mass spectrometry, to track the elimination of microbial fecal indicators throughout the treatments, and to assess diversity and spatial and temporal dynamics of both culturable and whole bacterial communities by MALDI-TOF MS and metabarcoding, respectively.

High bacterial diversity was observed in the DWTP and it varied throughout the treatments, especially in drinking water. The bacterial diversity was influenced by the treatment applied as well as the ambient temperature. It was also observed an effective reduction of microbial fecal indicators, as well the detection of sequences of pathogenic and potentially biofilm-forming bacteria at low proportions but undetectable by conventional regulated methods. The use of large water volumes concentrated by ultrafiltration improved sample representativity, especially in advanced treatment stages when indicator concentration is very low.

The strategy to improve the understanding of the DWTP microbiome is to combine different methodological approaches to assess water quality with a holistic vision for water resources management and future adaptation to the impacts of global warming.

RESUM

En els darrers anys l'augment poblacional així com els efectes d'un escalfament global estan minvant la capacitat i la qualitat dels recursos hídrics naturals disponibles. Els reptes als que ens enfrontem donen pas al desenvolupament d'estratègies que ens permetin l'ús de recursos hídrics alternatius que vagin de la mà de tecnologies de tractament d'aigua potable, així com a sistemes de control de la qualitat de l'aigua que garanteixin la seva seguretat a tota la població.

Les tecnologies de tractament d'aigua potable actualment poden assumir aquest repte constituint una barrera eficaç a contaminants i microorganismes patògens. Si bé els paràmetres microbiològics de la qualitat de l'aigua legislats segons la Directiva EU2020/2184 asseguren la seva innocuïtat alertant d'una possible contaminació fecal en els sistemes d'aigua potable, es necessiten mètodes complementaris que permetin detectar en detall les variacions de les comunitats bacterianes heteròtrofes que podrien influir en la qualitat i/o seguretat de l'aigua.

L'objectiu principal de la present tesi doctoral ha estat la caracterització de les comunitats bacterianes a les etapes de potabilització de l'estació de tractament d'aigua potable (ETAP) de Sant Joan Despí (Barcelona) mitjançant dues aproximacions, una basada en cultiu (proteòmica) i l'altra independent de cultiu (seqüenciació massiva o *metabarcoding*). Els estudis realitzats han permès posar a punt una metodologia per a identificar bacteris cultivables de l'aigua per espectrometria de masses MALDI-TOF, determinar l'eliminació dels indicadors de contaminació fecal al llarg dels tractaments, i avaluar la diversitat i les dinàmiques espacial i temporal de les comunitats bacterianes tant cultivables com totals per MALDI-TOF MS i metabarcoding, respectivament.

S'ha observat una gran diversitat bacteriana a l'ETAP que ha anat variant al llarg dels tractaments, especialment a l'aigua potable, i influenciada per alguns factors com el tipus de tractament aplicat i l'estacionalitat establerta per la temperatura, així com una reducció eficaç dels indicadors i la detecció de seqüències de bacteris patògens i potencials formadors de biofilms a baixes proporcions però no detectables pels mètodes legislats. La representativitat de les mostres corresponents a etapes on la concentració dels indicadors microbians era baixa, s'ha cobert amb mostres d'un gran volum d'aigua concentrades per ultrafiltració.

L'estratègia per a millorar el coneixement del microbioma dels processos de tractament de l'aigua potable consisteix a combinar diverses eines metodològiques per avaluar la qualitat de l'aigua i així disposar d'una visió holística per a la gestió dels recursos hídrics i la futura adaptació als impactes derivats de l'escalfament global.

INTRODUCCIÓ

1. Aigua, accés a l'aigua potable i microbiologia

L'aigua és imprescindible per entendre la vida al nostre planeta. És un recurs renovable que forma part dels diferents ecosistemes i és essencial a per l'evolució i la preservació d'aquests. La hidrosfera representa aproximadament un 71% del nostre planeta. Tot i que per quantitat és estable, l'aigua presenta un estat físic variable (sòlid, líquid i gasós), i distribució irregular amb variacions estacionals, podent formar part de l'atmosfera (vapor d'aigua), de la superfície terrestre (oceans, mars, maresmes, aiguamolls, estanys, rius, llacs i glaciers) o sota la superfície (aigües subterrànies). Del total d'aigua existent, la major part és salina, no obstant l'aigua dolça líquida que s'utilitza com a subministrament principal d'aigua potable, constitueix una fracció molt petita de l'aigua del planeta i amb volum irregular als diferents continents, essent superior a Àsia (3,7 milions km³) i molt inferior a Europa (541 mil km³), i Austràlia i Oceania (324 mil km³) (UNESCO, 2022). Al llarg dels anys, diversos estudis han estimat el volum global d'aigua dolça. Segons Shiklomanov i Rodda (2004), únicament un 2,5% correspon a aigua dolça, i la major part no és possible la seva utilització ja que en formen part els casquets polars i les glaceres d'alta muntanya (68,7%), i el permafrost (0,8%). No obstant, d'aquest petit percentatge, una important reserva d'aigua dolça es troba als aqüífers subterranis (30%) i la resta, molt més petita, està disponible com a aigua superficial i atmosfèrica (0,4%), que conjuntament representen un 0,7% dels recursos hídrics naturals d'aigua dolça disponibles (10,4 milions de km³). Estudis més recents apuntarien a que el volum global d'aigua dolça subterrània líquida seria molt superior al que s'havia estimat fins ara, amb un reservori de 15,9 milions de km³ (Ferguson et al., 2021).

Derivat de la nostra activitat (domèstica, industrial, agrícola, ramadera...) i a l'augment de la població, es genera una gran quantitat de productes de rebuig, entre els quals es troben els excrements humans i animals que sovint acaben contaminant el medi ambient i les aigües.

L'aigua amb contaminació microbiològica d'origen fecal ha estat històricament font de transmissió de malalties infeccioses com la colera, la diarrea, la disenteria, el tifus, l'hepatitis A o la poliomièlitis (WHO, 2018). Els patògens en l'aigua són un reflex de l'estat sanitari de les poblacions. Afortunadament al llarg dels anys s'han invertit fons públics i privats per desenvolupar nous sistemes de tractament d'aigua potable, subministrament d'aigües de consum, i sanejament ambiental per millorar-ne la qualitat i proveir aigua segura a la població. Els tractaments d'aigua aplicats a les estacions de tractament d'aigua potable (ETAP), formen una estratègia multibarrera capaç de produir aigua de beguda eliminant tant substàncies contaminants com microorganismes patògens que poden ser perjudicials per la salut humana i

animal, mentre que a les estacions depuradores d'aigües residuals (EDAR), es tracta l'aigua després dels diferents usos (residual, domèstica, industrial, etc.) disminuint els microorganismes patògens abans del seu retorn al medi natural de manera segura, o reutilitzant-la per diferents finalitats (mediambientals, agrícoles, urbanes, industrials, recreatives) (Real Decreto 1620/2007, 2007; Regulation (EU) 2020/741, 2020). Aquests avenços en els diferents tractaments, com ho va ser inicialment la cloració de l'aigua en les ETAP, han provocat una ràpida caiguda de les taxes de mortalitat per malalties de transmissió hídrica i els episodis de contaminació s'han anat reduït gradualment millorant la salut de la població.

La microbiologia de l'aigua, mitjançant l'ús d'indicadors de la contaminació fecal, ha estat clau per la millora i el control de la qualitat en l'aigua permetent avaluar el risc de la possible presència de patògens derivada d'una contaminació fecal. No obstant, en el present es calcula que a l'any es produeixen a prop de 500.000 morts (WHO, 2022) al món conseqüència del consum d'aigua contaminada, i és que 3 de cada 10 persones al món encara no tenen accés a aigua potable. Reduir la contaminació hídrica és un dels nostres majors reptes. Per això, en el marc dels Objectius de Desenvolupament Sostenible (ODS) de les Nacions Unides, l'objectiu 6 (aigua neta i sanejament) (United Nations, 2020) vetlla per un accés universal a l'aigua potable perquè sigui suficient, segura, assequible i accessible per a l'any 2030. Els esforços per garantir aigua potable a tot el món estan en constant evolució i requereixen un treball interdisciplinari a diferents nivells: recerca, polítiques i estratègies de gestió integral de l'aigua.

2. Estrès hídric i escalfament global a la regió mediterrània

El desenvolupament socioeconòmic de les activitats humanes està vinculat a un increment del consum d'aigua. Aquest, per diversos factors com l'augment de la població arrelada a una expansió industrial i agrícola-ramadera, i intensificada amb els efectes d'un escalfament global de l'atmosfera terrestre directament relacionat amb les millores de benestar de les poblacions humanes, fan preveure que s'agreuja la disponibilitat dels recursos hídrics en els propers anys provocant escassetat d'aigua. De fet, segons l'informe de l'Organització de les Nacions Unides (ONU) del 2022 (UNESCO, 2022), es preveu que el consum d'aigua augmenti un 1% anualment durant els propers 30 anys a causa de la demanda principalment dels sectors industrial i domèstic que podria induir un estrès hídric. Cal entendre l'estrès hídric com un dèficit de disponibilitat d'aigua en dos vessants diferents i que es donen conjuntament: perquè la demanda és superior als recursos disponibles o perquè el seu ús es veu restringit per la baixa qualitat (FAO and UN Water, 2021).

Alguns indrets són més susceptibles a experimentar estrès hídric, com ho és la conca Mediterrània (Giorgi and Lionello, 2007; MedECC, 2020), caracteritzada pel binomi de forts períodes de sequera i pluges torrencials que ha donat lloc a grans inundacions (Petrovic et al., 2010) marcant la vulnerabilitat del subministrament d'aigua fins i tot als països industrialitzats del nord de la Mediterrània (Burak and Margat, 2016; Petrovic et al., 2010).

Les projeccions del canvi climàtic preveuen un augment de la temperatura de 2,7°C per l'any 2100 amb les polítiques actuals (Stockwell et al., 2021), però la conca Mediterrània s'estima que s'està escalfant un 20% més ràpid que la mitjana mundial que podria, per l'any 2050, augmentar la temperatura de 2 a 4°C i reduir les precipitacions un 30% (MedECC, 2020).

En un escenari d'escalfament global on s'espera que la temperatura de l'aigua augmenti i en conseqüència es produeixi una eutrofització de les aigües superficials, es podria produir un canvi potencial en la diversitat microbiana amb l'aparició d'alguns bacteris oportunistes a l'aigua potable que podrien comprometre la qualitat de l'aigua o suposar un risc per a la salut pública (Calero Preciado et al., 2021; Whitehead et al., 2009).

3. One Water : canvis en les percepcions de la qualitat, el valor i la utilitat de les aigües

Les sequeres, el canvi climàtic, l'augment poblacional i la limitació dels recursos hídrics naturals, any rere any donen lloc a la necessitat de continuar desenvolupant solucions innovadores per abordar els reptes als quals ens enfrontem. Un element clau és el canvi d'enfocament dels recursos hídrics tradicionals de subministrament que són cada cop més limitats, amb el que cal trobar recursos hídrics alternatius, com la dessalinització i la reutilització de aigües residuals tractades (Gleick, 2018). Les tecnologies actuals de tractament d'aigua ens poden permetre fer realitat aquesta transició hídrica produint aigua de qualitat independentment del seu origen. Tanmateix, l'aigua subterrània, el recurs invisible, és una gran solució per alleugerir l'estrès hídric i els futurs efectes del canvi climàtic ja que és pràcticament ubiqüa i representa una de les reserves d'aigua dolça més importants del planeta (UNESCO, 2022), amb el que esdevé necessària una gestió integral dels recursos hídrics. L'ús integral dels diferents recursos hídrics ha donat lloc a una anàlisi holística anomenada *One Water* (una aigua) (Owen, 2015; US Water Alliance, 2016) que neix de la necessitat d'interconnectar l'aigua per gestionar-la de manera integrada donant valor a totes les aigües (potable, residual, industrial, pluvial, regenerada, etc.) per resoldre els reptes de l'aigua per a una sostenibilitat futura.

Davant aquest escenari de gran heterogeneïtat d'origens hídrics de subministrament, el control microbiològic de l'aigua resulta de gran importància per tal de garantir la seguretat de l'aigua a la població.

La present tesi doctorat s'engloba en el marc del pla de Doctorats Industrials de la Generalitat de Catalunya en col·laboració amb el grup de recerca Microbiologia d'aigües Relacionada amb la Salut de la Universitat de Barcelona, i Aigües de Barcelona, empresa metropolitana de gestió del cicle integral de l'aigua, i té com a objectiu el control microbiològic mitjançant indicadors de la contaminació fecal i la caracterització de les comunitats bacterianes dels processos de potabilització de l'ETAP de Sant Joan Despí (Barcelona).

4. Tractament de l'aigua potable a Barcelona

4.1. *L'abastament de l'aigua a Barcelona: Antecedents de la captació superficial i subterrània*

L'abastament d'aigua potable a Barcelona i l'àrea metropolitana fins l'any 1955 consistia en la captació d'aigua subterrània dels aqüífers del Llobregat i Besòs com a única font d'abastament. Amb el creixement demogràfic i urbanístic a la ciutat Comtal i rodalies, la demanda d'aigua s'incrementa per sobre dels recursos disponibles als aqüífers, i la sobreexplotació dels pous genera un gran descens en els seus nivells provocant la salinització de l'aqüífer deltaic del Llobregat deixant un 30% inutilitzable (Marcé et al., 2012). Per tant, esdevé necessari incrementar la producció d'aigua potable que no seria possible únicament amb l'explotació dels recursos hídrics subterranis. Així, el 1955, es posa en funcionament l'ETAP de Sant Joan Despí per tal d'utilitzar l'aigua superficial de la vall baixa del riu Llobregat com a font principal de captació per atendre la demanda d'aigua a la ciutat. Durant la dècada del 1970, la capacitat de tractament de l'ETAP ja s'havia doblat però no seria suficient per donar servei a la ciutat a les posteriors dècades, amb el que calia obtenir nous recursos hídrics. D'aquesta manera es posa en marxa una planificació hidràulica per a derivar cap a l'àrea metropolitana de Barcelona els cabals del riu Ter, mitjançant la construcció dels embassaments de Sau, Susqueda i el Pasteral (Armengol and Dolz, 2009) i un sistema de transvasament d'aigua de 56km de longitud des d'aquests fins a l'ETAP de Cardedeu (ATL, n.d.), així com els cabals del riu Llobregat aigües amunt amb l'ETAP d'Abrera (ATL, n.d.).

4.2. *Particularitats del riu Llobregat al punt de captació de l'ETAP de Sant Joan Despí*

El riu Llobregat, amb una longitud de 165 km, té origen a Castellar de n'Hug, als Pirineus catalans, i desemboca al El Prat de Llobregat, al mar Mediterrani. Des d'un punt de vista hidrològic, és un

riu típic mediterrani, amb un cabal caracteritzat per una gran variabilitat influenciada per períodes de precipitacions intenses i sequeres. Els trets distintius comencen pel seu cabal d'aigua, que pot passar de llargs períodes per sota de $5\text{m}^3/\text{s}$ a transicions abruptes que poden augmentar el cabal per un factor de 50 ($\sim 300\text{m}^3/\text{s}$) (Marcé et al., 2012). El cabal mitjà és de $22\text{m}^3/\text{s}$ (700hm^3 anuals) (Petrovic et al., 2010), un curs petit que es veu sobreexplotat per la gran demanda hídrica d'una població de més de 3 milions d'habitants només a Barcelona i rodalies, sumats a la resta d'habitants de les poblacions establertes a la conca des de Castellar de N'Hug, així com l'agricultura, la ramaderia, la indústria i el manteniment del Delta del Llobregat.

La vall mitja i baixa de la conca del riu Llobregat es veu afectada per diversos factors que conflueixen a les proximitats del punt de captació de la planta i riu amunt. El riu està sotmès a una gran pressió antropogènica derivada de les activitats industrials, agrícoles, la mineria de potassa de Cardona, Súria i Sallent i les depuradores aigües amunt que aboquen els seus efluent al riu i constitueixen una part important del seu flux (de mitjana, $137\text{hm}^3/\text{any}$ d'abocaments dels que un 92 % prové de la depuració d'aigües residuals) (Eljarrat and Barceló, 2012).

El riu Llobregat a prop de la zona de captació de l'ETAP té molts afluents torrencials, entre ells la riera de Rubí. El seu cabal ($0,75\text{m}^3/\text{s}$ de mitjana) és relativament baix en comparació amb el riu Llobregat, però en cas de fortes pluges el seu cabal pot augmentar ràpidament. La seva qualitat és molt baixa ja que condueix principalment aigües residuals tractades, tant d'origen urbà com industrial, així com pluvials. Per tal d'evitar que aquesta riera contaminada alteri la qualitat de les aigües superficials del Llobregat al punt de captació, l'aigua de la riera de Rubí es desvia a través d'un sistema de bypass que inclou antics canals de reg i infraestructures construïdes específicament amb aquest objectiu fins aigües avall de la captació de l'ETAP. No obstant això, en episodis de pluja intensa, les aigües de la riera de Rubí pot desbordar i s'incorporen al riu Llobregat, impactant negativament la seva qualitat i provocant sovint la interrupció del tractament d'aigua superficial a la potabilitzadora.

D'altra banda el riu Llobregat ha estat, des de començament del s. XX, afectat per una constant i elevada salinització a causa de l'extracció industrial de potassa aigües amunt de les captacions per abastament a Barcelona per les aportacions d'aigües extremament salines procedents de les mines de sal de Cardona, Súria i Sallent (Gorostiza Langa, 2014). Per tal de fer front als impactes negatius de les activitats mineres sobre els ecosistemes aquàtics fluvials i la qualitat de l'aigua, l'any 1989 entra en funcionament un col·lector de salmorra de 120 km de longitud que discorre paral·lel al riu Llobregat i condueix les salmorres al mar. Aquest col·lector també

recull els rebuigs del tractament per membranes (electrodiàlisi reversible a l'ETAP Abrera i osmosi inversa a l'ETAP Sant Joan Despí) que ha estat necessari implementar per eliminar les aportacions salines no derivades per aquest sistema. El col·lector de salmorra ha permès una clara millora de la qualitat de l'aigua del riu Llobregat (Fernández-Turiel et al., 2003; Martín-Alonso, 1994) i, conjuntament amb la necessitat d'instal·lació dels tractaments d'osmosi inversa de l'ETAP operatius des del 2007, s'ha aconseguit reduir la salinitat de l'aigua potable produïda a partir d'aquest recurs i, addicionalment, assegurar el compliment tècnic de la normativa sobre Trihalometans (THM) a l'aigua potable, generats a causa tant de la elevada salinitat del recurs com del seu important contingut en matèria orgànica. No obstant, el problema de la salinització del riu actualment persisteix per l'augment de la producció de potassa i per tant dels efluent salins posterior a la construcció del col·lector, que continua salinitzant el riu, a banda que al llarg dels anys s'han produït nombroses fuites del col·lector i trencaments parcial d'aquest provocant greus problemes locals per l'abocament de salmorres. Quan això succeeix, i per evitar que aquest vessament impacti la qualitat de l'aigua tractada, l'ETAP atura la captació superficial fins que es repara la fuga i el riu es recupera fins a les seves condicions precedents .

El conjunt de tots aquests factors determinen la complexitat dels tractaments de potabilització de l'ETAP de Sant Joan Despí, que actuaran de barrera als diferents perills que puguin comprometre la seguretat de l'aigua potable. Per això, la planta disposa d'una estació d'alerta al punt de captació amb analitzadors en línia que mesuren diversos paràmetres de la qualitat de l'aigua del riu (terbolesa, amoni, conductivitat, carboni orgànic total (TOC), níquel, UV 254 nm) per detectar possibles anomalies que puguin derivar ràpidament en un canvi d'operativa de planta per evitar situacions de risc. Per exemple, alguns paràmetres de consigna d'aturada de la captació superficial són l'amoni i la terbolesa, i es dona quan l'amoni al riu és superior a 2,5 ppm o si la terbolesa és superior a 500 NTU. Davant d'aquest escenari considerat de risc, s'atura la captació d'aigua superficial per evitar qualsevol possible impacte en la qualitat de l'aigua tractada i per evitar l'embrutiment dels processos de tractament. Aquest escenari es pot donar per diversos motius, principalment climatològics i antropogènics, com fortes pluges que aixequen els sediments del riu o causen intrusions d'aigües contaminades al riu per sobreiximents dels sistemes de sanejament, o per abocaments irregulars de diferents orígens (industrials, residuals, etc.). La planta però no atura totalment la seva producció, sinó que procedeix a tractar temporalment només l'aigua subterrània fins que la qualitat del riu torna als paràmetres de normalitat.

Per alleugerir aquestes particularitats que impacten la captació de l'ETAP i que afecten la qualitat de l'aigua del riu, com ja s'ha comentat anteriorment existeixen diverses

infraestructures connectades que deriven aquestes aigües contaminades aigües per sota del punt de captació (**Figura 1**), tal com el Canal de la Infanta (que deriva les aigües contaminants de la riera de Rubí i el riu Anoia cap a usos agrícoles), el Tub del Governador (assegura el bypass aigües avall de la captació de l'ETAP de les aigües contaminants sobrants al sistema de rec del Canal de la Infanta) o el Col·lector de Salmorres.

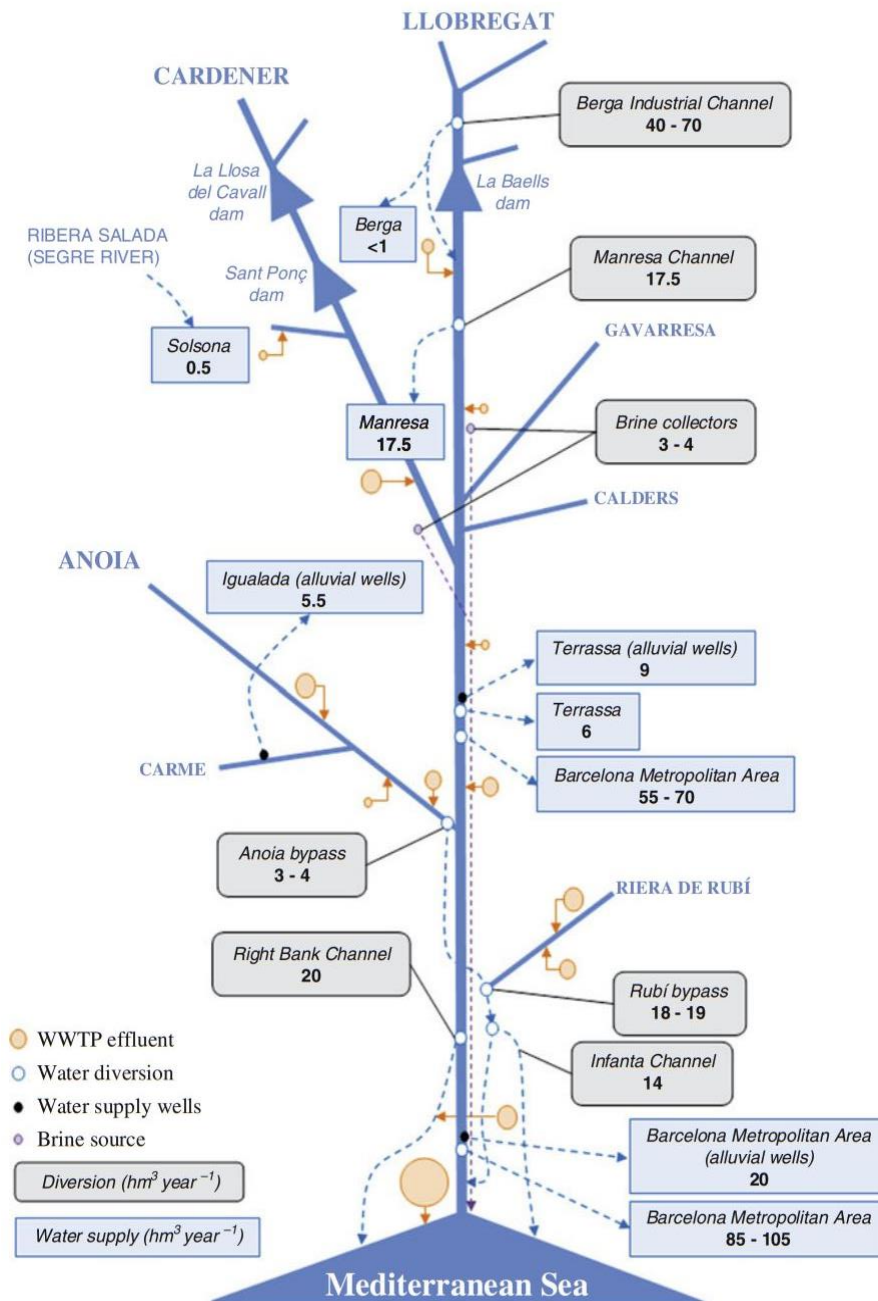


Figura 1. Esquema de les principals captacions (requadre blau), derivacions (requadre gris) i efluents d'aigües residuals (cercle taronja) de la conca baixa del riu Llobregat. Les fletxes blaves discontinües indiquen la direcció del flux de l'aigua o bé des dels punts de captació cap a les corresponents poblacions (requadre blau) o bé dels punts de derivació cap al mar. Els números indiquen el cabal en $hm^3/anyals$, i la mida del cercle taronja és proporcional al volum de l'efluent. Extret de Marcé et al. (2012)

4.3. Potabilització i distribució. La ETAP de Sant Joan Despí en l'actualitat

Des de la construcció de l'ETAP de Sant Joan Despí el 1955, la tipologia de tractament de la planta, així com les infraestructures de suport, s'han anat ampliant i evolucionant paral·lelament als estàndards de qualitat legistats per les corresponents directives i regulacions d'aigües de consum, la qualitat de l'aigua del riu, les particularitats que influeixen en la localització de la planta, la demanda hídrica i el context climàtic, entre d'altres. Així, on en els seus inicis es disposava únicament d'un pretractament i cloració, actualment es tracta d'una de les ETAP més complexes i avançades a nivell d'Europa, que està operada per Aigües de Barcelona, empresa metropolitana de gestió del cicle integral de l'aigua.

La captació de l'aigua superficial del riu Llobregat es realitza mitjançant unes reixes de formigó de 8 mm de gruix situades al llit del riu que donen pas a una galeria excavada amb comportes automàtiques que dirigeix l'aigua cap a l'interior de la central de tractament **Figura 2**.

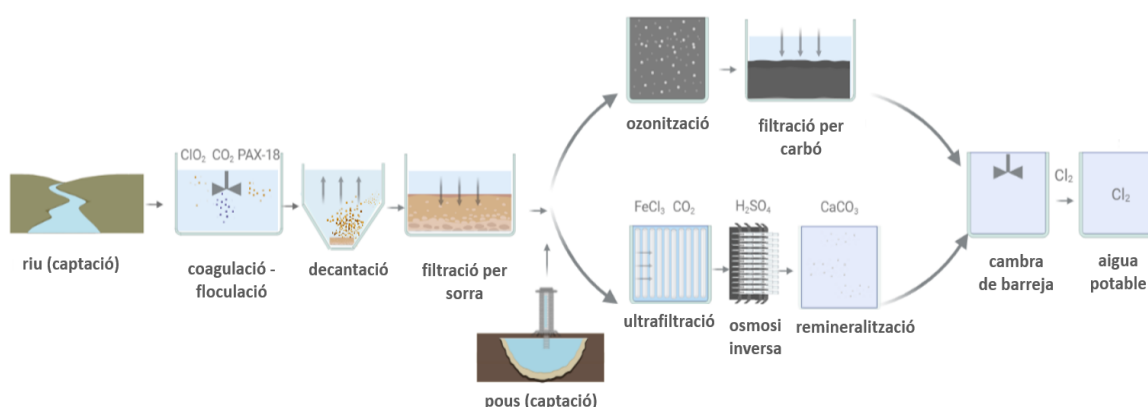


Figura 2. Esquema de les etapes del tractament de potabilització de l'ETAP de Sant Joan Despí

La primera etapa consisteix en un desarenat mitjançant l'ús de dragues per a l'eliminació de materials de gran mida (ex. graves). A continuació l'aigua s'eleva mitjançant un bombament cap a una cambra on es produirà la primera fase del pre-tractament: la coagulació de partícules flotants acompanyada de la primera desinfecció. Aquesta fase consisteix en la injecció de gas de diòxid de carboni (CO_2), per regular el pH adient pel procés de floculació, l'addició del reactiu coagulant basat en sals d'alumini (PAX-18) que aglutinarà les partícules en suspensió formant agregats, i la pre-oxidació amb diòxid de clor (ClO_2) que actuarà doblement com a desinfectant i oxidant de la matèria orgànica i certs metalls (ferro i manganès).

Aquesta aigua crua amb els reactius es fa circular per uns canals a cel obert paral·lels als decantadors que vessen cap a la fase de decantació. Es tracta de tancs en forma de piràmide invertida on la velocitat de l'aigua disminueix en funció de la secció que va travessant l'aigua, i on es produeixen dos processos paral·lels: una sedimentació de partícules coagulades que es dipositen com a fangs al fons del tanc, i una ascensió de l'aigua clarificada vers els canals de recollida existents a la part superior.

L'aigua clarificada es dirigeix cap a la fase filtració per sorra que permet l'eliminació de petites partícules a l'aigua i restes de flocs no eliminades per la decantació, així com de protozous com *Giardia* sp. i *Cryptosporidium* sp. i alguns bacteris resistents a la desinfecció com per exemple els clostridis. En el procés de filtració es fa circular l'aigua decantada a la superfície dels filtres, que s'infiltra per gravetat cap a l'interior del llit porós, constituït per sorra de granulometria molt homogènia. Per recuperar la capacitat filtrant, calen neteges periòdiques dels filtres de sorra per evitar la colmatació per saturació de partícules, l'afectació del rendiment del filtre amb l'augment de terbolesa o la formació de camins preferencials. D'aquesta manera, es fan rentats amb aigua a contracorrent cada 20.000-25.000 m³ d'aigua filtrada (variable en funció de els paràmetres d'exploració), i les aigües de rentat, conjuntament amb les purgues dels decantadors, es tracten a una planta de fangs destinada per aquest ús.

L'aigua filtrada per sorra surt clarificada per la part inferior dels filtres i es dirigeix cap a un dipòsit obert on es fa confluïr amb l'aigua subterrània provinent dels pous de l'aqüífer de Cornellà per tal d'optimitzar la qualitat de l'aigua tractada. La captació dels recursos subterranis es regula en funció de la qualitat de l'aigua superficial i permet cobrir la demanda d'aigua en cas d'una aturada de la captació superficial.

L'aigua resultant és bombejada mitjançant grans cargols d'Arquímedes que l'eleven cap a dues línies de tractament independents: el tractament convencional (ozonització i filtració per carbó actiu granulat), i el tractament avançat (ultrafiltració, osmosi inversa i remineralització).

El tractament convencional permet una millora de les propietats organolèptiques de l'aigua en quant l'olor, color i gust. A les cambres d'ozonització, l'ozó es difon en l'aigua que circula (mescla d'aigua filtrada per sorra i de pous), oxida els compostos orgànics, i actua com a biocida amb gran poder desinfectant per a la reducció dels microorganismes. L'aigua ozonitzada es distribueix per uns canals tancats que vessen a la part superior dels filtres de carbó amb la finalitat d'eliminar compostos orgànics i òxids metàl·lics com el ferro i el manganès. L'aigua circula per gravetat cap a l'interior del filtre, els compostos hi queden retinguts per l'elevada capacitat d'adsorció de les partícules del carbó i l'aigua clarificada surt per la part inferior del llit

filtrant. El rendiment dels filtres però va disminuint amb el seu funcionament i, per tal de mantenir un bon rendiment i assegurar l'eliminació dels compostos contaminants, es programen rentats a contracorrent per evitar la seva saturació, com succeeix amb els filtres de sorra. Amb la pèrdua de la capacitat adsorbent del filtre, detectable per la mesura de l'índex de iode, es duu a terme també una regeneració del carbó mitjançant forns de regeneració externs.

Les necessitats d'explotació defineixen el percentatge d'aigua que es deriva cap al tractament avançat (20-40%). Aquest s'inicia amb el procés d'ultrafiltració (UF) on, a l'entrada, hi ha una dosificació de clorur fèrric (FeCl_3) i CO_2 a l'aigua provinent de la mescla de filtració per sorra i pous. La ultrafiltració consta de cambres (trens UF) amb membranes d'ultrafiltració ($0,02 \mu\text{m}$ de diàmetre de porus) submergides a l'aigua que aspiren cap al seu interior reduint al mínim les partícules en suspensió, i retenint bacteris i alguns virus atenent la mida d'exclusió. L'objectiu principal d'aquesta etapa, i requisit pel bon funcionament de la osmosi inversa, és retenir l'alumini residual procedent del coagulant (PAX-18) de la decantació, que s'aconsegueix amb la disminució del pH de l'aigua acidificant-lo (6,6-6,8) amb l'addició de CO_2 per dificultar la seva solubilització i facilitar la retenció a les membranes d'ultrafiltració, mentre que el FeCl_3 s'addiciona per realitzar una post-coagulació de l'alumini residual afavorint la retenció a la UF i evitant que arribi a les membranes d'osmosi, on l'alumini podria provocar danys estructurals per obturació irreversible dels porus. Diàriament es realitzen operacions de rentat a contracorrent dels trens UF per millorar la capacitat de funcionament reduint les partícules que han quedat adherides a l'exterior de les membranes.

L'aigua permeada per ultrafiltració es bombeja cap al tractament d'osmosi inversa (OI), però prèviament a l'entrada, se li practica un pretractament per tal de protegir les membranes d'osmosi. Aquest pretractament consisteix en una desinfecció per radiació ultraviolada (UV), dosificació d'un agent dispersat, bisulfit sòdic i àcid sulfúric (H_2SO_4), filtració per cartutx (diàmetre de porus superior a UF) i de nou llum UV. La funció de l'UV i la filtració per cartutx és reduir o eliminar els microorganismes que han resistit els tractaments anteriors per evitar possibles recreixements al sistema de membranes d'OI. Respecte els reactius, el dispersant s'afegeix per reduir la cristallització de les sals de l'aigua que es concentren al rebuig del sistema, el H_2SO_4 reajusta el pH de l'aigua per evitar precipitacions i el bisulfit sòdic elimina els agents oxidants residuals.

L'etapa d'osmosi inversa consisteix en un bombament a alta pressió de l'aigua UF pretractada que travessa les membranes d'osmosi. Es tracta d'una membrana semipermeable que mitjançant pressió separa els constituents dissolts i deixa circular l'aigua permeada. Aquest

tractament suposa una gran barrera per bacteris i virus, i elimina de l'aigua les sals i pràcticament la totalitat dels compostos orgànics i inorgànics, assolint una conductivitat i carboni orgànic total (TOC) amb valors extremadament baixos. Aproximadament un 10% de l'aigua que entra a la osmosi correspondrà al rebuig generat pel funcionament de les membranes, que genera aigua concentrada en sals (salmorra). Aquest rebuig es descarta al col·lector de Salmorres de la conca del Llobregat.

L'aigua osmotitzada resultant és agressiva ja que per la seva desmineralització presenta una concentració de sals molt baixa i equivalent a l'aigua destil·lada ($\sim 100 \mu\text{S}/\text{cm}$), pel que es requereix una regulació del contingut salí. Així doncs, l'aigua es dirigeix cap a l'etapa de remineralització que consisteix en fer circular l'aigua osmotitzada per cambres amb llits de calcita (CaCO_3). En acabat, s'ajusta el pH de l'aigua remineralitzada amb hidròxid de sodi per assegurar que no sigui agressiva, i amb el pH estable es dirigeix a la cambra de mescla.

A la cambra de mescla hi conflueixen les aigües tractades de les dues línies de tractament: el convencional i l'avançat. En aquesta cambra es mesura la concentració d'amoni amb un analitzador automàtic per determinar la dosificació de clor que caldrà aplicar a les posteriors etapes per eliminar-lo o reduir-lo per sota de la consigna màxima de sortida (0,5 ppm), donat que l'amoni consumeix quantitats significatives de clor. Seguidament l'aigua de la cambra passa a un primer dipòsit de desinfecció (D4: 10.000 m^3) on es dosifiquen aproximadament 2 ppm de clor ($\sim 0,4 \text{ ppm}$ de clor lliure) amb un temps de contacte d'uns 30 minuts, assegurant la desinfecció de l'aigua i eliminant l'amoni. A continuació, l'aigua es condueix cap a un dipòsit d'estabilització (D3: 10.000 m^3) on s'efectuarà la cloració final ($\sim 1 \text{ ppm}$ de clor lliure) i l'estabilització de l'aigua abans de les impulsions a la xarxa de distribució. El temps de retenció en el dipòsit D3 és de 30 minuts.

L'aigua potable clorada final s'acumula en dos dipòsits (D1 i D2: 4.000 m^3), i mitjançant dues centrals de bombament, s'impulsa a quatre cotes d'alçada diferents (cota 10 m, 50 m, 70 m i 100 m) per proporcionar l'abastament a la xarxa de distribució garantint un mínim cost energètic. La impulsió d'aigua produïda és d'entre $3.000 - 3.500 \text{ L/s}$.

Els temps de retenció a cada etapa és relativament ràpid, ja que es triguen unes 6 hores a potabilitzar l'aigua des de la captació a la impulsió a la xarxa, però en algunes etapes el temps de contacte és superior als altres. Des de la captació del riu fins a l'efluent de la filtració per sorra és triga un total 3 de hores (2,5 h de decantació i 0,5 h a filtració per sorra). El temps de tractament convencional és més curt, amb 10 min d'ozonització seguits d'uns 23 min de filtració per carbó. El tractament avançat triga més temps, ja que requereix 42 minuts per la ultrafiltració

i 13 minuts per l'osmosi inversa, més uns 36 minuts de remineralització de l'aigua. Des de que l'aigua convergeix a la cambra de mescla, la posterior desinfecció amb clor, fins a poder impulsar l'aigua potable a la xarxa de distribució, passen aproximadament 2 hores.

5. Gestió integral de l'aigua

5.1. L'ús de recursos hídrics alternatius a Barcelona

Dessalinització

A la primavera del 2008, es va produir una forta sequera que va agreujar l'escassetat d'aigua a la ciutat de Barcelona, i va anar acompanyada de restriccions severes en l'ús de l'aigua i mesures excepcionals com el transport d'aigua potable des d'altres localitats i països en vaixells cap a Barcelona (ACA, 2019; Petrovic et al., 2010). Davant l'amenaça de l'escassetat d'aigua, l'estrès hídrics i la necessitat de consum (racional) d'aigua, s'adopten mesures alternatives de nous recursos hídrics. Així el 2009 entra en funcionament la dessalinitzadora ITAM al Prat de Llobregat (ATL, n.d.) que mitjançant osmosi inversa produeix aigua potable tractant aigua del mar, amb una producció anual de 60 hm³/any.

Reutilització

A la tardor de 2009, motivat per el període de sequera de l'any 2007-2008 que va provocar la promulgació del Decret de sequera a l'abril del 2007 i fins a la seva derogació (segona quinzena de gener del 2009), des de l'Agència Catalana de l'Aigua (ACA), actuant com a administració hidràulica, juntament amb l'administració sanitària, i amb l'entitat que explotava l'ERA del Prat (EMSHTR) i Aigües de Barcelona com a operadora de la ETAP de Sant Joan Despí, es va avaluar la incidència que podia tenir sobre el medi natural el fet de reutilitzar l'aigua regenerada procedent del tractament terciari de l'EDAR del Baix Llobregat, abocant-la a la llera del riu Llobregat a l'assut de Molins de Rei. Es va monitoritzar la qualitat de l'aigua regenerada a la sortida de l'ERA, en el mateix punt d'abocament, i aigües amunt i aigües avall del punt abocament, i es va concloure que la reutilització de l'aigua no generava alteracions importants en el medi receptor (Pérez et al., 2012; Rubiano et al., 2012). A l'estiu 2019 es va repetir un nou estudi sobre la qualitat de l'aigua regenerada en les mateixes condicions abocant de nou a l'assut de Molins de Rei. Els resultats van concloure el mateix: l'abocament aigua regenerada no deteriora la qualitat del riu en el punt abocament. En ambdues experiències es va dur a terme un extensiu pla de control que incloïa tant paràmetres fisicoquímics, químics i microbiològics.

Recàrrega artificial d'aqüífers

Des del 1969 s'injecta aigua potable als pous d'extracció per a la seva recàrrega, amb aigua potable per l'ús com a font de captació per potabilització, i amb aigua regenerada per evitar la intrusió salina. Recentment, es va provar a escala pilot la possibilitat de fer la recàrrega artificial dels pous amb aigua filtrada pel filtre de sorra per potabilització amb bons resultats tant a nivell operacional com de qualitat fisicoquímica i microbiològica (CETAQUA, 2018). Actualment mitjançant el projecte Queen (finançament Aigües de Barcelona) s'està posant a punt un nou model d'explotació de l'aqüífer que es preveu que en els propers anys operin injectant aigua filtrada per sorra per tal d'incrementar la disponibilitat de recursos subterranis de bona qualitat per cobrir la demanda d'aigua potable davant l'escenari d'estrès hídric.

5.2. Gestió i demanda hídrica a Barcelona

Per l'elevada densitat poblacional i la disponibilitat dels recursos hídrics, la gestió hídrica per l'abastament d'aigua actual a Barcelona és complexa ja que la xarxa de distribució rep aigua de diferents orígens. L'ETAP de Sant Joan Despí és la més rellevant en quant a complexitat i volum de producció a Barcelona, on es tracten fins a 5,3 m³/s d'aigua d'origen Llobregat.

Segons les dades d'explotació del 2020 (Aigües de Barcelona, 2020), l'aigua produïda, a l'ETAP de Sant Joan Despí procedia de la captació superficial del riu (85 hm³/any) i subterrània (34 hm³/any), i en menor proporció, de l'ETAP de les Estrelles i del Besòs, on es van tractar 6 hm³/any d'aigua subterrània dels aqüífers de Llobregat i Besòs, respectivament. Així el règim de producció de mitjana va ser de 342.000 m³ per dia (~125 hm³/any). Conjuntament amb la compra de cabals (~67 hm³/any) de tres orígens diferents principalment de la producció de l'ETAP d'Abrera (riu Ter) i l'ETAP de Cardedeu (riu Llobregat), i en menor mesura de la planta de dessalinització ITAM del Prat (aigua del mar Mediterrani) (ATL, n.d.), permet subministrar més de 192 hm³/any d'aigua potable a més de 3.200.000 habitants de l'àrea metropolitana de Barcelona.

La demanda anual d'aigua durant el 2020 a l'àrea metropolitana de Barcelona va ser lleugerament inferior (2,6%) en comparació a l'any 2019 degut a l'aturada de l'activitat industrial provocada per la pandèmia de la COVID-19. Els volums subministrats van ser de 158,7 hm³/any dels quals 116,4 hm³/any van correspondre a consum domèstic. El consum mitjà diari de les llars a l'àrea metropolitana va ser de 108,3 L/habitant.

6. Context legislatiu de la Qualitat de l'aigua de consum. Directiva EU 2020/2184

La legislació sobre qualitat de les aigües de consum es troba en un moment de canvi important: el desembre de 2020 es va aprovar la nova Directiva Europea (UE) 2020/2184, relativa a la qualitat de las aguas destinadas al consumo humano (Anonymous, 2020), que substitueix l'anterior Directiva 98/83/CE, adaptant-la al progrés científic-tècnic. En aquests moments, aquesta Directiva està en procés de transposició nacional, fet que suposarà la substitució del *Real Decret RD 140/2003, por el que se establecen los criterios sanitarios de la calidad del agua de consumo humano (versión consolidada 2018)*. El nou RD de transposició de la Directiva (UE) 2020/2184 ha d'entrar en vigor el gener de 2023.

La nova Directiva europea, amplia el focus de la directiva anterior incorporant tota una sèrie d'objectius globals i requeriments que van més enllà de la qualitat de l'aigua. Els grans objectius de la nova Directiva (UE) 2020/2184 són els següents:

- Millorar la qualitat i seguretat de les aigües, a través del control de nous paràmetres (addicionals als controlats fins a la data), tant químics com microbiològics
- Potenciar la gestió preventiva del risc (implantació de Plans sanitaris de l'aigua), ampliant-la a tot el cicle d'abastament, des de les captacions fins a les instal·lacions interiors
- Augmentar la confiança del ciutadà en l'aigua de consum i fomentar el seu consum
- Millorar la informació al ciutadà, tant sobre la qualitat de l'aigua consumida com sobre d'altres aspectes del servei
- Potenciar l'accés universal a l'aigua (Objectiu 6 dels ODS agenda 2030)
- Prevenir la contaminació en origen i minimització de tractaments

Pel que fa referència estricta a paràmetres de qualitat en l'àmbit microbiològic, la nova directiva també suposa un canvi important tant des del punt de vista d'incorporació de nous paràmetres (com *Legionella* i colífags somàtics), com des del punt de vista d'increments en la freqüència d'anàlisi d'alguns paràmetres: així, el recompte de bacteris aerobis a 22°C (HPC) i els enterococs passaran a fer-se en totes les anàlisis de *Control* i no únicament en les anàlisis *Completes*. A més a més, s'afegeix una nova mesura de control microbiològic després de cada neteja de dipòsits, concretament *Clostridium perfringens* (*C. perfringens*).

Actualment el control microbiològic de l'aigua de consum a nivell espanyol (segons RD 140/2003 consolidat) es basa en una verificació d'absència de la contaminació fecal mitjançant anàlisis d'indicadors de contaminació fecal com són coliforms totals i *Escherichia coli* (*E. coli*) en el cas

d'anàlisi de *Control*, o coliforms totals, *E. coli*, enterococs, *C. perfringens*, i recompte de bacteris aerobis heteròtrofs a 22°C (HPC) en el cas de les anàlisis *Completes*. Des d'aquest punt de vista, la nova Directiva europea representa un avenç en l'assegurament de la qualitat microbiològica, en quant incorpora paràmetres d'interès com *Legionella* (bacteri patogen en funció de l'espècie i de la concentració), i colífags somàtics com indicador de contaminació viral, que permet ampliar la seguretat en el cas d'aigües en què la desinfecció permeti eliminar la presència de tots els indicadors bacterians però no totalment la presència d'indicadors virals.

L'actual pla de control del procés de tractament i distribució d'Aigües de Barcelona, permet assegurar el compliment tant de la legislació vigent com de la nova Directiva (UE) 2020/2184, a més de disposar ja d'un Pla de Seguretat de l'Aigua (PSA) de tot el sistema (contemplant la producció i distribució de les aigües de consum), certificat per la norma de seguretat alimentària ISO 22000 des de l'any 2009.

Des del punt de vista legislatiu, tot i aquesta obertura a nous paràmetres microbiològics en el control de l'aigua de consum, sembla que encara queda lluny la incorporació de nous paràmetres derivats de noves tècniques analítiques que ja es poden considerar ben implantades a nivell metodològic, com les que, entre altres, es proposen a la present tesi doctoral i que no tenen altre objectiu final que portar a nivell operatiu el coneixement científic aplicat en altres àmbits i poder ser considerades tècniques complementàries a la legislació vigent oferint una visió global de la microbiologia del tractament de l'aigua.

6.1. Control microbiològic de la qualitat de l'aigua: indicadors de la contaminació fecal

La contaminació microbiològica de l'aigua es deu principalment a l'efecte de les activitats humanes i animals, suposant una especial preocupació per la qualitat de l'aigua les contaminacions microbiològiques d'origen fecal, i és que aquests residus aporten un gran nombre de microorganismes a l'aigua entre els que es troben els patògens que poden suposar un risc per la salut.

L'anàlisi en rutina de la totalitat dels microorganismes patògens presents en l'aigua no és viable degut, principalment, a la gran variabilitat que presenten, que la seva presència sol trobar-se en concentracions baixes, i que les tècniques per recuperar-los són complexes, laborioses, cares i de temps elevat. Per aquest motiu, la legislació vigent es basa en el control de microorganismes indicadors de contaminació fecal per garantir la innocuïtat del consum de l'aigua. Els bacteris indicadors de la contaminació fecal com els coliforms, *Escherichia coli* (*E. coli*) i enterococs s'han utilitzat històricament per avaluar la qualitat microbiana de l'aigua per la seva gran abundància en aigües residuals i femtes.

Els **indicadors microbiològics de contaminació fecal** són un grup de microorganismes relacionats amb la microbiota intestinal humana i d'animals de sang calenta. Per tant, la seva presència a l'aigua informa d'una deficiència de la qualitat per una contaminació fecal i alerta de la possible presència d'un microorganisme patogen. Així, els indicadors esdevenen una eina útil pel control de qualitat microbiològica de l'aigua permetent avaluar l'eficàcia dels processos de potabilització i de depuració. Un indicador microbiològic ha de complir els següents requisits (WHO, 2008):

- Estar relacionat ecològicament amb el patogen (femtes humanes i animals)
- No ser patogen
- Abundància superior al patogen
- Ser ràpid, senzill i econòmic de detectar pel mètode estandarditzat (cultiu)
- Resistència als processos de desinfecció i inactivació similar al patogen
- No present en aigües no contaminades

Per tal d'assegurar la innocuïtat de les aigües de consum a nivell microbiològic, actualment la Directiva, com s'ha esmenat a l'apartat 6, contempla l'assaig de 5 indicadors bacterians, 1 patogen bacterià de referència i 1 indicador víric. A la **taula 1** s'indiquen els paràmetres microbiològics i els valors paramètrics que han de complir les aigües destinades al consum (Anonymous, 2020).

TAULA 1. Paràmetres microbiològics, valors paramètrics i mètodes d'anàlisi segons la Directiva EU 2020/2184

Paràmetre	Valor paramètric	Mètode d'anàlisi
<i>E. coli</i>	0 UFC/100 mL	EN ISO 9308-1 ó 9308-2
Enterococs intestinals	0 UFC/100 mL	EN ISO 7899-2
<i>C. perfringens</i> (incloent espores)	0 UFC/100 mL	EN ISO 14189
Bacteris coliforms	0 UFC/100 mL	EN ISO 9308-1 ó 9308-2
Recompte de colònies a 22°C (HPC)		
a la sortida de l'ETAP	100 UFC/mL	EN ISO 6222
a la xarxa de distribució	Sense canvis anòmals	
<i>Legionella</i> (distribució domiciliaria) *	1000 UFC/L	EN ISO 11731
Colífags somàtics (a l'aigua crua)	50 PFU/100 mL	EN ISO 10705-2 i EN ISO 10705-3

* Davant un risc de *Legionella* per al seguiment de la verificació complementàriament es poden utilitzar mètodes ràpids alternatius de cultiu i mètodes moleculars (no basat en cultiu) com quantificació per qPCR(ISO/TS 12869)

Respecte els **indicadors bacterians**, els bacteris coliforms informen sobre les condicions sanitàries de l'aigua però, atès la seva gran ubiqüitat, per si sols no podrien utilitzar-se com a indicadors de contaminació fecal. En canvi, *E. coli*, tot i pertànyer al grup de coliforms, sí es troba estretament relacionada amb la microbiota intestinal i per tant constitueix un bon indicador de la contaminació fecal, així com permet indicar la possible presència de bacteris patògens entèrics com *Salmonella*, però no és un indicador especialment bo dels virus entèrics ni protozous ja que aquests són més resistents als tractaments d'inactivació. Tanmateix, els enterococs intestinals, tot i trobar-se en menor proporció a la microbiota intestinal humana, són considerats, conjuntament amb *E.coli*, els paràmetres nucli per alertar d'una contaminació fecal. D'altra banda, *C. perfringens*, també associat a femtes però en menor concentració que *E.coli*, representa un bon indicador (anaeròbic) de contaminació fecal i de l'efectivitat dels processos de tractament, ja que són formadors d'endòspores resistents a l'ambient, pel que la seva detecció a l'aigua potable indicaria problemes durant el procés de tractament. Per la seva resistència als processos de desinfecció es consideren bons indicadors de protozous paràsits com *Giardia* i *Cryptosporidium*. També, el recompte d'HPC és un indicador útil que permet monitoritzar canvis en la qualitat de l'aigua durant la seva potabilització o la seva distribució a la xarxa, com l'increment en nombre de les poblacions bacterianes per exemple en casos d'augment de l'activitat d'un biofilm, o en temps de retenció o estancament prolongats, o davant una ruptura de la integritat del sistema (Bartram et al., 2003).

Per altra banda, l'aparició de brots de malalties entèriques associades al consum d'aigua, en ocasions han estat relacionades amb aigües que presentaven paràmetres bacteriològics adequats. En aquest cas, la disposició d'**indicadors vírics** pel control de qualitat de l'aigua resulten de màxima importància com ja recull la nova Directiva. Així doncs, l'ús dels colífags somàtics (virus que infecten *E. coli* i altres coliforms intestinals) permet disposar d'un indicador de la contaminació fecal i de la presència de virus entèrics a l'aigua (Lucena and Jofre, 2014).

7. Comunitats bacterianes a l'aigua

Els ecosistemes d'aigua dolça són molt variables en quant a l'origen hídric (rius, aqüífers, llacs, etc.) i les condicions disponibles pel creixement microbià. Aquests medis aquàtics tenen una microbiota pròpia i alberguen una de les majors diversitats de tots els medis naturals (Thompson et al., 2017; Vitorino and Bessa, 2018; Walters and Martiny, 2020). Malgrat l'origen de l'aigua sigui diferent, s'han descrit aspectes comuns en quan els taxons majoritaris, que a nivell de fílums en aquests ecosistemes són principalment Proteobacteria, Bacteroidota,

Actinobacteriota, Firmicutes i Cyanobacteria (Aizenberg-Gershtein et al., 2012; Atnafu et al., 2021; Chao et al., 2015; Hou et al., 2018; Pinto et al., 2012; Thom et al., 2022; Vaz-Moreira et al., 2014; Wolf-Baca and Piekarska, 2020).

No obstant, les activitats antropogèniques poden alterar aquesta microbiota natural amb la introducció de microorganismes propis del intestí humà mitjançant la contaminació d'origen fecal a les aigües.

Les aigües dels processos de potabilització i les aigües de consum, tot i estar sotmeses a un tractament, i fins i tot complint els estàndards de qualitat microbiològics, són ecosistemes complexos que presenten una gran diversitat bacteriana (Hou et al., 2018; Pinto et al., 2012; Revetta et al., 2010; Wang et al., 2018). Malgrat que les concentracions bacterianes a la sortida de la planta de tractament solen ser baixes (<10 UFC/mL), es poden produir recreixements durant la seva distribució.

L'estudi d'aquesta diversitat pot ser de gran utilitat per millorar la qualitat de l'aigua, no obstant el seu control queda al marge de la legislació. Aquestes comunitats no tenen perquè suposar un risc per la salut però donades certes condicions, podrien donar lloc a l'aparició de bacteris que alterin les propietats organolèptiques de l'aigua i en alguns casos podrien comprometre la seva seguretat.

La gran majoria de bacteris que es troben als sistemes d'aigua són bacteris heteròtrofs que fan servir com a font d'energia el carboni orgànic. Alguns dels gèneres més abundants a rius, llacs, aigua potable i aigües residuals són: *Acidovorax*, *Acinetobacter*, *Aeromonas*, *Bacillus*, *Clostridium*, *Comamonas*, *Curvibacter*, *Enterobacter*, *Flavobacterium*, *Mycobacterium*, *Legionella*, *Sphingomonas*, *Pseudomonas*, *Rheinheimera*, *Rhodococcus*, entre d'altres (**Figura 3**).

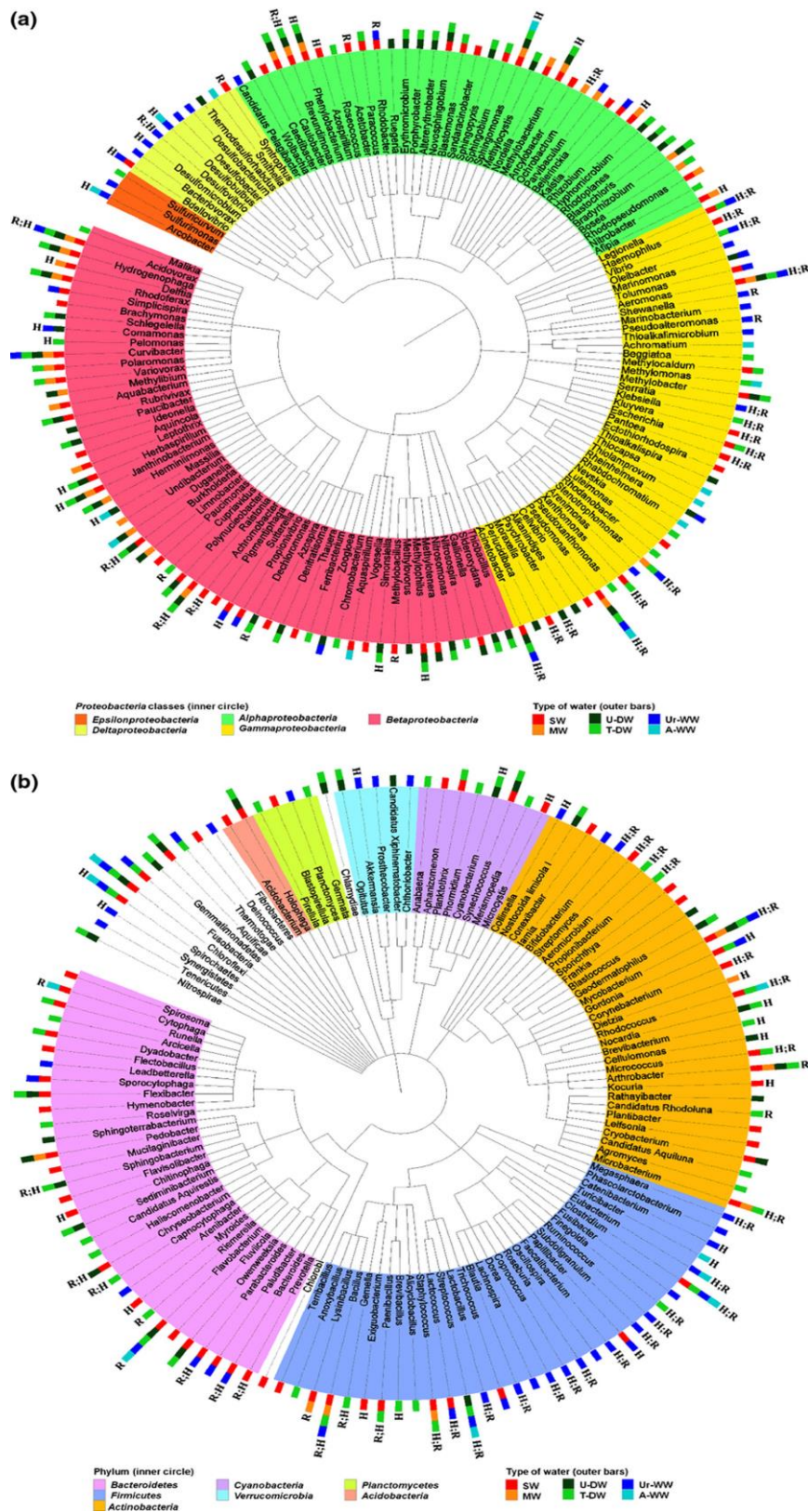


Figura 3. Representació gràfica de la diversitat bacteriana (a) Proteobacteria i (b) altres fílums observades diferents tipus d'aigua. Les diferents classes i fílums estan representades per colors diferents. A la part externa del cercle està representat amb barres les diferents matrius d'aigua (diferents colors) on s'han aïllat: : SW (aigües superficials: aiguamolls, rius i llacs); MW (aigua potable mineral); U-DW (aigua potable no tractada); T-DW (aigua potable tractada); Ur-WW (aigües residuals urbanes) i A-WW (aigües residuals animals). Extret de Vaz-Moreira et al. (2014).

No obstant també s'ha descrit la presència de bacteris autòtrofs que utilitzen una font d'energia inorgànica com els bacteris nitrificants (ex.: *Nitrosomonas*) que s'ha observat que recreixen rere l'ús de la desinfecció per cloramines o clor (Eichler et al., 2006; Regan et al., 2003; Thom et al., 2022), els oxidants de ferro (ex.: *Leptothrix*) que alguns gèneres estan associats a una acceleració de la corrosió de ferro (Sun et al., 2014), o els reductors de sofre (ex.: *Desulfovibrio*) aquests últims produint episodis d'olor i gust a l'aigua potable (Scott and Pepper, 2010).

Els canvis en la qualitat microbiològica de l'aigua són el resultat d'interaccions complexes entre els diversos microorganismes (bacteris, virus, etc.) i la influència de diversos factors que regulen el seu creixement, principalment l'efectivitat dels processos de potabilització, la disponibilitat i concentració de nutrients, la temperatura i la presència de desinfectants residuals a l'aigua.

Mitjançant els tractaments de potabilització a les ETAP s'aconsegueix reduir significativament la concentració de microorganismes a l'aigua final amb l'objectiu de proveir una aigua de qualitat i segura pel consumidor, no obstant es poden produir canvis durant la seva distribució a la xarxa que alterin l'estabilitat biològica de l'aigua. El concepte "estabilitat biològica" fa referència a un canvi mínim en la qualitat de l'aigua durant la distribució respecte la qualitat a la sortida de l'ETAP que no afecti la seguretat o alteri les propietats organolèptiques de l'aigua. Per aconseguir una aigua estable en el temps i l'espai, i limitar el creixement bacterià durant el transport, en molts països l'aigua potable es distribueix amb desinfectants residuals com el clor, que permet limitar el recreixement d'aquests a l'aigua potable. Diversos autors han assenyalat que nivells superiors a 0,2 ppm de clor lliure a la xarxa de distribució d'aigua potable redueixen significativament la microbiota i són adequats per a evitar la proliferació de microorganismes i la formació de biofilms (Fish and Boxall, 2018; Li et al., 2017, 2018). No obstant, els subproductes de desinfecció generats per l'ús de clor així com l'alteració del gust de l'aigua han fet que alguns països optin per distribuir aigua sense desinfectant residual. Així, per exemple als Països Baixos on es disposen de sistemes de distribució d'aigua potable sense cloració, l'estabilitat es basa en la pròpia microbiota de l'aigua que actuaria com a protecció natural contra la proliferació de patògens (Roeselers et al., 2015). Aquest efecte protector dels bacteris que colonitzen de manera natural el sistema d'aigües és un concepte que es coneix com a biofilm protector que tindria un efecte de protecció envers la presència de patògens a l'aigua (Wang et al., 2013). No obstant, als Països Baixos, la font principal de captació per a la producció d'aigua potable és l'aigua subterrània que es veu menys alterada a nivell microbiològic (Roeselers et al., 2015).

D'altra banda, el tipus de nutrients orgànics i inorgànics disponibles i les concentracions d'aquests a l'aigua potable també estan relacionades amb l'estabilitat biològica, ja que poden limitar o promoure el creixement bacterià a l'aigua. Els bacteris heteròtrofs, que constitueixen la majoria dels bacteris de l'aigua potable, principalment obtenen energia per la degradació dels compostos de carboni orgànic. El carboni és el compost que més limita el creixement bacterià, seguit dels nutrients inorgànics com el nitrogen i el fòsfor (Prest et al., 2016). Els oligoelements (ferro, magnesi, coure, etc.), també són necessaris per al creixement dels bacteris heteròtrofs però en concentracions molt menors que el carboni orgànic (Egli, 2010).

Tanmateix, una altre factor molt important és la temperatura que conjuntament amb la disponibilitats de nutrients té un efecte directe en el creixement de les poblacions bacterianes. Diversos autors han observat que davant d'una temperatura elevada i nivells de nutrients elevats, es produeix un augment de la taxa de creixement de les comunitats bacterianes, mentre que en un ambient amb baixes de concentracions nutrients el creixement bacterià disminueix (Degerman et al., 2013; Hou et al., 2018; Pinto et al., 2012). No obstant, la composició de nutrients als processos de tractament i a l'aigua potable és complexa, ja que disposa de diversos nutrients però a concentracions molt baixes pel que els bacteris que han resistit els tractaments estan ben adaptats i prolifereixen en aquest ambients oligotròfics, metabolitzant substrats de carboni diferents simultàniament per a la seva supervivència (Egli, 2010; Li et al., 2017; Pinto et al., 2012).

Així doncs, davant la gran heterogeneïtat de recursos hídrics disponibles per a la potabilització, les diferents estratègies de tractament i distribució de les aigües, així com els diferents factors que poden influenciar en aquests (aportació d'efluents residuals i industrials, temperatura, nutrients, ús de desinfectants, etc.) fa necessària la caracterització de les poblacions microbianes a l'aigua.

El coneixement sobre la microbiota autòctona pot suposar una eina molt útil per a identificar canvis en els punts de captació, en les etapes de potabilització així com en l'emmagatzematge de l'aigua en dipòsits o a la xarxa de distribució.

En els darrers anys, s'han desenvolupat mètodes analítics i moleculars d'alt rendiment que permeten una caracterització detallada de les comunitats bacterianes a l'aigua. Prèviament però, es requereix disposar d'una mostra representativa amb el que esdevé necessari disposar de mètodes de concentració.

7.1. Mètodes de concentració de microorganismes

Un paràmetre crític per l'estudi de les comunitats bacterianes a l'aigua és el procediment de mostreig que ha de garantir la integritat i la representativitat de la mostra. Per tal de detectar la microbiota que es troba a l'aigua cal fer una concentració de la mostra prèvia a l'anàlisi. No obstant, molts microorganismes com per exemple els patògens, es troben habitualment a baixes concentracions a les aigües, amb el que esdevé necessari emprar tècniques que permetin concentrar grans volums d'aigua per tal d'augmentar la sensibilitat en la seva detecció.

Existeixen diferents mètodes de concentració en mostres ambientals que varien en funció del tipus de microorganisme d'interès (bacteris, virus i protozous) (Bridle, 2014; Yates et al., 2016).

Tradicionalment s'han classificat com a mètodes de concentració primària (redueixen grans volums d'aigua a un volum final aproximadament de 1L) o secundària (o de reconcentració, permetent una reducció del volum de la mostra a pocs mil·lilitres) però a la pràctica la seva distinció és complexa ja que un mateix mètode pot formar part d'ambdues classificacions.

El més avantatjós és disposar d'un mètode que permeti simultàniament la concentració de virus, bacteris i protozous, que proporcioni recuperacions acceptables i consistents, que sigui de processament ràpid i senzill, reproduïble, que permeti processar diferents matrius d'aigua, i que sigui de baix cost.

Entre els diferents mètodes de concentració els més utilitzats es troben:

- **Centrifugació:** Emprat independentment per a bacteris, virus o protozous, segons el protocol, consisteix en compactar els microorganismes presents a la mostra en un pellet per sedimentació mitjançant forces gravitatòries que poden oscil·lar habitualment entre 1000 i 12.000 × g i temps variable (ex: 1100 × g 15 min per *Giardia spp.* en concentració secundària; 4000 × g 30 min per bacteris en concentració secundària després de concentració primària per ultrafiltració). La principal limitació d'aquesta metodologia resideix en la dificultat de concentrar volums relativament grans de mostra (1L) a grans velocitats sinó se'n fa ús de la ultracentrifugació, necessari per a la concentració de virus. També es pot emprar per reconcentració després d'una concentració primària per reduir el volum final de la mostra a pocs mil·lilitres o microlitres.
- **Floculació-Precipitació:** Emprat per a la concentració de virus, aquesta tècnica consisteix en l'ús de floculants (alumini, llet descremada, polietilenglicol (PEG), etc.) en que s'adhereixen els virus en suspensió per a la posterior precipitació per sedimentació, sense la necessitat de

concentració secundària (8-12h), o bé amb centrifugació posterior (Calgua et al., 2008; Hjelmsø et al., 2017; Lewis and Metcalf, 1988; Randazzo et al., 2020). El mètode té una limitació de volum mostral (10L) ja que es dificulta el seu rendiment en volums grans de mostra.

- **Filtració**

- **Adsorció-elució:** Emprat per a la concentració de virus, també conegut com a VIRADEL (*viral adsorption and elution*), es tracta de concentrar les partícules víriques, que en medi aquós presenten càrrega electronegativa, per adsorció electrostàtica sobre la superfície un filtre (acetat de cel·lulosa, nitrat de cel·lulosa o llana de vidre) amb que reté el virus i posteriorment s'elueix amb tampons alcalins (extracte de carn, glicina, etc.) amb filtres que retenen els virus (Wallis and Melnick, 1967). Permet filtrar volums grans de mostra (10-50L). Hi ha diverses metodologies com els discs MAF (monolithic affinity filtration), la filtració amb llana de vidre (columna amb llana de vidre empaquetada a una determinada densitat que actua d'absorbent de virus), o el filtre electropositiu NanoCeram (Argonide, US) (combinació de microvidre i cel·lulosa que permet l'adsorció de virus, bacteris i protozous en grans volums de mostres de diferent terbolesa).
- **per membrana:** Aquesta tècnica permet concentrar bacteris, virus o protozous per separat (en funció del tipus de filtre), i consisteix en fer passar un volum de mostra d'aigua (<10L) a través d'un filtre de membrana (nitrocel·lulosa, policarbonat, acetat de cel·lulosa, etc.) microporós amb una determinada mida de diàmetre de porus (0,22 µm ó 0,45 µm per bacteris) on hi quedaran retinguts els microorganismes de mida superior. El seu bon rendiment depèn d'una baixa terbolesa a la mostra (adequat per aigua potable). Pot utilitzar-se en concentració primària o secundària. Actualment pel control de qualitat de l'aigua està estandarditzat l'ús de la filtració membrana mitjançant la concentració de volums de 100ml per tots els indicadors bacterians (excepte HPC).
- **Ultrafiltració:** És l'únic mètode de concentració, a dia d'avui, que permet la recuperació simultània de virus, bacteris i protozous basant-se en l'exclusió de mida amb recuperacions descrites superiors al 50% amb volums grans de mostra (Gunnarsdottir et al., 2020; Hill et al., 2007; Rhodes et al., 2011). Principalment es classifiquen en:
 - **CeUF:** La ultrafiltració centrífuga (CeUF), que consisteix en una concentració per centrifugació mitjançant un ultrafiltre amb una mida d'exclusió (5-100 KDa en funció del dispositiu), permet la concentració primària i/o secundària de tots els microorganismes. Durant la centrifugació, l'aigua i les molècules més petites passaran a través del filtre i els

microorganismes hi quedaran retinguts. A causa de la petita mida dels porus, les mostres amb elevada terbolesa (ex. riu, aigua residual, etc.) poden obstruir el filtre augmentant el temps de processament de la mostra. Un exemple és l'ultrafiltre Centricon® Plus-70 (Merck, Alemanya) (Gunnarsdottir et al., 2020; Medema et al., 2020; Rusiñol et al., 2020) que permet concentrar de manera ràpida (70 mL en 25 min) en el rang de 50X a 200X mitjançant repetits volums de càrrega de 70 mL, arribant a un volum final inferior a 350 µL.

- **DEUF:** La ultrafiltració sense sortida (DEUF) és basa en filtració per l'exclusió de mida (10-70kDa) amb volums grans de mostra (>100 L) (Gunnarsdottir et al., 2020; Hill et al., 2007; Rhodes et al., 2011). La tècnica DEUF permet concentrar grans volums d'aigua *in-situ* connectant directament a l'aixeta durant el mostreig a diferència de la resta de tècniques (excepte NanoCeram) que requereixen del transport de grans volums d'aigua al laboratori per a l'anàlisi. La tècnica és senzilla i permet concentrar diferents matrius d'aigua (regenerades, processos de tractament, riu, aigua potable, etc.). L'anàlisi de grans volums pot ser un enfocament eficaç per detectar microorganismes a baixes concentracions. Tot i que mitjançant el mètode d'adsorció-elució amb llana de vidre o NanoCeram permet la concentració simultània de bacteris, virus i protozous el primer mètode té una reproductibilitat variable i ambdós presenten una bona recuperació en virus (<80%) però baixa per bacteris i protozous (<12%) en comparació amb els mètodes d'ultrafiltració DEUF (>50%) (Francy et al., 2013).

Un exemple de filtre és l'ultrafiltre d'hemodiàlisi **Rexeed-25A** (Akahi Kasei Medical, Japó) (Gunnarsdottir et al., 2020; Pascual-Benito et al., 2020). Es tracta d'un cartutx de filtració de fibra buida (Polisulfona) de 185 µm i una mida de porus de 30 kDa que es pot connectar directament a una aixeta per la filtració *in situ* (o amb bomba peristàtica) permetent una concentració de mostra de fins a 100L/hora (depèn de la matriu d'aigua) tolerant una pressió màxima de 66 kPa que si se supera pot danyar les fibres, pel que sempre s'ha d'utilitzar durant la filtració un manòmetre per, un cop regulat el flux, comprovar que no augmenti la pressió. Permet concentrar volums grans de mostra (>1100 L). Durant la filtració, les partícules i microorganismes superiors a la mida de tall (30 kDa) queden retingudes a la superfície de l'entramat de fibres, mentre que l'aigua i partícules més petites circulen per l'interior de les fibres cap a l'exterior del filtre (permeat) (**Figura 5**). Un cop finalitzada la filtració, per tal de recuperar el concentrat, al laboratori es procedeix a fer un rentat a contracorrent amb una solució d'elució mitjançant bomba peristàtica que permet recuperar els microorganismes en un volum reduït de 500-600 mL que es pot reconcentrar per CeUF.

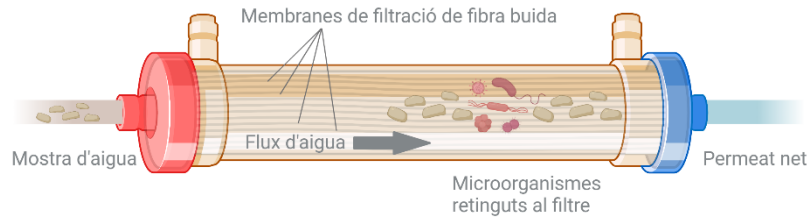


Figura 5. Esquema del procés de concentració d'un filtre d'ultrafiltració DEUF durant el mostreig d'aigua

El mètode DEUF per a la concentració de mostres d'aigua potable, així com altres matrius d'aigua, ja és una realitat que progressivament es va incorporant a estudis ambientals d'indicadors de contaminació fecal i patògens d'interès que es troben a baixes concentracions, per la seva alta capacitat de filtració, i recuperació simultània de tots els microorganismes (Cuevas-Ferrando et al., 2020; Ferrari et al., 2019; Gunnarsdottir et al., 2020; Pascual-Benito et al., 2020; Rhodes et al., 2011). No obstant, l'ús d'aquesta metodologia no està actualment explorada per l'estudi de les comunitats bacterianes per NGS ni per proteòmica a processos de potabilització i distribució d'aigües de consum. Per aquest motiu, la present tesi doctoral ha avaluat l'ús d'aquesta metodologia per a l'estudi de la diversitat bacteriana.

7.2. Mètodes de detecció i identificació de microorganismes

Una de les tasques fonamentals als laboratoris de microbiologia de l'aigua és l'aplicació d'una metodologia precisa i robusta que permeti detectar, enumerar i identificar els microorganismes implicats en processos de deteriorament o contaminació de l'aigua. Pel que fa a la qualitat de l'aigua, aquesta metodologia, està estandarditzada i recollida sota la Directiva europea com s'ha mencionat a l'anterior apartat 6 i es detalla a continuació a l'apartat 7.2.1.

Tradicionalment la microbiologia d'aigües ha emprat mètodes manuals basats en cultiu per la detecció, recompte, i identificació de microorganismes d'interès. En l'actualitat els mètodes que s'utilitzen pel control de qualitat microbiològica segueixen basant-se en mètodes dependents de cultiu, que tenen una elevada fiabilitat però una de les limitacions és el temps de resposta dels resultats ja que requereixen d'una incubació de 24 hores o superior en funció del microorganisme.

En els darrers anys, la microbiologia en l'àmbit ambiental i alimentari, ha fet un gran salt cap a l'automatització i ha donat lloc a mètodes ràpids, kits, miniaturitzacions (Fung, 2002), així com tècniques de seqüenciació d'alt rendiment (McDaniel et al., 2021; Zhang and Liu, 2019) per l'assegurament de la qualitat i per aprofundir en el coneixement sobre els microorganismes i les seves funcions fent ús d'aquestes noves eines com a monitorització més enllà del que està estrictament legislat. A la **Figura 6** es resumeixen els principals mètodes d'anàlisi de microorganismes que es troben actualment disponibles.

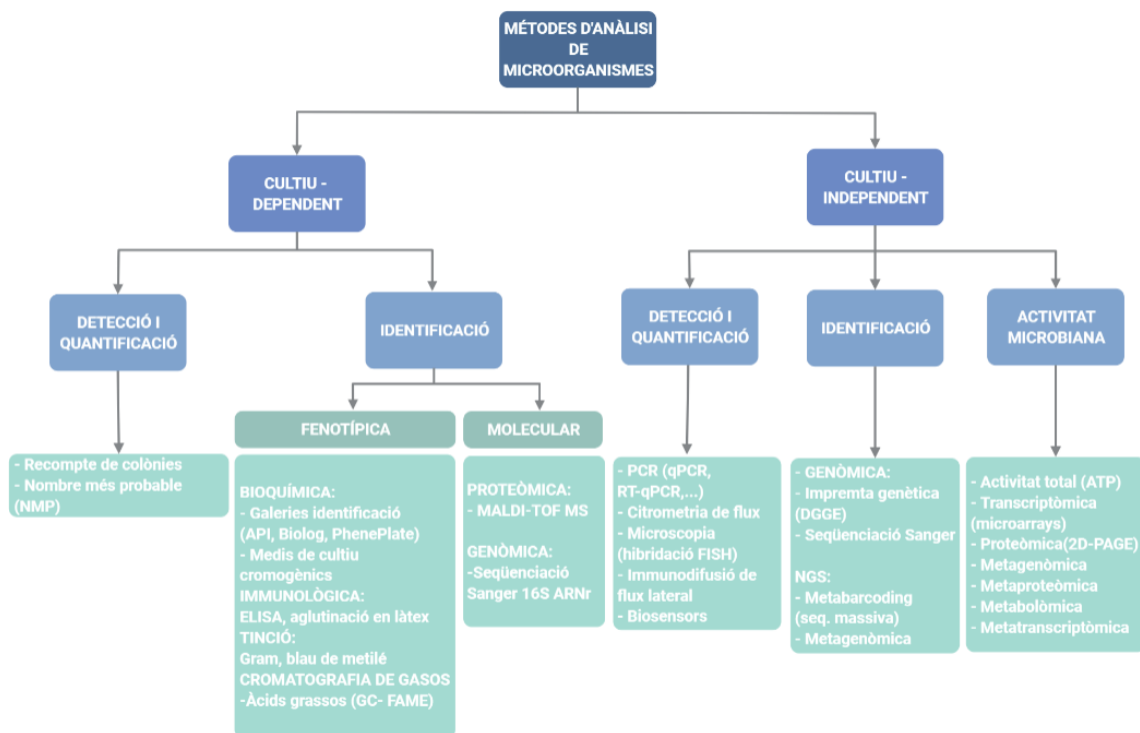


Figura 6. Diagrama dels principals mètodes d'anàlisi de microorganismes per la seva detecció, quantificació i detecció segons si es basen en cultiu o si són independents de cultiu

Les diferents tècniques permeten l'anàlisi de les comunitats bacterianes mitjançant dues estratègies diferents: dependent de cultiu o independent de cultiu. Ambdues aproximacions presenten diverses metodologies per a la detecció i/o quantificació (o semi-quantificació) dels microorganisme d'interès, així com la identificació i caracterització d'aquestes comunitats. Per exemple, per controlar la densitat microbiana a l'aigua potable, depenent del cultiu s'utilitza majoritàriament el recompte de colònies en placa, mentre que pel recompte cel·lular independent del cultiu, la citometria de flux. En canvi, per caracteritzar les poblacions bacterianes cultivables, es poden utilitzar aproximacions com la proteòmica per espectrometria de masses MALDI-TOF o per l'anàlisi de la totalitat de poblacions (cultivables i no cultivables)

tècniques de seqüenciació d'última generació (NGS) com per exemple la seqüenciació massiva d'amplicons (*metabarcoding*).

Adicionalment, les tècniques independents de cultiu també permeten analitzar l'activitat microbiològica pel que fa a la detecció de molècules o gens relacionats amb l'activitat microbiana, que per exemple mitjançant les tècniques *òmiques* ens permet veure quins microorganismes es troben, què poden fer, i quin microorganisme fa què (metagenòmica), o bé quin microorganisme està activament fent una funció determinada (metatranscriptòmica), o quin microorganisme està produint quins productes metabòlits (metabolòmica), etc.

A continuació es detallen les tècniques emprades en la present tesi doctoral per a l'estudi de les comunitats bacterianes als processos de potabilització i distribució de les aigües de consum.

7.2.1. Espectrometria de masses amb desorció/ionització làser assistida per matriu amb detecció de masses per temps de vol (MALDI-TOF MS) per a la identificació de bacteris

L'espectrometria de masses MALDI-TOF és una tècnica emprada originalment al camp de la química orgànica per l'anàlisi de biomolècules i polímers sintètics. Per primer cop, l'any 1975 Anhalt i Fenselau van suggerir l'ús de l'espectrometria de masses pel camp de la microbiologia, concretament per la caracterització bacteriana, mitjançant llargues extraccions amb cloroform-metanol d'algunes espècies de bacteris patògens a l'observar que aquests disposaven d'un perfil de masses únic. Més endavant, al 1996 Holland i altres autors aconseguiren per primer cop obtenir els espectres directament de cèl·lules bacterianes mitjançant l'ús de MALDI-TOF MS. Des d'aleshores es comença a introduir en el sector hospitalari i des de fa aproximadament 15 anys, aquesta tecnologia és àmpliament emprada en rutina per diagnòstic clínic, i més recentment s'ha anat implementant de forma progressiva a diverses disciplines del camp de la microbiologia amb un alt grau d'acceptació per la seva robustesa, rapidesa, baix cost i facilitat d'ús en la identificació bacteriana (Singhal et al., 2015).

Actualment, la tècnica MALDI-TOF MS té diverses aplicacions en diferents àmbits:

En l'àmbit clínic destaquen la detecció precoç de bacteris resistents a antibiòtics reduint el temps d'actuació per a la teràpia antibiòtica adequada (Florio et al., 2020), el diagnòstic clínic mitjançant la identificació de patògens en cultius de sang, orina, femtes, fluid cerebroespinal, etc. (Tsuchida et al., 2020), així com els estudis epidemiològics, per exemple pel seguiment de l'epidèmia causada recentment per la COVID-19 permetent un diagnòstic ràpid i fiable del virus SARS-CoV-2 en saliva (Deulofeu et al., 2021; Hernandez et al., 2021).

En l'àmbit alimentari, és molt emprat en rutina per garantir la seguretat i la qualitat dels productes, per exemple identificant els bacteris fermentadors d'aliments (Angelakis et al., 2011), o els bacteris oportunistes o patògens implicats en el seu deteriorament (Barreiro et al., 2010), o per a identificar possibles bacteris responsables de brots transmesos per aliments com el causat l'any 2011 a Alemanya per *E.coli* productora de toxina Shiga O104:H4 associada a germinats vegetals (Christner et al., 2014), etc. Tanmateix el seu ús es recomana com a confirmació i identificació de patògens transmesos per aliments i aigua com *Salmonella spp.*, *Campylobacter spp.* o *Cronobacter spp.* com alternativa als mètodes de referència basats en proves bioquímiques ja que presenten resultats equivalents i més ràpidament en comparació amb els procediments de confirmació especificats per cada mètode de referència (Bastin et al., 2019).

En el camp de la microbiologia ambiental aquesta tècnica s'ha explorat en diversos ambients com el marí, per exemple determinant els bacteris simbiotes de corals, esponges i peixos (Vidal et al., 2020) o per controlar el transport no desitjat de bacteris patògens a través de l'aigua de llast (Emami et al., 2012), a la rizosfera per identificar bacteris aïllats de sòls contaminats (Christner et al., 2014), o a aqüífers d'aigua dolça per determinar les espècies que viuen en ambients contaminats (Santos et al., 2017), etc.

Per tant, les aplicacions mitjançant l'ús de MALDI-TOF MS permeten monitoritzar diversos ambients amb el que pot resultar de gran utilitat pel camp de la microbiologia d'aigües en relació al control dels processos de producció d'aigua potable i la xarxa de distribució per una ràpida detecció dels bacteris que resideixen en aquestes matrius poc explorades actualment per aquesta tecnologia. Per exemple, davant d'un esdeveniment de contaminació microbiològica d'enterococs a la xarxa, les proves de confirmació regulades no estarien disponibles fins a 5 hores després d'obtenir el resultat presumptiu. Tot i que majoritàriament es respon amb les corresponents accions correctives davant la notificació d'un resultat presumptiu, amb MALDI-TOF MS en pocs minuts es podria confirmar la seva presència per tal d'actuar immediatament. Més enllà de la identificació de la presència d'indicadors, permet la identificació ràpida de les colònies que han crescut a les plaques dels controls de qualitat microbiològics de rutina legiscats, com per exemple els HPC, aportant una informació complementària que resultaria molt laboriosa, lenta i amb un cost més elevat mitjançant la identificació amb proves bioquímiques.

Plataformes MALDI-TOF MS i descripció de la tècnica

Actualment hi ha diverses plataformes de MALDI-TOF MS, entre les que destaquen Microflex LT (Bruker Daltonics, Alemanya), VITEK MS PRIME (BioMérieux, França), AXIMA (Schimadzu, EEUU) i Autof ms1000 (Autobio Diagnostics, Xina).

En la present tesi doctoral s'ha emprat aquesta tecnologia per a la identificació bacteriana de les plaques d'heteròtrofs (HPC) mitjançant l'equip Microflex LT MALDI-TOF MS (Bruker Daltonics, Alemanya).

La tècnica d'espectrometria de masses MALDI-TOF es basa en la detecció de la relació massa-càrrega de les proteïnes majoritàries, principalment ribosòmiques (2-20 KDa), dels bacteris que componen un perfil d'espectre de masses únic que permet la seva identificació. L'espectre adquirit mitjançant aquesta tècnica es compara amb una base de dades de col·lecció d'espectres de masses de soques de referència, anomenat també *llibreria*, i en funció de les similituds amb aquestes s'aconsegueix una identificació a nivell de gènere, espècie i fins i tot subespècies.

La metodologia requereix l'ús de material biològic de microorganismes cultivables partint d'una colònia d'un cultiu pur fresc. El procediment de preparació del material biològic per a la seva identificació pot comportar protocols diferents en funció del tipus de microorganisme en quant a la composició i estructura de la membrana o paret cel·lular, estat d'esperulació, etc. La preparació de mostres en general poden englobar-se en dos tipus de mètodes: transferència directa o extracció de proteïnes.

La transferència directa és el mètode més fàcil i ràpid i permet abordar aproximadament el 90-95% de les mostres segons el fabricant. Consisteix en dipositar sobre un suport metàl·lic (la placa d'anàlisi d'acer inoxidable) una colònia aïllada (o una porció d'aquesta) que es cobreix amb una solució orgànica d'àcid alfa-ciano-4-hidroxicinàmic (matriu HCCA), que penetra dins la paret cel·lular dels microorganismes i cristal·litza mentre s'asseca ràpidament a temperatura ambient, deixant lliures les proteïnes intracel·lulars.

El protocol de transferència directa ampliada amb àcid fòrmic és el més emprat per la rapidesa i efectivitat en l'obtenció de proteïnes (Theel et al., 2012) ja que, a l'incloure un pas previ afegint àcid fòrmic sobre la colònia abans de posar la matriu HCCA, ofereix una extracció parcial sobre la placa d'anàlisi millorant la lisi cel·lular i el corresponent alliberament de proteïnes. Abans de l'anàlisi requereix d'un ràpid assecatge de pocs minuts a temperatura ambient.

El mètode d'extracció de proteïnes és més exhaustiu pel que fa a la lisi cel·lular i permet l'extracció de les proteïnes de cèl·lules bacterianes mitjançant l'ús de dissolvents de diferents

tipus (Šedo et al., 2011). Un dels més emprats es basa en l'extracció amb etanol/àcid fòrmic/acetoni-tril en tub, però en comparació amb la transferència directa, és un mètode més lent i laboriós, i es recomana per a la identificació de microorganismes amb cèl·lules rígides o parets complexes (llevats, *Mycobacterium* spp., etc.), o si es requereixen espectres d'alta qualitat per a la creació de llibreries. Amb aquesta metodologia, es parteix, com amb les dues anteriors, d'una colònia aïllada que se sotmet a una extracció en tub i amb passos de centrifugació. L'eluat resultant de la lisi cel·lular es diposita sobre el suport metàl·lic i es cobreix amb la matriu HCCA per iniciar l'anàlisi després del seu assecatge casi immediat a temperatura ambient.

Independentment del mètode de preparació de la mostra, el procediment d'anàlisi que segueixen és comú (**Figura 7**). Conjuntament amb les mostres es recomana introduir en una de les posicions de la placa d'anàlisi una solució de calibratge anomenada *bacterial test standard* (Bruker Daltonics) que consisteix en una preparació de proteïnes d'*E. coli* DH5 alfa que crea un perfil proteic per calibrar l'instrument i validar l'experiment.

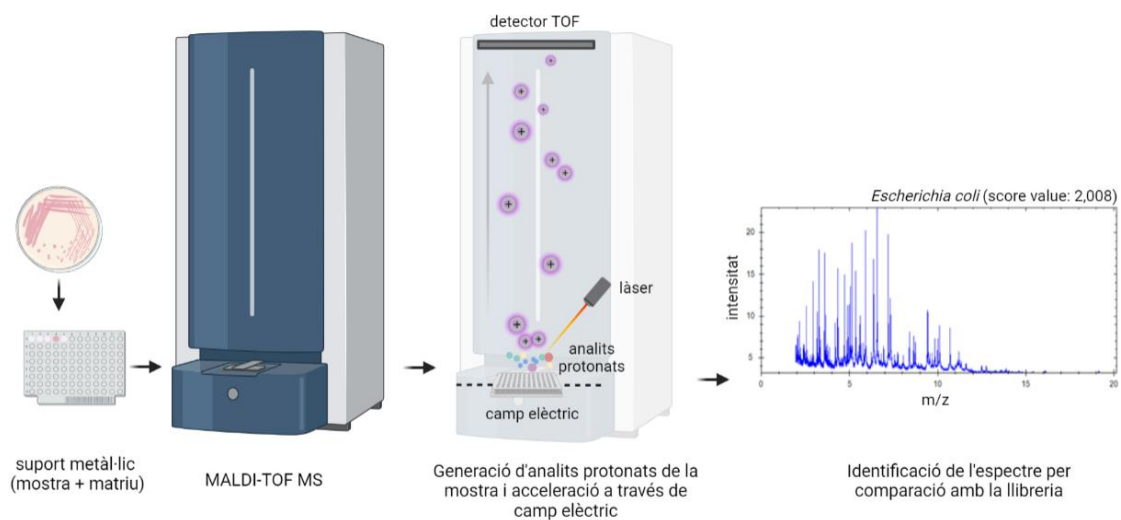


Figura 7. Esquema del flux de treball per a la identificació bacteriana mitjançant la tècnica MALDI-TOF MS

Una vegada el suport metàl·lic disposa de la mostra i la matriu ja cristal·litzada, s'introdueix dins l'equip de MALDI-TOF MS i es programa amb el software de l'equip la seva lectura i posterior identificació amb la llibreria. Primerament se sotmet la mostra al buit, necessari per l'anàlisi, i un cop s'assoleixen les condicions òptimes, la matriu cristal·litzada, que és capaç d'absorbir energia, s'irradia amb un làser i es generen analits protonats a la mostra. Se sotmet la mostra a camps elèctrics que permet separar els analits per la seva massa i càrrega. Els analits carregats es desplacen al llarg d'un tub al buit i són detectats per un analitzador de masses (detector TOF,

Time of flight). En funció del temps que triguen en recórrer el tub i arribar al detector TOF es determina la ràtio massa/càrrega (m/z) de la mostra donant lloc a un conjunt de pics de masses que generen un perfil proteic característic de la soca analitzada. La identificació de la soca es dona per comparació de similituds entre el perfil proteic obtingut de la mostra i els perfils proteics de soques de referència de la llibreria de l'equip, i és representada amb una puntuació (*score*) classificada amb tres colors diferents: identificació no fiable (vermell; $\text{score} \leq 1,699$), identificació probable del gènere (groc; $\text{score} 1,700\text{--}1,999$), identificació segura de gènere i espècie probable (verd; $\text{score} 2,000\text{--}2,299$) i la identificació molt probable d'espècie (verd; $\text{score} \geq 2,300\text{--}3,000$).

Segons el fabricant (Bruker Daltonics), una classificació addicional que acompanya a la puntuació establerta (*score*) per avaluar la consistència dels resultats en quant a l'assignació taxonòmica, consisteix en classificar els 10 millors resultats de coincidències d'identificació basant-se en 3 categories de consistència: A (coherència de l'espècie: les 10 millors coincidències en verd són de la mateixa espècie ó les 10 millors coincidències en groc són del mateix gènere), B (coherència del gènere: els 10 primers coincideixen en el mateix gènere però no en la mateixa espècie), o C (sense consistència: les 10 millors coincidències no són del mateix gènere). A la **figura 8** es mostra un breu d'informe de resultats a mode d'exemple.

Analyte Name	Analyte ID	Organism (best match)	Score Value	Organism (second best match)	Score Value
<u>A1</u> (+)(B)	CR_N8_1	Rheinheimera sp.	1.842	not reliable identification	1.302
<u>A11</u> (++)(B)	CR_N8_11	<u>Aeromonas caviae</u>	2.266	Aeromonas hydrophila	2.238
<u>D8</u> (+++)(A)	DEC_N8_14	Acidovorax temperans	2.331	Acidovorax temperans	2.21
<u>B5</u> (-)(C)	CR_N8_17	not reliable identification	1.554	not reliable identification	1.514

Figura 8. Exemple d'informe breu de resultats amb la consistència en negreta i parèntesi (primera columna), *score* (quarta columna) i identificació (tercera columna) per a 3 soques aïllades del riu (CR) i 1 soca de la decantació (DEC) a la vuitena campanya de mostreig (N8).

A la primera fila de resultats es pot observar com la soca número 1 aïllada del riu al vuitè mostreig (CR_N8_1), ha estat assignada a nivell de gènere amb una consistència B i un *score* de 1,842 que correspon a *Rheinheimera* sp. A la segona fila, la soca CR_N8_11 ha obtingut una identificació a nivell d'espècie amb un *score* de 2,266 però amb una consistència B atès que la identificació de la segona millor coincidència (second best match: *Aeromonas hydrophila*) no correspon a la mateixa espècie que la millor identificació (best match: *Aeromonas caviae*). La tercera fila corresponent a una soca aïllada de la decantació (DEC_N8_14), va assolir una

identificació a nivell d'espècie amb un *score* de 2,331 i consistència A identificada com a *Acidovorax temperans* segons les 10 primeres millors coincidències. Per últim, la quarta columna, corresponent a la soca CR_N8_17 no va ser possible la seva identificació a l'obtenir un *score* de 1,554 i una consistència C.

Les similituds entre els organismes i un nombre limitat d'espectres a la llibreria poden provocar una mala discriminació entre espècies, així com identificacions errònies. Aquests errors es produeixen amb una freqüència relativament mitja o baixa en funció del camp d'estudi i normalment es poden superar emprant un mètode d'extracció més exhaustiu, proves d'identificació addicionals o mitjançant la creació de llibreries. Per exemple, la tècnica MALDI-TOF MS actualment no pot diferenciar *E. coli* de *Shigella*, no obstant alguns taxònoms suggereixen que podrien tractar-se de la mateixa espècie i no de dues espècies diferents (Pupo et al., 2000). Tanmateix, continuant amb l'exemple de la **figura 8**, per la soca CR_N8_11 corresponent al gènere *Aeromonas* no s'ha pogut determinar l'espècie atès que aquest gènere conté espècies amb patrons molt similars i per tant la diferenciació entre espècies és difícil, com han descrit alguns autors als seus estudis (Párez-Sancho et al., 2018).

Com s'ha dit anteriorment, un altre motiu pel qual s'obté un espectre de bona qualitat però no és possible la seva identificació, és la manca o nombre limitat d'espectres de referència a la llibreria. En aquests casos, és possible obtenir una identificació incorrecta a nivell d'espècie o cap identificació.

Les llibreries que incorpora l'equip són actualitzades anualment pel fabricant incorporant nous perfils principals d'espectres (MSP: *Main Spectrum Profile*) de soques de referència per millorar el poder d'identificació. Atès el gran interès que ha despertat aquesta tecnologia a altres àrees de la microbiologia com l'alimentària i l'ambiental, poc a poc a s'han anat incorporant soques d'origen no clínic a les llibreries, amb el que les actualitzacions periòdiques podrien ajudar a superar aquest problema. No obstant, en funció del tipus de mostra, sobretot pel que fa a les aigües, és freqüent que la llibreria no disposi de MSPs suficients relacionats amb aquestes matrius i per tant podrien no ser adequats per una correcta identificació de les soques d'interès. És per aquest motiu que les llibreries estan obertes a modificacions i per tant es poden adaptar a les necessitats de cada laboratori mitjançant la creació de llibreries específiques d'interès. Molts autors han descrit la necessitat de crear llibreries personalitzades incorporant MSP de les soques d'interès per tal d'aconseguir una correcta identificació de les soques (De Carolis et al., 2014; Kopcakova et al., 2014; Santos et al., 2016; Seuylemezian et al., 2018).

En la present tesi doctoral s'ha creat una llibreria amb MSP de soques aïllades de mostres d'aigua i soques de referència ambientals per tal de millorar la capacitat d'identificació dels anàlisis amb MALDI-TOF MS a mostres d'aigua de diferents orígens.

7.2.2. Caracterització de comunitats bacterianes totals per seqüenciació massiva del gen 16S ARNr (metabarcoding)

En els últims anys, l'aparició de plataformes de seqüenciació de nova generació (NGS) d'alt rendiment (*High-Throughput Sequencing*) ha revolucionat l'estudi de comunitats microbianes. Les diverses aproximacions en funció de l'objectiu de l'estudi van des d'identificar tots els membres que conformen la comunitat bacteriana i caracteritzar la seva diversitat per seqüenciació d'amplicons del gen 16S ARNr (metabarcoding), discernir la capacitat funcional de la comunitat de microorganismes analitzant la totalitat dels genomes (metagenòmica), l'anàlisi de l'expressió gènica que estan manifestant aquests microorganismes (metatranscriptòmica), etc.

Els estudis de caracterització bacteriana en mostres ambientals mitjançant l'ús de plataformes NGS tenen un gran potencial per explorar el microbioma de diferents ambients, com el d'aigua dolça, ja que permeten observar la totalitat de la composició microbiana i no només aquella cultivable, que pot contenir patògens oportunistes i una altra diversitat poc explorada. Per tant, resulten de gran utilitat com enfocaments alternatius i complementaris al control d'indicadors de contaminació fecal legiscats per a una millor comprensió dels microbiomes als sistemes d'aigua per millorar la gestió de l'aigua.

Tanmateix, l'anàlisi de les comunitats bacterianes és metodològicament complexa i no està estandarditzada, amb el que la varietat de metodologies que es poden utilitzar són molt diverses. Cada etapa, des del mostreig fins a l'anàlisi de dades, pot introduir biaixos en la identificació o les abundàncies relatives que poden afectar la interpretació de la diversitat de la comunitat (Pollock et al., 2018).

La majoria dels estudis de la microbiota bacteriana es basen en *metabarcoding*, que consisteix en amplificar i seqüenciar múltiples fragments d'ADN a la vegada de diferents mostres, proporcionant dades tant qualitatives com semi-quantitatives (proporcions o abundància relativa) de les comunitats bacterianes.

La tècnica de *metabarcoding* neix de la unió de dos processos: el *DNA barcoding* (Hebert et al., 2003), que consisteix en la identificació taxonòmica d'individus mitjançant l'ús d'una seqüència d'ADN curta com a "codi de barres" (*barcodes*), i la seqüenciació massiva.

El *metabarcoding* (o seqüenciació massiva d'amplicons) permet la caracterització de les comunitats microbianes obtenint milions de seqüències en un mateix experiment i permetent la seqüenciació de múltiples mostres alhora mitjançant l'ADN extret d'una mostra ambiental.

El processament de les mostres per metabarcoding consta dels següents passos:

Mostreig i concentració de la mostra: El mostreig és una de les parts més importants per a la correcta interpretació dels resultats finals així com el processament de la mostra (punt de mostreig, temps i temperatura de conservació de la mostra, etc.). Les mostres d'aigua es poden concentrar segons la metodologia comentada a l'anterior apartat 7.1., i és rellevant atès que aquesta pot tenir influència en el tipus de microorganisme que es recuperarà. Per exemple, en la concentració per filtració hi ha diverses mides de porus que retindran uns microorganismes però d'altres no, i en funció del volum de mostra serà possible o no la detecció de taxons amb baixa abundància.

Extracció d'ADN: Existeixen diferents mètodes d'extracció que poden ser manuals o automatitzats, basats en kits comercials o en protocols tradicionals (ex: fenol-cloroform) que es diferencien pel seu tipus de lisi (mecànica, física, química, enzimàtica, etc.), el volum de processament de mostra (habitualment 100 µL-10 mL), etc. Durant l'extracció cal tenir en compte que algunes cèl·lules bacterianes poden tenir més resistència a la lisi com les endòspores o els bacteris gram positius (Riffiani et al., 2015) i per tant pot afectar a l'eficiència de l'extracció i a l'ADN que se n'obindrà de tota la mostra. Tanmateix està descrit que durant l'extracció alguns reactius poden romandre a l'extracte d'ADN donant lloc a inhibicions a la posterior PCR, així com la presència del propi microbioma dels kits d'extracció que poden aportar ADN addicional que no es trobava a la mostra original (Brandt and Albertsen, 2018; Salter et al., 2014).

Selecció de primers i amplificació per PCR: L'ADN extret de les mostres s'amplifica prèviament a la seqüenciació per obtenir els amplicons. Aquests estan dirigits a un fragment curt del gen 16S ARNr, que presenta 9 regions hipervariables (V1-V9), que s'utilitzen per inferir la taxonomia i que presenten una cobertura i resolució taxonòmiques diferents en funció dels tàxons (Baker et al., 2003; Bukin et al., 2019; Morgan and Huttenhower, 2012). És important destacar que l'elecció de les regions hipervariables i el disseny d'encebadors (primers) de PCR tenen un efecte en la resolució filogenètica. Tot i que no hi ha un consens respecte quina és la regió "universal",

les més comuns són les regions V2/V3, V3/V4 o V4, ja que contenen la màxima heterogeneïtat de nucleòtids i mostren el màxim poder discriminatori. Segons Bukin, la regió V2/V3, presenta una resolució més alta per a tàxons de rang inferior (gèneres i espècies), mentre que altres autors indiquen que la regió V4 representa millor el conjunt diversitat bacteriana (Caporaso et al., 2012; Liu et al., 2020).

Selecció de la plataforma de seqüenciació: Actualment al mercat hi ha disponibles diverses plataformes de segona generació (seqüenciació massiva per síntesi: Illumina, Ion Torrent, etc. o piroseqüenciació: 454 Roche) i tercera generació (seqüenciació per nanopor) que permeten l'anàlisi de metabarcoding (D'Amore et al., 2016; Heather and Chain, 2016). A grans trets es diferencien per la capacitat d'analitzar fragments llargs (fins a 100 Kb) o curts (fins a 400 pb), si es requereix amplificació prèvia o no, el tipus de reacció de seqüenciació, etc. La tercera generació d'Oxford Nanopore Technologies (MinION, GridION, etc.), que disposa de seqüenciadors portables a camp com MinION, permet analitzar fragments llargs en poques hores sense la necessitat d'amplificació, però actualment l'error de seqüenciació es troba al voltant d'un 14% (Sahlin et al., 2021). La tecnologia Illumina (principalment el sistema MiSeq) s'ha convertit en la plataforma de seqüenciació més comuna pels estudis de metabarcoding del gen 16S ARNr, ja que és d'alt rendiment, robust i amb una taxa d'error baixa (1%), però permet analitzar fragments d'un màxim de 300 pb.

A continuació, a la **figura 9**, es detalla el protocol de *metabarcoding* pel que fa a la **seqüenciació per Illumina MiSeq** (D'Amore et al., 2016).

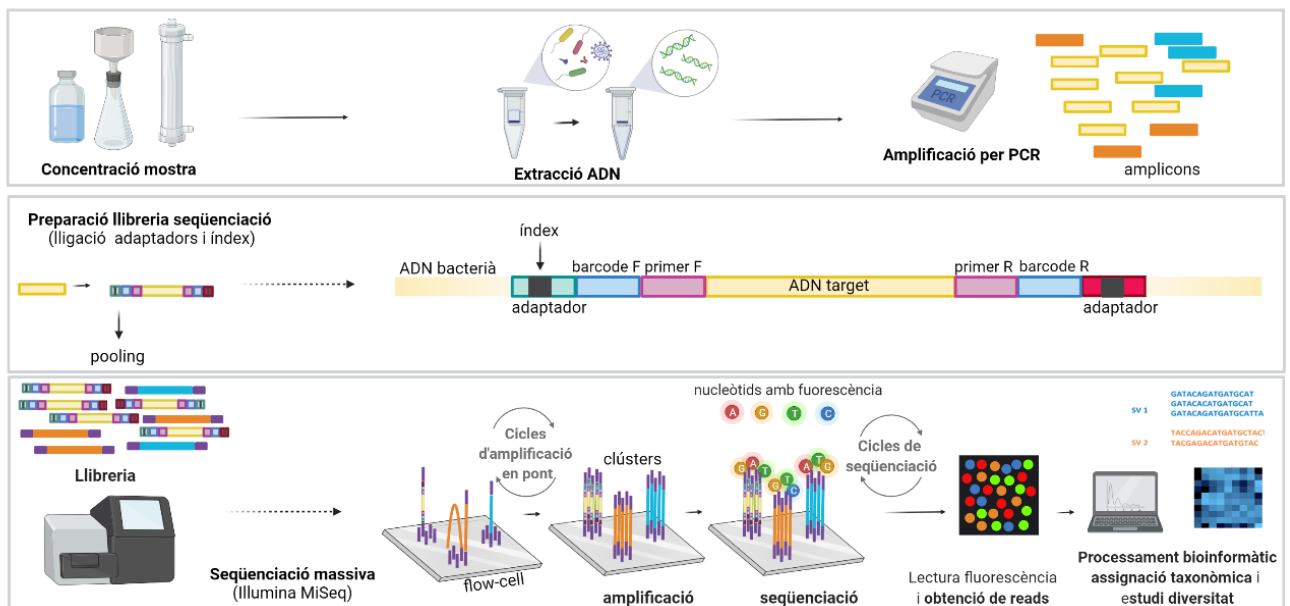


Figura 9. Processament de les mostres per seqüenciació massiva del gen 16S ARNr amb la plataforma Illumina MiSeq

Partint d'una prèvia concentració de la mostra i la posterior extracció de l'ADN contingut en aquestes, aquest ADN s'amplifica mitjançant una PCR dirigida a la regió genòmica d'interès obtenint els amplicons. Els encebadors (*primers*) usats en aquesta PCR contenen un codi de barres únic (*barcode*) que permetrà la posterior identificació de cada mostra i els seus corresponents amplicons, ja que per seqüenciar amplicons de diverses mostres al mateix temps, s'agruparan tots junts en una sola mostra al seqüenciador (*multiplexing* o *pooling*). Es recomana que per obtenir una bona cobertura de seqüències, el màxim nombre de mostres que es multiplexin siguin 96, on s'obtindran aproximadament 230.000 seqüències crues/mostra. Si s'agrupen més mostres, s'obtindran un menor número de seqüències per mostres rere la seqüenciació.

A continuació es prepara la **llibreria de seqüenciació** que consisteix en l'agrupament de totes les mostres (*pooling*), una purificació dels productes de la PCR i la lligació dels adaptadors de seqüenciació amb l'índex corresponent. Si a més a més es preparen diferents llibreries per un mateix experiment (*run*) de seqüenciació, l'índex de la llibreria ens permetrà diferenciar-les. La llibreria es quantifica per qPCR per tal de carregar la concentració idònia a la cel·la de flux (*flow-cell*), que és el suport físic on es produiran totes les reaccions durant la seqüenciació. La cel·la de flux conté oligonucleòtids fixats a la superfície que són complementaris a les seqüències del adaptadors que tenen una doble funció: capturar els adaptadors immobilitzant els amplicons, i mitjançant una polimerasa, serveixen de primers per l'amplificació en pont de la cadena complementària de l'amplicó. Durant els diferents **cicles d'amplificació** es produeixen milers de còpies clonals dels amplicons originals que generen clústers. Aquests clústers d'amplificació seran seqüenciats mitjançant la síntesi d'ADN amb nucleòtids marcats amb un fluoròfor que a l'incorporar-se emeten un color diferent en funció del tipus de nucleòtid que s'hi ha unit (A, C, G o T). La **seqüenciació per síntesi** incorpora 1 sol nucleòtid a cada cicle de seqüenciació. Després de cada incorporació, els clústers són monitoritzats mitjançant la lectura de la fluorescència corresponent a cada posició de la seqüència. Aquest procés es produeix amb milers de clústers emetent senyals a la vegada, generant per tant milions de dades de les seqüències alhora. A cada seqüència però la lectura seqüencial d'un únic nucleòtid per cicle permet una baixa taxa d'error (~1%). No obstant, atès que l'addició de nucleòtids no es fa alhora en totes les seqüències que es sintetitzen (primer se sintetitzen les cadenes *forward* i després les *reverse*), la qualitat del senyal que s'obté va disminuint al llarg dels cicles de seqüenciació per un deteriorament dels reactius ja que un experiment (*run*) de seqüenciació triga aproximadament 48h hores. Aquest fet condiciona un dels principals desavantatges d'aquesta

tecnologia i és que la impossibilita per treballar amb seqüències més llargues de 300 parells de bases (Tan et al., 2019).

Un cop finalitzada la seqüenciació, en un *run* típic de MiSEQ v3 (2x300 pb) s'obtenen al voltant de 20-22 milions de seqüències crues (*reads*). L'anàlisi de les dades obtingudes requereix l'ús d'eines bioinformàtiques (*pipelines*). Hi ha molts tipus de *pipelines* bioinformàtiques disponibles per processar i analitzar les dades de seqüenciació, com per exemple QIIME2, Bioconductor, Mothur, Phyloseq, etc. (Marizzoni et al., 2020). El **processament bioinformàtic** permet eliminar els adaptadors, primers, barcodes, reads de baixa qualitat o seqüències errònies (*denoising*), assignar cada *read* a la seva mostra corresponent (*demultiplexing*), agrupar totes les seqüències segons les seves similituds (*dereplicating*), alinear les seqüències per parells, eliminar seqüències quimèriques i agrupar les seqüències úniques finals en clústers (*clustering*) segons les variants de seqüència d'amplicons (ASV), o bé segons un percentatge de tall de similitud constant (*cut-off*).

El pas final és l'anàlisi dels ASVs, que ens permetran veure la seva abundància relativa i obtenir la **identificació taxonòmica**, que es realitza mitjançant la comparació amb seqüències de referència de bases de dades. Tot i que hi ha diverses bases de dades disponibles, les més emprades són SILVA, Greengenes i RDP que regularment s'actualitzen i es revisen (Sahlin et al., 2021).

Tot i la complexitat de l'ús de les eines bioinformàtiques per a l'estudi de la caracterització de les comunitats bacterianes, les eines van evolucionant per a facilitar a l'investigador el processament de les dades. Les tècniques de seqüenciació d'alt rendiment, com el metabarcoding 16S ARNr, tenen un gran potencial per a proporcionar informació detallada que complementi la qualitat microbiològica de l'aigua.

A la present tesi doctoral s'ha posat a punt la tècnica de metabarcoding 16S ARNr de la regió v4 mitjançant l'ús de la plataforma Illumina MiSeq per ajudant a generar una imatge més precisa de les comunitats microbianes poc explorades en les diferents etapes de l'ETAP i l'aigua potable.

7.3. El paper dels microorganismes als sistemes d'aigua potable

Les comunitats bacterianes al medi natural són omnipresents i tenen un paper important en la majoria de processos biològics, com en la descomposició de la matèria orgànica, formen part de diversos cicles biogeoquímics (fixació del nitrogen, carboni, etc.), són capaços de degradar contaminants, inhibeixen el creixement d'altres microorganismes, etc.

La presència d'aquestes comunitats als sistemes de tractament d'aigua potable no tenen perquè suposar un risc per la salut, de fet poden tenir un paper positiu facilitant l'eliminació de contaminants químics presents i reduint la matèria orgànica (Berry et al., 2006; Blázquez-Pallí et al., 2019; Bouwer and Crowe, 1988), però per contra també poden tenir un impacte negatiu donades certes condicions. Per exemple, s'ha observat en sistemes d'aigua amb desinfecció per monocloramines que alguns bacteris nitrificants pertanyents a *Alfaproteobacteria* i *Gammaproteobacteria* com *Nitrosomonas spp.* poden contribuir a l'esgotament de monocloramines residuals donant lloc a la formació de nitrats a l'aigua potable i en conseqüència un augment de HPC (Regan et al., 2002; Zhang and DiGiano, 2002).

D'altra banda, el recreixement bacterià pot donar lloc a l'aparició de bacteris oportunistes o potencials patògens que contribueixin al deteriorament de la infraestructura, alterin les propietats organolèptiques de l'aigua i en alguns casos podrien comprometre la seva seguretat (Berry et al., 2006; Li et al., 2017). Per exemple, es va observar durant un esdeveniment molt extrem a Beijing una coloració vermella de l'aigua potable causada principalment per *Gallionella spp.* (bacteri oxidant de ferro) (Li et al., 2010). Tanmateix altres autors han reportat episodis d'olor i gust de l'aigua potable causat per alguns *Streptomyces*, que poden produir grans quantitats de compostos orgànics volàtils, inclosos terpenoides (geosmina) (Asquith et al., 2018) o causats per alguns cianobacteris (Izaguirre et al., 1982).

Es creu que els **biofilms** són la font principal de microorganismes a la xarxa de distribució. Els biofilms són poblacions de microorganismes estructuralment complexes que es troben incrustades en una matriu de substàncies polimèriques extracel·lulars (Asquith et al., 2018). Aquesta estructura és dinàmica i sol presentar una colonització successiva de diferents taxons bacterians i proporciona una major protecció als desinfectants facilitant un reservori a l'aparició de possibles patògens oportunistes.

Els **patògens** transmesos per l'aigua són una de les majors preocupacions ja poden tenir una virulència moderada o gran i poden causar malalties especialment a persones amb immunodeficiències amb símptomes com gastroenteritis, infeccions respiratòries, conjuntivitis i erupció cutània. La contaminació fecal és una de les principals fonts d'aquests patògens com *E. coli*, *Salmonella*, *Cryptosporidium*, *Giardia* i norovirus. No obstant altres espècies poden actuar de patògens oportunistes com *Aeromonas*, *Arcobacter*, *Campylobacter*, etc. (Pruden et al., 2019). Algunes espècies que formen sovint biofilms als sistemes d'aigua potable i són patògens oportunistes són *Legionella pneumophila*, *Pseudomonas aeruginosa*, *Stenotrophomonas maltophilia* i *Mycobacterium spp.* (van der Wielen and van der Kooij, 2013). Segons diversos

estudis, s'ha documentat que els patògens oportunistes i potencials formadors de biofilms, així com els indicadors de contaminació fecal, augmenten a temperatures més altes (Calero Preciado et al., 2021; Favere et al., 2021; Hou et al., 2018). Per tant, s'espera que davant d'un escenari d'augment de temperatura global i el conseqüent augment de nutrients a les aigües indueixi un augment el potencial de recreixement d'aquestes comunitats.

Per tant, no es pot ignorar el paper dels bacteris en els sistemes d'aigua i el seu aïllament, identificació i caracterització són importants per entendre millor l'estat actual dels sistemes per tal de proveir noves eines de control de la qualitat de l'aigua així com per fer front al repte de la gestió de l'aigua davant el context d'estrès hídric i canvi climàtic.

OBJECTIUS

L'objectiu principal de la present tesi doctoral ha estat el control microbiològic de l'aigua mitjançant indicadors i l'aplicació de procediments basats en proteòmica i seqüenciació massiva (genòmica) per a la descripció de la diversitat bacteriana associada a les etapes de potabilització (des de la captació fins a la impulsó a la xarxa de distribució) de l'estació de tractament d'aigua potable (ETAP) de Sant Joan Despí.

Els objectius específics han estat:

1. Creació d'una **llibreria d'espectres de masses de soques ambientals** per a la identificació per espectrometria de masses MALDI-TOF MS de bacteris presents a les aigües de consum (Article 1).
2. Optimització dels mètodes de **concentració i extracció d'ADN** de mostres d'aigua per l'anàlisi de seqüenciació massiva (Article 2).
3. Control de qualitat microbiològic mitjançant l'anàlisi convencional d'**indicadors de la contaminació fecal** per a la determinació de la cinètica d'eliminació al llarg dels tractaments de l'ETAP (Article 2).
4. Caracterització i monitorització de les **poblacions bacterianes cultivables** de les plaques d'heteròtrofs mitjançant **MALDI-TOF MS** (Article 3).
5. Caracterització i monitorització de la **totalitat de les poblacions bacterianes** mitjançant **metabarcoding** (seqüenciació massiva) (Article 2).
6. **Comparació de la biodiversitat** segons les aproximacions per les tècniques de **MALDI-TOF MS i metabarcoding** i avaluació del seu ús conjunt addicional als paràmetres legiscats (Article 3).

INFORMES DE LES PUBLICACIONES

1. Llistat de publicacions

La present tesi Doctoral s'ha basat en les següents publicacions en revistes revisades per parells:

Article I:

Pinar-Méndez, A., Fernández, S., Baquero, D., Vilaró, C., Galofré, B., González, S., Rodrigo-Torres, L., Arahál, D.R., Macián, M.C., Ruvira, M.A., Aznar, R., Caudet-Segarra, L., Sala-Comorera, L., Lucena, F., Blanch, A.R., Garcia-Aljaro, C., 2021. Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database. *Water Res.* 203. <https://doi.org/10.1016/j.watres.2021.117543>

Article II:

Pinar-Méndez, A., Wangensteen, O.S., Præbel, K., Méndez, J., Galofré, B., Blanch, A.R., Garcia-Aljaro, C., 2022. Monitoring bacterial community dynamics in a drinking water treatment plant : an integrative approach using metabarcoding and microbial indicators in large water volumes. *Water*. 14 (9), 1435. <https://doi.org/10.3390/w14091435>

Article III:

Pinar-Méndez, A., Galofré, B., Blanch, A.R., Garcia-Aljaro, C. Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant. *Water Res.* *Under review* (sotmès a *Water Research* -WR69746- el 19/05/2022)



Dra. Cristina García Aljaro
Barcelona, maig 2022



Prof. Dr. Anicet Blanch i Gisbert

2. Informe sobre el Factor d'impacte

Els articles que constitueixen la memòria de la present Tesi Doctoral han estat publicats, o es troben en procés de revisió, a revistes científiques internacionals indexades al *Journal Citation Reports*, i rellevants per a la línia de recerca en la que ha treballat la doctoranda.

Article I:

“Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database” va ser publicat l'any 2021 a la revista *Water Research*. La revista es troba al primer decil (D1) i primer quartil (Q1) a totes les seves categories (Water Resources, Environmental Sciences i Environmental Engineering) i disposa d'un factor d'impacte el 2020 de 11,236.

Article II:

“Monitoring bacterial community dynamics in a drinking water treatment plant: an integrative approach using metabarcoding and microbial indicators in large water volumes” ha estat publicat l'any 2022 a la revista *Water*. La revista es troba al segon quartil (Q2) a totes les seves categories (Water Resources i Environmental Sciences) i disposa d'un factor d'impacte el 2020 de 3,103.

Article III:

“Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant” es troba actualment (maig 2022) en procés de revisió a la revista *Water Research*. La revista es troba al primer decil (D1) i primer quartil (Q1) a totes les seves categories i disposa d'un factor d'impacte el 2020 de 11,236.



Dra. Cristina García Aljaro
Barcelona, maig 2022



Prof. Dr. Anicet Blanch i Gisbert

3. Informe de participació en les publicacions

La doctoranda Anna Pinar Méndez ha participat en totes les publicacions que formen part de la seva Tesi Doctoral com es detalla a continuació. Cap dels coautors presents en aquestes publicacions ha fet servir dades i/o els articles per a l'elaboració de la seva Tesi Doctoral.

Article I:

Pinar-Méndez, A., Fernández, S., Baquero, D., Vilaró, C., Galofré, B., González, S., Rodrigo-Torres, L., Arahal, D.R., Macián, M.C., Ruvira, M.A., Aznar, R., Caudet-Segarra, L., Sala-Comorera, L., Lucena, F., Blanch, A.R., Garcia-Aljaro, C., 2021. Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database. *Water Res.* 203. doi: 10.1016/j.watres.2021.117543

L'estudi ha format part del projecte nacional RETOS Drinking Water Library (RTC-2015-4496-2) en el que la doctoranda ha participat activament en el processament de les mostres (concentració de mostres, enumeració i aïllament de soques, extracció, identificació mitjançant MALDI-TOF MS i preservació de soques), l'anàlisi dels resultats obtinguts i ha coliderat la redacció de l'article.

Article II:

Pinar-Méndez, A., Wangenstein, O.S., Præbel, K., Méndez, J., Galofré, B., Blanch, A.R., Garcia-Aljaro, C., 2022. Monitoring bacterial community dynamics in a drinking water treatment plant : an integrative approach using metabarcoding and microbial indicators in large water volumes. *Water*. 14 (9), 1435. doi.org/10.3390/w14091435

La doctoranda ha contribuït en el disseny experimental de l'estudi i ha realitzat la totalitat dels mostreigs, processament de les mostres (concentració, anàlisi dels paràmetres microbiològics de la qualitat de l'aigua, extracció d'àcids nucleics i preparació llibreries de seqüenciació), ha participat activament en el processament bioinformàtic i anàlisi estadístic de les dades, així com l'anàlisi dels resultats obtinguts i ha liderat la redacció de l'article.

Article III:

Pinar-Méndez, A., Galofré, B., Blanch, A.R., Garcia-Aljaro, C. Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant. *Water Res.* *Under review* (sotmès a *Water Research* -WR69746- el 19/05/2022)

La doctoranda ha contribuït en el disseny experimental de l'estudi i ha realitzat la totalitat dels mostreigs, processament de les mostres (concentració de mostres, enumeració i aïllament de soques, extracció, identificació mitjançant MALDI-TOF MS, processament d'espectres, preservació i seqüenciació Sanger de les soques), processament de les dades, així com l'anàlisi dels resultats obtinguts i ha liderat la redacció de l'article.



Dra. Cristina García Aljaro

Barcelona, maig 2022



Prof. Dr. Anicet Blanch i Gisbert

PUBLICACIONES

1. CAPITOL 1**Article 1: Identificació ràpida i millorada de bacteris d'aigua potable mitjançant la llibreria *Drinking Water Library*, una base de dades dedicada a l'espectrometria de masses MALDI-TOF**

“Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database”

Anna Pinar-Méndez, Sonia Fernández, David Baquero, Carles Vilaró, Belén Galofré, Susana González, Lidia Rodrigo-Torres, David R. Arahall, M. Carmen Macian, María A. Ruvira, Rosa Aznar, Laia Caudet-Segarra, Laura Sala-Comorera, Francisco Lucena, Anicet R. Blanch, Cristina Garcia-Aljaro.

Water Research (2021); 203, 117543, DOI: 10.1016/j.watres.2021.117543

L'aigua destinada al consum humà prové de diferents orígens: minerals naturals i de fonts (envasades), i de xarxa de subministrament (aixeta). La seva qualitat, a nivell microbiològic, es controla mitjançant la detecció i quantificació de bacteris indicadors de contaminació fecal i bacteris aerobis heterotròfics, d'acord amb les Directives Europees 2009/54/EC i 2020/2184, respectivament. Mentre que l'absència d'indicadors deriva en un ús apte pel consum, l'enumeració dels bacteris heterotròfics permet la lectura de l'eficàcia dels tractaments de potabilització (no aplicable a aigües envasades) i l'estat de deteriorament de l'aigua. La presència de bacteris heterotròfics a l'aigua potable no suposa directament un risc per a la salut humana, però aquestes comunitats podrien comprometre la qualitat de l'aigua, amb el que és de gran importància disposar d'eines d'identificació ràpides i fiables dels bacteris residents pel seu control rutinari, com permet l'espectrometria de masses MALDI-TOF. No obstant, una gran limitació en l'ús d'aquesta tecnologia resideix en la falta de espectres de soques d'origen ambiental a la base de dades de l'equip (Bruker BDAL), amb el que esdevé necessari la creació de biblioteques específiques amb espectres d'interès per tal d'obtenir identifications satisfactòries. D'altra banda, per tal de treballar amb aquestes soques de les que no tenim informació del risc biològic, cal adoptar mesures de precaució i bioseguretat en els mètodes de preparació de mostres que inactivin possibles patògens o les seves espores, i a la vegada permetin la correcta lisi cel·lular per la obtenció de les proteïnes.

L'objectiu d'aquest estudi es va centrar en la creació de la primera base de dades d'espectres de masses de soques ambientals per la identificació específica per MALDI-TOF MS de bacteris presents a les aigües de consum: la Drinking Water Library (DWL). La construcció de la base de dades es va definir seguint dos metodologies diferents: una basada en l'adquisició de soques de

referència d'origen ambiental i de relació amb l'aigua, i l'altra basada en l'aïllament de soques directament de mostres d'aigua.

Per la primera aproximació, 199 soques de referència de la CECT, que complien amb la descripció, es van seleccionar, processar amb un protocol d'extracció de proteïnes exhaustiu (àcid fòrmic/acetoneitril) i es va generar un perfil principal d'espectres (MSP) robust (18-24 espectres de qualitat) per cada un d'ells, que es van incorporar a la llibreria DWL.

Per la segona estratègia, un total de 209 mostres d'aigua de diferents orígens (xarxa de distribució, diferents processos de tractament d'aigua potable, minerals naturals i de font) es van processar, segons el seu origen, per filtració de membrana, inoculació en massa de diferents volums o ultrafiltració amb Rexeed i es van incubar en medi agar de recompte en placa per aigües ISO (Oxoid, UK) a 22°C durant 72h i R2A a 22°C durant 120h. Els 3.809 aïllaments de bacteris es van seleccionar i re-aïllar en medi ISO a 22°C per obtenir un cultiu fresc i pur per l'anàlisi. Per a cada aïllat, es va processar una colònia amb el mètode de transferència directa amb àcid fòrmic al 70% i es va analitzar per MALDI-TOF MS. El resultat d'identificació amb la llibreria original de l'equip Bruker BDAL només va permetre la identificació del 45% de totes les soques (1.721), mentre que el 55% no es van poder resoldre (2.088).

Les 2.088 soques no identificades esdevenen el grup d'interès per a la creació de la llibreria per tal d'incorporar els seus espectres en una base de dades definida específicament per a la seva aplicació en l'anàlisi de mostres d'aigua. Així, les soques es van re-processar pel mètode d'extracció amb àcid fòrmic/acetoneitril i els espectres resultants es van re-analitzar i processar per MALDI-TOF MS, i agrupar en clústers segons les seves similituds. Un total de 120 soques representants de cada clúster es van caracteritzar per seqüenciació parcial del gen 16S rRNA i GC-FAME, i es va generar un MSP per cada una de les soques, que es van incorporar a la llibreria DWL amb la seva identificació corresponent.

En total 319 MSP formen part de la DWL incorporant un total de 96 gèneres diferents, dels quals 44 gèneres i 189 espècies, no es trobaven a la llibreria BDAL. Aquesta nova configuració va permetre una millora significativa en la identificació de les soques que restaven sense identificar i que es van re-identificar mitjançant l'ús conjunt de BDAL i DWL, reduint d'un 55% a un 8% les soques sense identificar, i un augmentant d'un 29% a un 76% la identificació a nivell d'espècie.

Com a conclusió principal, l'ús de MALDI TOF MS amb la base de dades adequada, permet una identificació ràpida de les soques d'interès i gràcies a la seva robustesa, facilitat d'ús i baix cost es proposa com a eina de control en rutina per l'anàlisi de les aigües.



Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database

Anna Pinar-Méndez^{a,d,e,1,*}, Sonia Fernández^{b,1}, David Baquero^b, Carles Vilaró^a, Belén Galofré^a, Susana González^b, Lidia Rodrigo-Torres^c, David R. Arahal^c, M. Carmen Macián^c, María A. Ruvira^c, Rosa Aznar^c, Laia Caudet-Segarra^{d,e}, Laura Sala-Comorera^{d,e}, Francisco Lucena^{d,e}, Anicet R. Blanch^{d,e}, Cristina Garcia-Aljaro^{d,e,*}

^a Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain

^b Cetaqua, Water technology center, Cornellà de Llobregat, Spain

^c Departamento de Microbiología y Ecología and Colección Española de Cultivos Tipo (CECT), Universitat de València, Valencia, Spain

^d Department of Genetics, Microbiology and Statistics, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain

^e The Water Research Institute, University of Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Bacterial identification
MALDI-TOF MS
Drinking water
Distribution network
Bottled water
Spectra database

ABSTRACT

According to the European Directives (UE) 2020/2184 and 2009/54/EC, which establishes the sanitary criteria for water intended for human consumption in Europe, water suitable for human consumption must be free of the bacterial indicators *Escherichia coli*, *Clostridium perfringens* and *Enterococcus* spp. Drinking water is also monitored for heterotrophic bacteria, which are not a human health risk, but can serve as an index of bacteriological water quality. Therefore, a rapid, accurate, and cost-effective method for the identification of these colonies would improve our understanding of the culturable bacteria of drinking water and facilitate the task of water management by treatment facilities. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) is potentially such a method, although most of the currently available mass spectral libraries have been developed in a clinical setting and have limited environmental applicability. In this work, a MALDI-TOF MS drinking water library (DWL) was defined and developed by targeting bacteria present in water intended for human consumption. This database, made up of 319 different bacterial strains, can contribute to the routine microbiological control of either treated drinking water or mineral bottled water carried out by water treatment and distribution operators, offering a faster identification rate compared to a clinical sample-based library. The DWL, made up of 96 bacterial genera, 44 of which are not represented in the MALDI-TOF MS bacterial Bruker Daltonics (BDAL) database, was found to significantly improve the identification of bacteria present in drinking water.

1. Introduction

Water quality for human consumption is regulated by the European Directives (UE) 2020/2184 and 2009/54/EC, which cover tap water and natural mineral water, respectively. In both types of water, an absence of fecal bacterial indicators is an indicator of safety, and heterotrophic bacteria are enumerated to manage water quality (Bartram et al., 2003), serving as an indicator of the efficacy of water treatment processes, mainly disinfection. Though not directly a risk for human health,

heterotrophic communities can influence water quality, as they may include opportunistic pathogens and can form biofilms on surfaces. The complex ecosystems of drinking waters also contain viable but non-culturable bacteria, which represent a minor fraction of the overall drinking water communities and are still highly underexplored.

A wide range of techniques are available for bacterial identification: they can be culture-based, involving the use of chromogenic media or biochemical testing, or molecular, such as high-throughput sequencing of the 16S rRNA gene. Differing considerably in performance and

* Corresponding authors at: Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, General Batet 1-7, Barcelona 08028, Spain and Department of Genetics, Microbiology and Statistics, University of Barcelona, Avinguda Diagonal, 643, 08028 Barcelona, Spain.

E-mail addresses: apinar@aiguesdebarcelona.cat (A. Pinar-Méndez), cr Garcia@ub.edu (C. Garcia-Aljaro).

¹ Both the authors contributed equally to this work.

<https://doi.org/10.1016/j.watres.2021.117543>

Received 26 May 2021; Received in revised form 23 July 2021; Accepted 7 August 2021

Available online 12 August 2021

0043-1354/© 2021 Elsevier Ltd. All rights reserved.

requirements, culture-based methods can be time-consuming, whereas molecular methods are more expensive and need specific technical skills and expertise. Proteomics offers a well-balanced approach through the application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which analyses the constantly expressed high-abundance proteins (mainly ribosomal) of a microbial cell. This promising tool has equivalent identification power to Sanger 16S rRNA gene sequencing and provides faster results (Sala-Comorera et al., 2016; Sárvari et al., 2018; Timperio et al., 2017). Moreover, according to previous studies (Tan et al., 2012), the use of MALDI-TOF can produce annually net savings of 87.8% in reagent costs compared to traditional methods, or up to 57.1% savings when including technologist and maintenance. Even though initial investment in the equipment is high, it is offset after a short period, typically three years at a reasonable use. Identification can be performed by non-skilled personnel and the time to obtain results is shortened by one working day compared to biochemical methods (Tsuchida et al., 2020). Meanwhile, in genomic technologies, a high expertise in molecular techniques is needed and also the results take longer, it depends on sequencing type, within few days to two weeks. Despite the power of genomics, some discrepancies have been observed between MALDI-TOF analysis of cultivable bacteria and 16S rRNA high-throughput sequencing since they are targeting different populations: heterotrophic populations versus total bacteria (Comorera et al., 2020). In this case, although genomics is thought to be more powerful, a lower diversity index was observed in samples analyzed by 16S rRNA-sequencing compared to heterotrophic bacteria by MALDI-TOF MS, probably because the dominance of certain species hindered the identification of minor genera by high-throughput sequencing. Furthermore, other biases including primer bias and the lack of a standardized methodology are still a major shortcoming which difficult its use in routine analysis (Boers et al., 2019).

On the other hand, current water quality regulations still rely on culture-based methods to assess water safety. In addition to the mandatory monitoring of fecal indicator bacteria and heterotrophic plate counts, a more in-depth characterization of heterotrophic bacteria by MALDI-TOF MS could provide a better understanding of these bacterial communities and improve water management and distribution.

Studies have demonstrated the reliability, speed, and easy-to-use features of MALDI-TOF MS in different areas, including food (Angelakis et al., 2011; Pavlovic et al., 2013; Yu et al., 2019), groundwater (Santos et al., 2017), wastewater (Eddabra et al., 2012), mining (Avanzi et al., 2017) or even in space craft and associated surfaces (Seuylemezian et al., 2018). Recent research (Sala-Comorera et al., 2017) showed the suitability of MALDI-TOF MS for the routine monitoring of heterotrophic bacteria in a drinking water treatment plant, as well as in mineral water (Sala-Comorera et al., 2020). However, there is general agreement that the resolving power of MALDI-TOF MS is currently limited by the lack of environmental spectra in commercial databases and there is a need to create in-house libraries with the bacterial spectra of interest (De Carolis et al., 2014; Kopcakova et al., 2014; Rahi et al., 2016; Santos et al., 2016; Seuylemezian et al., 2018). Another drawback of this approach is that identification can be affected by differences in analytical variables, i.e., the extraction method, culture conditions, matrix and database (Rahi et al., 2016; Ruelle et al., 2004).

Among the different MALDI sample preparation methods (direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction), the formic acid extended direct transfer has been described as the most cost-effective and time-saving (Ghosh et al., 2015; Theel et al., 2012). However, other authors have achieved better results when using extraction methods in fungal cells (Chalupová et al., 2014) and Gram-positive bacteria (Alatoom et al., 2011), as the complex and rigid cell walls need extra lysis to access their proteins. In addition, manufacturer-recommended protocols may also vary depend on their ability to inactivate pathogenic organisms. Safety measures with respect to handling MALDI-TOF MS samples, requires at least work under a biosafety level 2 conditions (BSL-2) and extraction procedures reducing

risk of exposure to potential pathogens. According to some studies, certain chemical or physical treatment used in sample preparation before processing on MALDI-TOF MS, contribute to biological inactivation of samples: e.g. trifluoroacetic acid (TFA) inactivates potential pathogens especially bacterial endospores (Lasch et al., 2008; Drevinec et al., 2012), tube-based ethanol-formic acid-acetonitrile extraction followed by filtration is recommended for security-sensitive biological agents (Tracz et al., 2016) or 70% ethanol is suggested to be sufficient for non-spore forming bacteria inactivation (Cunningham and Patel, 2015). The choice of an appropriate extraction method is therefore crucial for good results and biological safety.

We here report, to our knowledge, the first MALDI-TOF MS database constructed specifically for the identification of bacteria present in water for human consumption, named the Drinking Water Library (DWL). It was created using spectra from a selection of related environmental reference strains from the Spanish Type Culture Collection (CECT), as well as drinking water isolates, properly identified by 16S rRNA gene sequencing. The spectra were generated by standardized procedures and compiled in a dedicated database. The DWL will assist the monitoring of culturable bacterial communities in drinking water by allowing a rapid and accurate identification of isolates. In addition, this study provides the basis to increase the DWL by further characterization of drinking water heterotrophic bacteria.

2. Materials and methods

2.1. Samples and sampling conditions

Two hundred and nine samples were taken of different types of water, including water from distribution networks (109), process water (68), spring water (2) and bottled natural mineral water (30). The samples were grouped into two categories based on their origin: treated (from distribution networks and process water) and non-treated (natural mineral water and spring water). Sampling and processing conditions were slightly different for each group, as described below.

For the treated water category, 109 chlorinated drinking water samples (0.2–1 ppm free chlorine) from distribution networks in Catalonia (North-East Spain), Andorra and South France were analyzed; 68 process water samples were collected at nine different stages of treatment in a drinking water treatment plant (DWTP) in Barcelona corresponding to river water, groundwater, sand filters, ozonation, granular activated carbon (GAC) filters, ultrafiltration, cartridge filters, reverse osmosis and chlorinated water. For all 177 samples, 1 liter of water was collected in polyethylene sterile bottles with sodium thiosulfate (24 mg/L) and transported to the laboratory at 4°C for further analysis within 24 h.

For the non-treated water category, 12 different bottled water brands (natural mineral water and spring water) were selected and a total of 30 samples were analyzed. Natural mineral water samples consisted of water bottled in 1.5 L polyethylene terephthalate bottles acquired from different retailers and stored in the dark at room temperature (22 ± 2 °C) until analysis within one month after bottling. In the case of spring water, two different sources were sampled in sterile 1 liter polyethylene bottles, kept at 4 °C and analyzed within 24 h.

2.2. Isolation of heterotrophic bacteria from treated and non-treated drinking water

Heterotrophic bacteria from drinking water samples were recovered by concentration through membrane filtration, hollow fiber membrane ultrafiltration, or mass inoculation.

2.2.1. Treated water

Out of 177 treated water samples, 112 were processed by membrane filtration. Volumes of 100 mL were filtered through 0.22 or 0.45 µm pore-sized mixed cellulose ester filters (Millipore, Germany). Filters

were incubated on Water Plate Count Agar ISO (WPCA) (Oxoid, Spain) at 22 ± 2 °C for 72 ± 3 h.

For the process water samples, as concentration methods for early-stage samples from the DWTP were not required, decimal serial dilutions between 0.01 mL and 1 mL were cultured by mass inoculation in WPCA at 22 ± 2 °C for 72 ± 3 h (ISO, 1999). Spread plating of the samples onto WPCA by a mass inoculation method was also done.

To recover a higher diversity in chlorinated matrices, larger volumes were sampled. Thus, 25 samples from distribution networks were filtered through the hollow fiber filter Rxeed™ 25-A (Asahi Kasei Medical Co, Japan), which is a haemodialysis ultrafilter typically used for clinical purposes but recently applied to concentrate water samples. The main advantage of this method is its capacity to concentrate bacteria, viruses and protozoa in large volumes of water (from 10 to 1000 L) (Hill et al., 2007; Gunnarsdottir et al., 2020). Briefly, samples ranging between 50 and 700 L were processed by this technique. After filtration and elution, 0.2 mL of the sample was inoculated onto WPCA by the spread plate technique and incubated at 22 ± 2 °C for 72 ± 3 h.

Those samples that showed no growth on WPCA were re-analyzed using R2A plates (Becton Dickinson, U.S.) at 22 °C for 120 ± 4 h. R2A agar is a low-nutrient medium reported to improve the recovery of stressed and chlorine-tolerant bacteria from drinking water systems (Reasoner and Geldreich, 1985). Colonies were picked up and subcultured on the corresponding culture media and they were further characterized by MALDI-TOF MS analysis.

2.2.2. Non-treated water

Samples from spring water and bottled natural mineral water were concentrated by membrane filtration. Different sample volumes (0.1 mL, 1 mL, 10 mL, 100 mL) were filtered through 0.22 µm pore size nitrocellulose membrane filters (Millipore, Germany) and incubated on WPCA (Oxoid, Spain) supplemented with 0.5 g/L 2,3,5-triphenyltetrazolium chloride (TTC) at 22 °C for 120 ± 3 h. Due to bacterial metabolism, colorless TTC is reduced to formazan, a red compound helping the visualization of the colonies.

Colonies were subcultured in WPCA for purification prior to MALDI-TOF MS analysis.

2.3. Identification of bacterial isolates by MALDI-TOF MS using the Bruker library

Bacterial isolates were obtained by subculturing one single well-isolated colony on new WPCA in order to achieve fresh pure cultures. Sample preparation for MALDI-TOF MS was carried out using the formic acid extended direct transfer method recommended by Bruker Daltonics. Thus, biological material (a single colony) from fresh agar cultures was smeared directly onto a spot on a MALDI target plate and then overlaid with 1 µL of 70% formic acid. After air-drying, 1 µL of matrix solution (saturated solution of α -cyano-4-hydroxycinnamic acid in a standard solvent (Sigma-Aldrich)) was added. The Bruker Bacterial Test Standard (Bruker Daltonics, Germany) containing an extract of *Escherichia coli* DH5 alpha peptide with a protein profile was included for each plate to calibrate the instrument and validate the run.

Mass spectra ranging from 2000 to 20,000 Da of each isolate were automatically acquired, using the Microflex LT MALDI-TOF MS device and MALDI BioTyper software, version 3.1 (Bruker Daltonics, Germany) with FlexControl software package (Bruker Daltonics, Germany). The obtained spectra were analyzed using MALDI BioTyper Real-Time Classification software package (Bruker Daltonics, Germany) with the bacterial Bruker Daltonics (BDAL) database MBT Compass Library DB-7311, and checked using the FlexAnalysis software (Bruker Daltonics, Germany). The Bruker identification results are classified with a score based on matches between the mass spectra of the analyzed sample and the mass spectra database. The identification categories were as recommended by Bruker: unreliable identification was in red (score ≤ 1.699); probable genus identification in yellow (score 1.700–1.999); secure

genus and probable species identification in green (score 2.000–2.299); and highly probable species identification also in green (score ≥ 2.300 –3.000).

2.4. Selection of isolates to be included in the Drinking Water Library

The DWL consisted of spectra obtained from strains isolated in different water matrices and the spectra of reference strains from the CECT. The selection procedure was as follows:

2.4.1. Selection of non-successfully identified isolates (red category) recovered from the different water matrices under study

Non-successfully identified isolates were grouped into similarity clusters, subcultured and reanalyzed using the acid/acetonitrile extraction method (Bruker Daltonics, Germany): ca. 10 mg biomass from agar cultures was first suspended in 300 mL water by careful mixing, and then the suspension was mixed with 900 mL ethanol. The biomass was collected by centrifugation and the pellet was re-suspended in 1–80 µL 70% formic acid. The suspension was mixed carefully with 1–80 µL acetonitrile. Immediately after centrifugation, the supernatant was removed and aliquots of 1 µL were placed on each spot of a stainless-steel target plate. After air-drying, 1 µL of matrix solution was added.

The isolates were analyzed, generating 9 spectra per isolate, which constitute the mini Main Spectrum Profile (mMSP), a useful approximation to simplify the clustering. Spectra were acquired, processed, and compared by MALDI BioTyper software, version 3.1 (Bruker Daltonics, Germany), and checked for quality using the FlexAnalysis software (Bruker Daltonics, Germany). All spectra were processed using the default program settings for smoothing (Savitzky-Golay), baseline subtraction (TopHat) and normalization. An averaged mass spectrum was created by eliminating those with the higher deviation. The parameters used were a mass range from 3000 to 10,000 Da and 500 ppm as a maximum error for the main spectrum profile (MSP). The mMSPs were incorporated into an *ad hoc* database for clustering purposes.

Clustering of non-identified bacterial isolates was carried out in two steps. First, a dendrogram was created by the standard MALDI BioTyper 3.1 MSP creation method (Bruker Daltonics, Germany), where distance values are relative and normalized to a maximum value of 1.000. Second, a comparison based on log score values where each mMSP was compared with the *ad hoc* mMSP database. On this basis, isolates representing each cluster were selected for identification by 16S rRNA gene sequencing analysis.

2.4.2. Selection of CECT reference bacterial strains

For the construction of the DWL database, reference environmental and water-related bacterial strains were selected from the CECT catalogue (Table S1). The selection was based on the source of isolation (environmental and water-related), growth temperature (20–30 °C) and growth culture medium (mostly WPCA and R2A) or other media such as GSP (Glutamate Starch Phenol-red agar, also known as *Pseudomonas Aeromonas* Selective Agar), Glycine Vancomycin Polymyxin Cycloheximide Agar, Xylose Lysine Deoxycholate agar and Brain Heart Infusion agar. Related species (same genus) to those selected were also included, even though they had not been isolated from the environment. All the selected reference strains were cultured according to the growth conditions described for each bacterial strain in the CECT catalogue and processed with the acid/acetonitrile extraction method recommended by Bruker Daltonics to obtain their MSP as described below in Section 2.6.

2.5. Identification of isolates selected as representative for DWL spectra

2.5.1. Sequencing of the 16S rRNA gene

Representative bacterial strains from each cluster were identified by partial 16S rRNA gene sequencing on a Genius thermocycler (Techn, Burlington, NJ, USA). The amplification mixture (50 µL) comprised 1 µL

(50 pmol/μL) each of universal primers amplifying a 1000-bp region of the 16S rRNA gene 616 V (forward): 5'-AGAGTTTGATYMTGGCTCAG-3' and P699R (reverse): 5'-GGGTYKCGCTCGTTR-3' (Integrated DNA Technologies), 0.25 μL (5 U/μL) of Taq DNA polymerase (Takara, Clontech Laboratories, Inc.), 5 μL of 10X reaction buffer, 4 μL of dNTP mixture (10 mM), 33.75 μL of sterile filtered water (Milli-Q purification system, Millipore, Billerica, MA, USA), and 5 μL of DNA template (40 ng/μL). The DNA templates were amplified by initial denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 56°C for 30 s, extension at 72°C for 45 s, and a final extension at 72°C for 10 min. Negative controls, devoid of DNA, were simultaneously included in the amplification process.

PCR amplicons of the partial 16S rRNA gene were verified by visualization in 1% agarose gel electrophoresis. Amplicons were purified and sequenced by the MacroGen Company Inc. Madrid, Spain. Subsequent sequencing reactions were done on an Abi Prism 3700 automated sequencer using the Big Dye Terminator v3.1 cycle sequencing kit. Sequencing primers were the same ones as used in the amplification reaction but diluted five-fold (10 pmol/μL). The taxonomic classification of bacterial isolates was performed using EzTaxon (Kim et al., 2012; Yoon et al., 2017) and BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) (Altschul et al., 1997). Identification criteria for species delimitation varied depending on the phyla, ranging from 98.2% for strains of the phylum *Bacteroidetes* to 99.0% for *Actinobacteria* and those sequences with less than the respective cut-off values were considered potential new species, as in Meier Kolthoff et al. (2013). All strains are currently deposited at the public catalogue of the CECT and their identification results are shown in Table S2.

2.5.2. GC-FAME analysis

Complementary to the 16S rRNA gene sequencing, GC-FAME (Gas Chromatographic Fatty Acid Methyl Esters) analysis of bacterial strains was carried out as a method of strain authentication. Fatty acid methyl esters were extracted from biomass grown for 48 h on WPCA (Oxoid, Spain) at 26°C and prepared according to standard protocols as described for the MIDI Microbial Identification System (Sasser, 1990) at the CECT. Cellular fatty acid content was analyzed by GC with an Agilent 6850 chromatographic unit, with the MIDI Microbial Identification System using the RTSBA6 method (MIDI Inc, Newark, US) and identified using the Microbial Identification Sherlock software package. Only for those strains that did not show growth under these conditions was the analysis performed at their optimal growth conditions.

2.6. Generation of reference MSP for inclusion in the DWL

Selected isolates were inoculated in WPCA at 22 ± 2 °C and fresh cultures were used for MSP creation. To generate MSPs for inclusion in the DWL, a total of eight independent replicates of each isolate were processed, obtaining a collection of 24 spectra, three per replicate. Sample preparation for MALDI-TOF MS was carried out using the acid/acetonitrile extraction method recommended by Bruker Daltonics, as described before in section 4.b. The spectra measurements were recorded using the UltrafleXtreme (Bruker Daltonics, Germany) at VISAVET (UCM, Madrid, Spain). The instruments are equipped with a nitrogen laser. All spectra were recorded in linear positive mode with an acceleration voltage of 20 kV. Spectra were acquired by accumulating 250 laser shots across a spot. A mass range of 2000–20,000 *m/z* was used for analysis. The mass spectrometer was externally calibrated with the Bruker Bacterial Test Standard (Bruker Daltonics, Germany). Flex-Analysis software (Bruker Daltonics, Germany) was used for baseline subtraction, smoothing of spectra and mass labelling of peaks. MSPs in the DWL were represented by 18–24 good quality spectra.

2.7. Validation of the DWL

To evaluate the DWL performance, five internal and four external

validation exercises were carried out by different technicians and different MALDI-TOF spectrometers. Validation exercises consisted of blind studies for the identification of bacterial isolates whose MSP had been included in the DWL. For internal validation, selected strains were analyzed to evaluate the identification results based on different protein extraction methods: direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction. Corresponding spectra were obtained with the online acquisition method Biotyper Real Time Classification and the Microflex mass spectrometer (Bruker Daltonics, Germany). For identification, DWL and BDAL libraries were used. In addition, six other research centers in Spain performed the external validation. Samples consisted of 1.5 mL microcentrifuge tubes of bacterial pellet resuspended in ethanol:water 1:3 and stored at –20°C until analysis. In this trial, 30 strains were processed in parallel for MALDI-TOF MS analysis using an extraction method and different mass spectrometer devices: Ultraflex III (Bruker Daltonics, Germany), UltrafleXtreme (Bruker Daltonics, Germany) and Microflex (Bruker Daltonics, Germany). All spectra were obtained with the online acquisition method Biotyper Real-Time Classification (Bruker Daltonics, Germany). Matrix blaster function was previously performed by firing eight initial laser shots at 25% intensity to get rid of the very first layer on sample. Then, mass spectra were automatically acquired in steps of 50 shots for a total of 250 shots accumulated, with initial laser power at 15% and maximal at 25%. The results based on log score value and identification were compared with the DWL and BDAL libraries.

3. Results and discussion

3.1. Identification of isolates using the Bruker Daltonics (BDAL) database

A total of 3809 bacterial isolates were analyzed by MALDI-TOF MS using the BDAL database: 1085 from distribution networks, 442 from process water, 356 from spring water and 1926 from bottled natural mineral water.

Regarding the treated water samples, 288 isolates from distribution networks were identified up to species level (26.5%), 266 at the genus level (24.5%) and 531 were not reliably identified (49%). In process water samples, 143 out of 442 isolates were identified at the species level (32.4%), 115 at the genus level (26%) and 184 were unidentified (41.6%) (Fig. 1A).

In the non-treated water samples, 182 out of 356 isolates (51.1%) from spring water were identified at the species level, 72 (20.2%) at the genus level and 102 (28.7%) remained unidentified. In bottled natural mineral water, 490 out of 1926 isolates (24.5%) were identified at the species level, 165 (8.5%) at the genus level, whereas 1271 isolates (66%) were unidentified, representing the highest number of unknown strains among the different samples (Fig. 1A).

In summary, 1721 isolates (45%) were successfully identified with the BDAL database at the species and/or genus level, whereas 2088 isolates remained unidentified (55%).

3.2. Identification of isolates not matched to the BDAL database

A total of 134 cluster-representative isolates were selected for identification by partial 16S rRNA gene sequencing. However, after preservation, only 120 grew satisfactorily and could be sequenced and further characterized by FAME profiles (data not shown), which established affiliation to 53 genera (Table S2). The isolates identified by 16S rRNA gene sequencing corresponded to distribution networks (45), process water (28), bottled natural mineral water (35) and spring water (12) (Fig. 2B). Therefore, the water from distribution networks accounted for the highest number of species not represented in the BDAL database.

Fourteen out of the 53 identified genera were absent in the BDAL database. In addition, some isolates (67) belonged to species not included in the BDAL database or to yet-to-be described. Identification of non-matched isolates by 16S rRNA sequencing allowed the BDAL

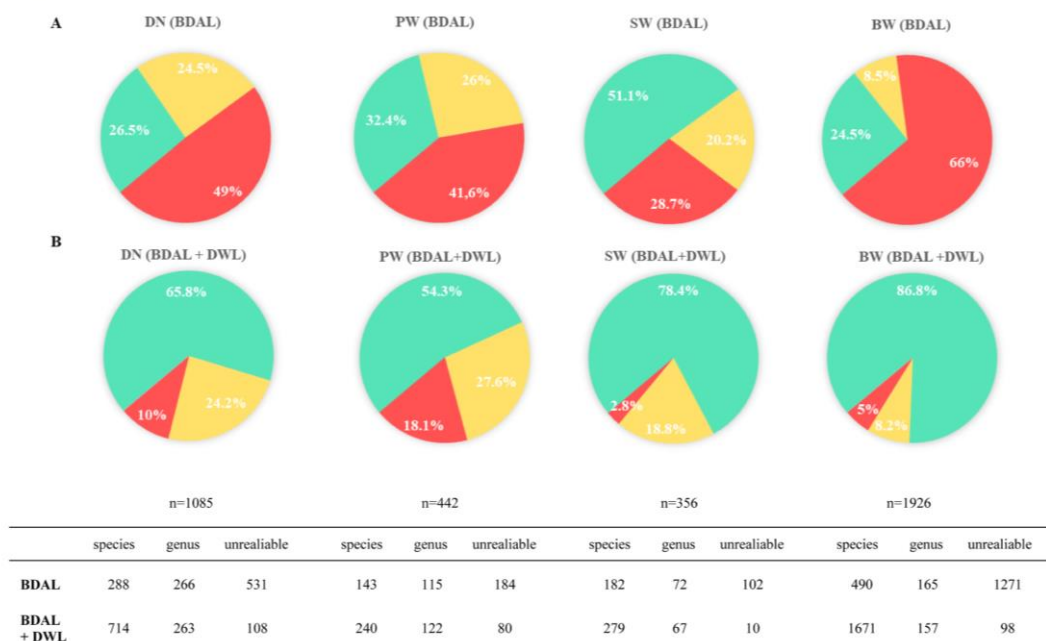


Fig. 1. Percentage of isolates identified using the original Bruker Daltonics (BDAL) database (A) and the BDAL database expanded with the Drinking Water Library (DWL) (B) from different sources of water: distribution networks (DN), process water (PW), spring water (SW) and bottled natural mineral water (BW). Number of isolates (n) per source and according to identification results are shown in the table below the graph. When using the extended database, better identification results were obtained at the level of species (green) and genus (yellow), and unreliable identification (red) was reduced.

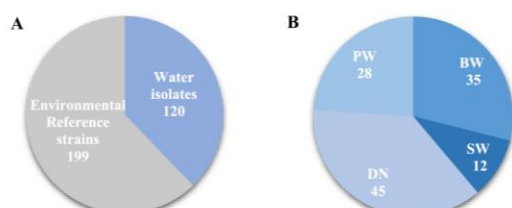


Fig. 2. Contents of the Drinking Water Library (DWL) and origin of the water isolates. (A) Composition of the 319 MSPs of the DWL according to the number of bacterial strains included in the database from environmental reference strains and water isolates. (B) Number of bacterial strains included in the new database according to origin: process water (PW), distribution network (DN), bottled mineral water (BW) and spring water (SW).

database to be expanded after the incorporation of their MSPs to the DWL (Table 1). A high variability in origin was observed among these isolates; some of the genera were recovered from multiple sources and others from only one. These results might be affected by a bias in the colony selection, as the main aim of the study was to recover the highest bacterial diversity from the four water sources. Nineteen genera were recovered from both treated (distribution network and process water) and non-treated water (spring water and bottled mineral water): *Acidovorax*, *Bacillus*, *Bosea*, *Brevundimonas*, *Caulobacter*, *Chryseobacterium*, *Deinococcus*, *Dyadobacter*, *Flavobacterium*, *Hydrogenophaga*, *Methylbacterium*, *Microbacterium*, *Mycobacterium*, *Paenibacillus*, *Pedobacter*, *Pseudomonas*, *Sphingobium*, *Sphingomonas* and *Sphingopyxis*. Nine genera

were isolated only from distribution networks: *Acinetobacter*, *Blastomonas*, *Domibacillus*, *Ensifer*, *Micrococcus*, *Novosphingobium*, *Porphyrobacter*, *Pseudoxanthomonas* and *Shinella*. Eight genera were recovered only from process water: *Cloacibacterium*, *Ideonella*, *Massilia*, *Nocardioideis*, *Novispirillum*, *Rheinheimera*, *Rhodoferax* and *Roseomonas*.

Similarly, for non-treated water samples, a total of six genera were exclusively recovered in spring water: *Aeromicrobium*, *Hermiimonas*, *Nocardia*, *Psychrobacillus*, *Williamsia* and *Xanthomonas*. And finally, regarding bottled natural mineral water samples, 11 different genera were identified as *Aquabacterium*, *Bradyrhizobium*, *Janibacter*, *Limnobacter*, *Methylbium*, *Pararhizobium*, *Phyllobacterium*, *Polaromonas*, *Psychrobacter*, *Rhizobium* and *Variovorax*.

3.3. The Drinking Water Library

The DWL constructed in this study (Fig. 2) includes 319 MSPs generated for isolates unmatched to the BDAL database and identified by 16S rRNA gene sequencing (120), representing the different clusters in which the isolates were grouped (Table S2). Also included are spectra from reference strains (199) selected for their environmental or water-related origins (Table S1).

The MALDI-TOF MS identification capacity increased when using the DWL with the newly incorporated MSPs, as half the genera in the library were absent from the BDAL database. At the beginning of the study in 2016, the BDAL database included 428 bacterial genera, which allowed the identification of 35% of drinking water isolates. The MSPs included in the DWL correspond to species belonging to 96 bacterial genera (Table 1), 44 of which were absent from the BDAL database and one was potentially a new genus. For the remaining 52 genera, 74 MSPs corresponding to one or more species were included, reaching a total of 319

Table 1

Contents of libraries BDAL 7311, DWL, and BDAL plus DWL according to the number of species per genera, and the origin of the main spectrum profiles (MSPs) that configured the new in-house DWL database (BW: Bottled mineral water, SW: Spring water, DN: Distribution network, PW: Process water and Ref: reference strains from the CECT catalogue). In brackets, number of new MSPs added. Genera not included in the BDAL database are shown in bold.

Genera included in the database	BDAL		DWL		BDAL+ DWL No. MSP	Origin of MSP (DWL)				
	No. MSP	No. species	No. MSP	No. species		BW	SW	DN	PW	Ref
<i>Acidovorax</i>	9	7	+ (2)	1	11	X		X		
<i>Acinetobacter</i>	120	22	+ (3)	2	123			X		X
<i>Aeromicrobium</i>	1	1	+ (1)	1	2		X			
<i>Aeromonas</i>	46	18	+ (38)	9	84					X
<i>Alcanivorax</i>	–	–	+ (1)	1	1					X
<i>Ampullimonas</i>	–	–	+ (2)	1	2					X
<i>Ancylobacter</i>	–	–	+ (1)	1	1					X
<i>Aquabacterium</i>	–	–	+ (3)	3	3	X				
<i>Arcicella</i>	–	–	+ (1)	1	1					X
<i>Arcobacter</i>	15	6	+ (2)	2	17					X
<i>Azoarcus</i>	3	3	+ (1)	1	4					X
<i>Azonexus</i>	–	–	+ (1)	1	1					X
<i>Bacillus</i>	145	102	+ (4)	1	149			X	X	
<i>Belliella</i>	–	–	+ (1)	1	1					X
<i>Blastomonas</i>	2	2	+ (5)	4	7			X		X
<i>Bosea</i>	–	–	+ (5)	5	5	X		X	X	
<i>Bradyrhizobium</i>	2	2	+ (1)	1	3	X				
<i>Brevundimonas</i>	15	8	+ (9)	5	24	X		X		X
<i>Caulobacter</i>	2	2	+ (5)	4	7	X		X		X
<i>Chitinimonas</i>	–	–	+ (3)	2	3					X
<i>Chromobacterium</i>	–	–	+ (1)	1	1					X
<i>Chryseobacterium</i>	29	15	+ (22)	16	51	X		X	X	X
<i>Cloacibacterium</i>	–	–	+ (1)	1	1				X	
<i>Cronobacter</i>	9	1	+ (1)	1	10					X
<i>Deinococcus</i>	4	2	+ (2)	2	6			X	X	
<i>Domibacillus</i>	–	–	+ (1)	1	1			X		
<i>Duganella</i>	–	–	+ (1)	1	1					X
<i>Dyadobacter</i>	–	–	+ (2)	2	2			X	X	
<i>Enticicia</i>	–	–	+ (2)	2	2					X
<i>Ensifer</i>	–	–	+ (1)	1	1			X		
<i>Ferruginibacter</i>	–	–	+ (1)	1	1					X
<i>Flavobacterium</i>	15	14	+ (18)	12	33		X		X	X
<i>Flectobacillus</i>	–	–	+ (2)	2	2					X
<i>Helimonas</i>	–	–	+ (2)	1	2					X
<i>Herbaspirillum</i>	17	11	+ (1)	1	18					X
<i>Hermiimonas</i>	2	2	+ (2)	2	4		X			
<i>Hydrogenophaga</i>	2	2	+ (2)	2	4	X			X	
<i>Hymenobacter</i>	–	–	+ (2)	2	2					X
<i>Ideonella</i>	–	–	+ (1)	1	1				X	
<i>Iodobacter</i>	1	1	+ (2)	1	3					X
<i>Janibacter</i>	2	2	+ (1)	1	3	X				
<i>Janthinobacterium</i>	1	1	+ (1)	1	2					X
<i>Kinneretia</i>	–	–	+ (1)	1	1					X
<i>Klebsiella</i>	31	3	+ (1)	1	32					X
<i>Lacihabitans</i>	–	–	+ (1)	1	1					X
<i>Lactobacillus</i>	249	97	+ (1)	1	250					X
<i>Limnobacter</i>	–	–	+ (1)	1	1	X				
<i>Massilia</i>	4	2	+ (4)	4	8				X	X
<i>Methylibium</i>	–	–	+ (1)	1	1	X				
<i>Methylobacterium</i>	22	11	+ (11)	8	33	X		X		X
<i>Methylosinus</i>	–	–	+ (2)	2	2					X
<i>Microbacterium</i>	55	38	+ (2)	2	57	X			X	
<i>Micrococcus</i>	21	4	+ (1)	1	22			X		
<i>Mucilaginibacter</i>	–	–	+ (10)	10	10					X
<i>Mycobacterium</i>	74	39	+ (7)	6	81		X	X		X
<i>Nevskia</i>	–	–	+ (2)	1	2					X
<i>Nocardia</i>	117	47	+ (1)	1	118		X			
<i>Nocardioides</i>	2	2	+ (1)	1	3				X	
<i>Novispirillum</i>	–	–	+ (1)	1	1				X	
<i>Novosphigobium</i>	18	12	+ (5)	4	23			X		X
<i>Paenibacillus</i>	137	69	+ (6)	5	143	X	X	X		X
<i>Pararhizobium</i>	1	1	+ (1)	1	2	X				
<i>Parasediminibacterium</i>	–	–	+ (1)	1	1					X
<i>Pedobacter</i>	2	2	+ (10)	10	12	X			X	X
<i>Pelomonas</i>	2	2	+ (1)	1	3					X
<i>Pheaeospirillum</i>	–	–	+ (1)	1	1					X
<i>Phyllobacterium</i>	–	–	+ (1)	1	1	X				
<i>Polaromonas</i>	–	–	+ (3)	2	3	X				
<i>Porphyrobacter</i>	–	–	+ (1)	1	1			X		
<i>Propionivibrio</i>	–	–	+ (1)	1	1					X
<i>Pseudomonas</i>	174	93	+ (10)	5	184		X	X		X

(continued on next page)

Table 1 (continued)

Genera included in the database	BDAL		DWL		BDAL+ DWL No. MSP	Origin of MSP (DWL)				
	No. MSP	No. species	No. MSP	No. species		BW	SW	DN	PW	Ref
<i>Pseudoxanthomonas</i>	5	3	+ (2)	2	7			X		
<i>Psychrobacillus</i>	3	3	+ (1)	1	4		X			
<i>Psychrobacter</i>	3	2	+ (8)	5	11	X				X
<i>Rheinheimera</i>	1	1	+ (4)	4	5				X	X
<i>Rhizobium</i>	16	3	+ (1)	3	17	X				X
<i>Rhodococcus</i>	96	27	+ (5)	3	101			X		X
<i>Rhodoferax</i>	–	–	+ (1)	1	1				X	
<i>Roseateles</i>	–	–	+ (2)	2	2					X
<i>Roseomonas</i>	4	1	+ (2)	1	6				X	X
<i>Runella</i>	–	–	+ (2)	2	2					X
<i>Sediminibacterium</i>	–	–	+ (1)	1	1					X
<i>Shinella</i>	–	–	+ (1)	1	1			X		
<i>Sphingobacterium</i>	16	6	+ (1)	1	17					X
<i>Sphingobium</i>	15	11	+ (6)	6	21			X	X	
<i>Sphingomonas</i>	62	32	+ (10)	10	75	X		X		X
<i>Sphingopyxis</i>	5	5	+ (8)	8	15	X	X	X		X
<i>Sphingorhabdus</i>	–	–	+ (1)	1	1					X
<i>Tabrizicola</i>	–	–	+ (1)	1	1					X
<i>Taeseokella</i>	–	–	+ (1)	1	1					X
<i>Undibacterium</i>	–	–	+ (1)	1	1					X
<i>Variovorax</i>	7	1	+ (2)	2	9	X				
<i>Vibrio</i>	92	54	+ (2)	2	95					X
<i>Williamsia</i>	–	–	+ (1)	1	1		X			
<i>Xanthomonas</i>	36	17	+ (1)	1	37		X			
<i>Yersinia</i>	71	12	+ (1)	1	72					X
TOTAL	1783	824	319	240	2102	23	11	25	18	62

MSPs. The predominant genera were *Mucilaginibacter* with 10 MSPs, followed by *Bosea* with 5 MSPs and *Aquabacterium*, *Chitinimonas* and *Polaromonas* with 3 MSPs each. The newly added genera are shown in Table 1. Of the 44 new genera included in the DWL, 30 correspond to reference strains and 14 to isolates from treated water and non-treated water. At the species level, the DWL database provides 189 species not represented in the BDAL database, 67 of which are potentially new taxa.

Independently of their origin, whether environmental or reference strains, the new MSPs expanded the in-house database and improved the capacity of MALDI-TOF MS to identify drinking water isolates. Distribution of the DWL MSPs per genera and origin is shown in Table 1. Briefly, the genera included in the DWL were represented by variable numbers of MSPs. The most highly represented genus was *Aeromonas*, with 38 MSP entries (11.9%), followed by *Chryseobacterium* with 22 MSPs (6.9%), *Flavobacterium* with 18 MSPs (5.6%), *Sphingomonas* with 13 MSPs (4.1%) and *Methylobacterium* with 11 MSPs (3.4%). The genera *Pedobacter*, *Pseudomonas*, *Sphingopyxis* and *Mucilaginibacter* contributed 10 MSPs (3.1% each), whereas *Brevundimonas* provided 9 MSPs (2.8%), *Psychrobacter* 8 MSPs (2.5%), *Mycobacterium* 7 MSPs (2.2%) and *Sphingobium* and *Paenibacillus* 6 MSPs (1.9% each). The genera *Blastomonas*, *Rhodococcus*, *Bosea*, *Novosphigobium* and *Caulobacter* provided 5 MSPs (1.6% each) and the genera *Bacillus*, *Rheinheimera* and *Massilia* 4 MSPs (1.3% each).

3.4. Validation of the DWL database

The newly developed DWL was validated by (a) internal assays with reference CECT strains and drinking water isolates; (b) external blind trials; (c) re-identification of the drinking water isolates recovered in this project, and (d) identification of new isolates from routine laboratory analysis.

3.4.1. Internal assays

An internal assay was carried out with 9 CECT reference strains (CECT 317 *Brevundimonas diminuta*, CECT 7302 *Chryseobacterium aquaticum*, CECT 7791 *Flavobacterium tructae*, CECT 5998 *Methylobacterium aquaticum*, CECT 7550 *Mucilaginibacter myungsuensis*, CECT 7273 *Mycobacterium latzerense*, CECT 153 *Paenibacillus polymyxa*, CECT 7114 *Pedobacter aquatilis* and CECT 8016 *Sphingopyxis italica*). The

selection was based on the bacterial diversity found in drinking water and protein extraction difficulty in order to challenge the identification capacity of the new database. The strains were analyzed by MALDI-TOF MS, as described in Material and Methods (direct transfer, formic acid extended direct transfer and acid/acetonitrile extraction methods). Six out of the nine strains were successfully identified at the species level with scores of 2.04–2.55; two strains were identified at the genus level (1.73–1.94) and one was not identified (1.48–1.56) (Table S3). In general, differences between extraction methods were not observed, the score values being very close. Strain CECT 7273 *Mycobacterium latzerense* was only identified to the genus level by the extraction method (score 1.73–1.80). As described in other studies, due to their complex cell walls, *Mycobacterium* spp. require a specific extraction method for a better protein recovery, which may affect MALDI-TOF results (Alcolea-Medina et al., 2019). Strain CECT 153 *Paenibacillus polymyxa* was not identified (scores 1.49–1.56) when using the DWL but genus level identification was achieved with the BDAL database (scores 1.84–1.98). It is well known that sporulation can affect identification by MALDI-TOF analysis, as spores differ in protein content compared to vegetative cells (Lasch et al., 2009).

3.4.2. External blind trials

For external validation, four rounds of proficiency testing of the DWL were carried out, each round involving at least three Spanish research centers and 29 strains (Table S4), which represents about 10% of the DWL content. One batch of cell culture was prepared per strain and distributed among the centers for the parallel analysis by MALDI TOF MS using the acid/acetonitrile extraction method and the mass spectrometer device available in the center (in all, three different instruments were used). As a result, 81% of measurements resulted in identification at probable species level (score >2.000) and 13% at probable genus level (score 1.700–1.999), 6% not leading to identification (score <1.699). In general, spectra were consistent, although occasional inter-laboratory and even intra-laboratory variability was observed. Differences in log score values between research centers can be explained either because of the cell composition (sporulated cultures, cell wall complexity), the elapsed time before sample processing, the conditions of sample preservation, or sample handling.

Likewise, inter-laboratory differences in analytical sensitivity were

registered for certain taxonomic groups that are nearly identical by 16S rRNA gene sequencing (Janda and Abbott, 2007). This was the case of strain *Flavobacterium* sp. CECT 9288, for which the highest log score value matched a closely related phylogenetic species of the genus *Flavobacterium*. As a result, two out of six centers only achieved identification at the genus level.

Slightly different results were observed depending on the spectrometer device used. For some strains, profiles clustered separately in Microflex vs UltraflexXtreme/Ultraflex II. Moreover, the quality of spectra also directly affected the results. The higher laser intensity of automatic compared to manual acquisition may produce background noise that impairs sensitivity, resulting in false or unreliable classification. For instance, one of the six laboratories in the trial acquired three spectra for strain CECT 9470 *Phyllobacterium myrsinacearum*, yet identification at species level was achieved with only one spectrum.

3.4.3. Re-identification of the drinking water isolates recovered in this study

The extended database (BDAL plus DWL) allowed offline re-analysis of isolates unidentified with previously acquired spectra. The performances of both databases are shown in Fig. 3. The BDAL database could identify 35% of isolates to levels of probable genus (score 1.700–1.999), secure genus and probable species (score 2.000–2.299), or highly probable species identification (score ≥ 2.300 –3.000), which increased up to 95% when the DWL was used as well. The BDAL library alone failed to identify 2088 (54.8%) of the isolates, whereas 618 (16.2%) were identified at the genus level and 1103 (29%) at the species level. In contrast, when the same spectra were searched against the BDAL database plus the DWL, the identification performance improved by 47%, 2904 of the isolates being identified at the species level (76.2%) and 296 remaining unidentified (7.8%). The unidentified fraction corresponded to single representatives (non-clustered isolates) not included in the DWL as they were detected only once throughout the study.

In general, the addition of new MSP entries in the DWL increased the coverage of genus and species diversity, thereby improving the discrimination capacity of the original database. As a result (see Fig. 1B), 714 isolates from the distribution networks were successfully identified at the species level (65.8%), 263 at the genus level (24.2%) and 108 isolates remained unknown (10%). The results for process water samples were similar: from a total of 1085 isolates, 240 were identified at species level (54.3%), 122 at genus level (27.6%) and 80 remained unknown (18.1%). In the two groups, the percentage of unknown samples decreased by 38.9% and 23.5%, respectively, when using DWL. Surprisingly, this reduction was significantly greater for the bottled

natural mineral water samples. When the 1926 spectra were searched against the extended database, 1671 were identified at the species level (86.8%) and another 157 (8.2%) at the genus level; only 98 isolates remained unidentified, which represented an improvement of 60.9% over the BDAL database (Fig. 1A). The extended database also gave better results in spring water samples: a total of 97.2% (346 out of 356) of isolates were successfully identified, 78.4% at species level and 18.8% at genus level, with only 10 isolates unidentified.

As shown, the DWL contributed new MSPs on environmental strains lacking in the original database and constitutes a good tool for the routine testing of drinking water samples. Although databases are frequently updated by the providers, results of the present study highlighted, in accordance with other authors (De Carolis et al., 2014; Kopicakova et al., 2014; Rahi et al., 2016; Seuylemezian et al., 2018) that there is a need for continuously updated in-house databases to increase taxonomic resolution, especially for microbial ecology studies.

4. Conclusions

The Drinking Water Library presented here is composed of 319 MALDI-TOF MS profiles of 120 bacteria isolated from different drinking water sources (distribution networks and mineral water) of different origins, mainly Catalonia (North-East Spain), Andorra and the South of France, and a selection of 199 reference strains of environmental or water-related origin from the Spanish Type Culture Collection. The new library, created using Bruker UltraflexXtreme MALDI-TOF MS equipment, contains MSPs for 164 species of 53 genera already included in the Bruker Daltonics database increasing the species coverage, and 67 possible new species belonging to 44 genera not previously included. Strains of 120 bacteria, including representative strains of potential new species, were deposited at the CECT for public accession, following the established protocols. Further studies are in progress to elucidate the taxonomic status of strains that might belong to new taxa.

The new DWL, which extended the original Bruker Daltonics database with 319 new MSPs for 96 genera from water and other related environments, improved the resolution power for bacterial identification by up to 76%.

Overall, bacterial identification by MALDI-TOF MS is an easy and fast high-throughput technique with low running costs and remarkable specificity. In the present study, there were slight differences in the results between spectrometers, but the method still proved a good choice for interlaboratory collaboration. Moreover, using the formic acid extended direct transfer protocol allowed the operation time to be significantly reduced and when combined with the DWL, it provided a rapid and successful identification of drinking water bacteria. Able to improve identification, MALDI-TOF MS and the DWL constitute an excellent tool for multiple applications, ranging from basic research on water bacterial communities to routine analysis, i.e., in water testing laboratories and water supply companies.

Supplementary material

Additional file 1: tables and figures

CRedit authorship contribution statement

Anna Pinar-Méndez: Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Sonia Fernández:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **David Baquero:** Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **Carles Vilaró:** Formal analysis, Methodology, Writing – review & editing. **Belén Galofré:** Conceptualization, Supervision, Writing – review & editing. **Susana González:** Conceptualization, Supervision, Writing – review & editing. **Lidia Rodrigo-Torres:** Formal analysis, Methodology, Writing – review & editing. **David R. Arahál:**

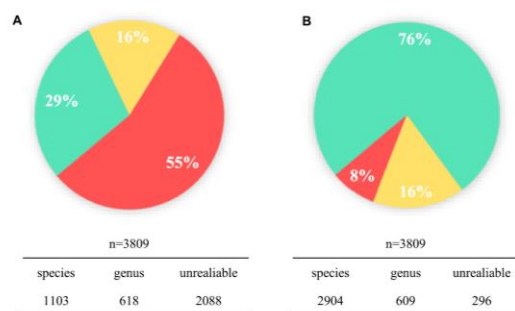


Fig. 3. The percentage of isolates identified using the original Bruker Daltonics (BDAL) database (A) compared with the BDAL database plus the Drinking Water Library (B). In green: highly probable species (score ≥ 2.300 –3.000), secure genus and probable species (score 2.000–2.299); in yellow: probable genus (score 1.700–1.999), and in red: unreliable results (score ≤ 1.699). Number of isolates (n) according to identification results are shown in the table below the graph.

Conceptualization, Supervision, Writing – review & editing. **M. Carmen Macián**: Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **María A. Ruvira**: Conceptualization, Supervision, Formal analysis, Methodology, Writing – review & editing. **Rosa Aznar**: Conceptualization, Supervision, Writing – review & editing. **Laia Caudet-Segarra**: Formal analysis, Methodology, Writing – review & editing. **Laura Sala-Comorera**: Formal analysis, Methodology, Writing – review & editing. **Francisco Lucena**: Conceptualization, Supervision, Writing – review & editing. **Anicet R. Blanch**: Conceptualization, Supervision, Writing – review & editing. **Cristina García-Aljaro**: Conceptualization, Formal analysis, Methodology, Supervision, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study was supported by the Spanish Ministry of Science and Innovation (grant number RTC-2015-4496-2) and Pla de Doctorats Industrials of the Catalan Government (2016 DI 083). We thank Dr. Francesc Márquez Garrido (Bruker Daltonics Spain) for technical assistance in constructing the customized Drinking Water Library with MALDI-TOF MS during the three years of the study. We also thank all the technicians and students from the institutions involved in this work for their support: María José Arnedo, Ivan Manero, Eva Duran, Daniel Toribio, Judit Martínez, Miquel Jorba, Miriam Navarro, Beatriz Pinto and Rosa M^a Gimenez.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.watres.2021.117543](https://doi.org/10.1016/j.watres.2021.117543).

References

- Alatoom, A.A., Cunningham, S.A., Ihde, S.M., Mandrekar, J., Patel, R., 2011. Comparison of direct colony method versus extraction method for identification of gram-positive cocci by use of Bruker biotyper matrix-assisted laser desorption/ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 49, 2868–2873. <https://doi.org/10.1128/JCM.00506-11>.
- Alcolea-Medina, A., Fernandez, M.T.C., Montiel, N., García, M.P.L., Sevilla, C.D., North, N., Lirola, M.J.M., Wilks, M., 2019. An improved simple method for the identification of Mycobacteria by MALDI-TOF MS (Matrix-Assisted Laser Desorption-Ionization mass spectrometry). *Sci. Rep.* 9, 5–10. <https://doi.org/10.1038/s41598-019-56604-7>.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
- Angelakis, E., Million, M., Henry, M., Raoult, D., 2011. Rapid and accurate bacterial identification in probiotics and yoghurts by MALDI-TOF mass spectrometry. *J. Food Sci.* 76, M568–M572. <https://doi.org/10.1111/j.1750-3841.2011.02369.x>.
- Avanzi, I.R., Gracioso, L.H., Baltazar, M.D.P.G., Karolski, B., Perpetuo, E.A., do Nascimento, C.A.O., 2017. Rapid bacteria identification from environmental mining samples using MALDI-TOF MS analysis. *Environ. Sci. Pollut. Res.* 24, 3717–3726. <https://doi.org/10.1007/s11356-016-8125-8>.
- Bartram, J., Cotruvo, J., Exner, M., Fricker, C., Glasmacher, A., 2003. Expert Consensus in Heterotrophic Plate Counts and Drinking Water Safety (Eds.). IWA Publishing, London, UK, pp. 1–11.
- Boers, S.A., Jansen, R., Hays, J.P., 2019. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 1059–1070. <https://doi.org/10.1007/s10096-019-03520-3>.
- Chalupová, J., Raus, M., Sedlářová, M., Šebela, M., 2014. Identification of fungal microorganisms by MALDI-TOF mass spectrometry. *Biotechnol. Adv.* 32, 230–241. <https://doi.org/10.1016/j.biotechadv.2013.11.002>.
- Cunningham, S.A., Patel, R., 2015. Standard matrix-assisted laser desorption/ionization-time of flight mass spectrometry reagents may inactivate potentially hazardous bacteria. *J. Clin. Microbiol.* 53, 2788–2789. <https://doi.org/10.1128/JCM.00957-15>.
- De Carolis, E., Vella, A., Vaccaro, L., Torelli, R., Posteraro, P., Ricciardi, W., Sanguinetti, M., Posteraro, B., 2014. Development and validation of an in-house database for matrix-assisted laser desorption/ionization-time of flight mass spectrometry-based yeast identification using a fast protein extraction procedure. *J. Clin. Microbiol.* 52, 1453–1458. <https://doi.org/10.1128/JCM.03355-13>.
- Drevinek, M., Dresler, J., Klimentova, J., Písa, L., Hubalek, M., 2012. Evaluation of sample preparation methods for MALDI-TOF MS identification of highly dangerous bacteria. *Lett. Appl. Microbiol.* 55, 40–46. <https://doi.org/10.1111/j.1472-765X.2012.03255.x>.
- Eddabara, R., Prévost, G., Scheffel, J.M., 2012. Rapid discrimination of environmental Vibrio by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Microbiol. Res.* 167, 226–230. <https://doi.org/10.1016/j.micres.2011.09.002>.
- Ghosh, A.K., Paul, S., Sood, P., Rudramurthy, S.M., Rajbanshi, A., Jillwin, T.J., Chakrabarti, A., 2015. Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry for the rapid identification of yeasts causing bloodstream infections. *Clin. Microbiol. Infect.* 21, 372–378. <https://doi.org/10.1016/j.cmi.2014.11.009>.
- Gunnarsdottir, M.J., Gardarsson, S.M., Figueras, M.J., Puigdomènech, C., Juárez, R., Saucedo, G., Arnedo, M.J., Santos, R., Monteiro, S., Avery, L., Pagaling, E., Allan, R., Abel, C., Eglitis, J., Hamsch, B., Hügler, M., Rajkovic, A., Smigic, N., Udovicki, B., Albrechtsen, H.J., López-Avilés, A., Hunter, P., 2020. Water safety plan enhancements with improved drinking water quality detection techniques. *Sci. Total Environ.* 698, 134185. <https://doi.org/10.1016/j.scitotenv.2019.134185>.
- Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D., Cromeans, T.L., 2007. Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-Liter tap water samples. *Appl. Environ. Microbiol.* 73, 4218–4225. <https://doi.org/10.1128/AEM.02713-06>.
- Janda, J.M., Abbott, S.L., 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J. Clin. Microbiol.* 45, 2761–2764. <https://doi.org/10.1128/JCM.01228-07>.
- Kim, O.S., Cho, Y.J., Lee, K., Yoon, S.H., Kim, M., Na, H., Park, S.C., Jeon, Y.S., Lee, J.H., Yi, H., Won, S., Chun, J., 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* 62, 716–721. <https://doi.org/10.1099/ij.s.0.038075-0>.
- Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., Pristas, P., 2014. Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry. *Chem. Pap.* 68, 1435–1442. <https://doi.org/10.2478/s11696-014-0612-0>.
- Lasch, P., Nattermann, H., Erhard, M., Stämmeler, M., Grunow, R., Bannert, N., Appel, B., Naumann, D., 2008. MALDI-TOF mass spectrometry compatible inactivation method for highly pathogenic microbial cells and spores. *Anal. Chem.* 80, 2026–2034. <https://doi.org/10.1021/ac701822j>.
- Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.P., Göker, M., 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinform.* 14, 60. <https://doi.org/10.1186/1471-2105-14-60>.
- Pavlovic, M., Huber, L., Konrad, R., Busch, U., 2013. Application of MALDI-TOF MS for the identification of food borne bacteria. *Open Microbiol. J.* 7, 135–141. <https://doi.org/10.2174/1874285801307010135>.
- Rahi, P., Prakash, O., Shouche, Y.S., 2016. Matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) based microbial identifications: challenges and scopes for microbial ecologists. *Front. Microbiol.* 7, 1–12. <https://doi.org/10.3389/fmicb.2016.01359>.
- Reasoner, D.J., Geldreich, E.E., 1985. A new medium for the enumeration and subculture of bacteria from potable water. *Appl. Environ. Microbiol.* 49, 1–7. <https://doi.org/10.1128/aem.49.1.1-7.1985>.
- Ruelle, V., El Moulaj, B., Zorzi, W., Ledent, P., De Pauw, E., 2004. Rapid identification of environmental bacterial strains by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 18, 2013–2019. <https://doi.org/10.1002/rcm.1584>.
- Sala-Comorera, L., Blanch, A.R., Vilaró, C., Galofré, B., García-Aljaro, C., 2017. Heterotrophic monitoring at a drinking water treatment plant by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry after different drinking water treatments. *J. Water Health* 15, 885–897. <https://doi.org/10.2166/wh.2017.090>.
- Sala-Comorera, L., Caudet-Segarra, L., Galofré, B., Lucena, F., Blanch, A.R., García-Aljaro, C., 2020. Unravelling the composition of tap and mineral water microbiota: divergences between next-generation sequencing techniques and culture-based methods. *Int. J. Food Microbiol.* 334, 108850. <https://doi.org/10.1016/j.ijfoodmicro.2020.108850>.
- Sala-Comorera, L., Vilaró, C., Galofré, B., Blanch, A.R., García-Aljaro, C., 2016. Use of matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry for bacterial monitoring in routine analysis at a drinking water treatment plant. *Int. J. Hyg. Environ. Health* 219, 577–584. <https://doi.org/10.1016/j.ijheh.2016.01.001>.
- Santos, I., Martín, M., Carlton, D., Amorim, C., Castro, P., Hildenbrand, Z., Schug, K., 2017. MALDI-TOF MS for the identification of cultivable organic-degrading bacteria in contaminated groundwater near unconventional natural gas extraction sites. *Microorganisms* 5, 47. <https://doi.org/10.3390/microorganisms5030047>.
- Santos, I.C., Hildenbrand, Z.L., Schug, K.A., 2016. Applications of MALDI-TOF MS in environmental microbiology. *Analyst* 141, 2827–2837. <https://doi.org/10.1039/c6an00131a>.
- Sárvári, K.P., Soki, J., Iván, M., Miszti, C., Latkoczy, K., Meleg, S., Urbán, E., 2018. MALDI-TOF MS versus 16S rRNA sequencing: minor discrepancy between tools in identification of bacteroides isolates. *Acta Microbiol. Immunol. Hung.* 65, 173–181. <https://doi.org/10.1556/030.64.2017.025>.

- Sasser, M., 1990. Identification of bacteria by gas chromatography of cellular fatty acids. *USFCC Newsl.* 20, 1–6.
- Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W., Vaishampayan, P., 2018. Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. *Front. Microbiol.* 9, 1–8. <https://doi.org/10.3389/fmicb.2018.00780>.
- Tan, K.E., Ellis, B.C., Lee, R., Stamper, P.D., Zhang, S.X., Carroll, K.C., 2012. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: a bench-by-bench study for assessing the impact on time to identification and cost-effectiveness. *J. Clin. Microbiol.* 50, 3301–3308. <https://doi.org/10.1128/JCM.01405-12>.
- Theel, E.S., Schmitt, B.H., Hall, L., Cunningham, S.A., Walchak, R.C., Patel, R., Wengenack, N.L., 2012. Formic acid-based direct, on-plate testing of yeast and *Corynebacterium* species by Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 50, 3093–3095. <https://doi.org/10.1128/JCM.01045-12>.
- Timperio, A.M., Gorrasi, S., Zolla, L., Fenice, M., 2017. Evaluation of MALDI-TOF mass spectrometry and MALDI BioTyper in comparison to 16S rDNA sequencing for the identification of bacteria isolated from Arctic sea water. *PLoS ONE* 12, 1–15. <https://doi.org/10.1371/journal.pone.0181860>.
- Tracz, D.M., Antonation, K.S., Corbett, C.R., 2016. Verification of a matrix-assisted laser desorption ionization-time of flight mass spectrometry method for diagnostic identification of high-consequence bacterial pathogens. *J. Clin. Microbiol.* 54, 764–767. <https://doi.org/10.1128/JCM.02709-15>.
- Tsuchida, S., Umemura, H., Nakayama, T., 2020. Current status of matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS) in clinical diagnostic microbiology. *Molecules* 25, 4775.
- Yoon, S.H., Ha, S.M., Kwon, S., Lim, J., Kim, Y., Seo, H., Chun, J., 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* 67, 1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.
- Yu, Z., Peruzzy, M.F., Dumolin, C., Joossens, M., Houf, K., 2019. Assessment of food microbiological indicators applied on poultry carcasses by culture combined MALDI-TOF MS identification and 16S rRNA amplicon sequencing. *Food Microbiol.* 82, 53–61. <https://doi.org/10.1016/j.fm.2019.01.018>.
- Lasch, P., Beyer, W., Nattermann, H., Stämmler, M., Siegbrecht, E., Grunow, R., Naumann, D., 2009. Identification of *Bacillus anthracis* by using matrix-assisted laser desorption ionization-time of flight mass spectrometry and artificial neural networks. *Appl. Environ. Microbiol.* 75 (22), 7229–7242. <https://doi.org/10.1128/AEM.00857-09>.

Supplementary Material

Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database.

Anna Pinar-Méndez ^{a,d,e,1*}, Sonia Fernández ^{b1}, David Baquero^b, Carles Vilaró^a, Belén Galofré^a, Susana González^b, Lidia Rodrigo-Torres L^c, David R. Arahal^c, M. Carmen Macián^c, María A. Ruvira^c, Rosa Aznar^c, Laia Caudet-Segarra^{d,e}, Laura Sala-Comorera^{d,e}, Francisco Lucena^{d,e}, Anicet R. Blanch^{d,e}, Cristina Garcia-Aljaro^{d,e*}

^a Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, Barcelona, Spain

^b Cetaqua, Water technology center, Cornellà de Llobregat, Spain

^c Departamento de Microbiología y Ecología and Colección Española de Cultivos Tipo (CECT), Universitat de València, Valencia, Spain

^d Department of Genetics, Microbiology and Statistics, University of Barcelona, Barcelona, Spain

^e The Water Research Institute, University of Barcelona, Barcelona, Spain

¹ These authors contributed equally to the manuscript

TABLE S1. Reference bacterial strains corresponding to the MSPs included in the Drinking Water Library

Collection Acronym	Strain No	Name	Isolated from
CECT	4632	<i>Acinetobacter sp.</i>	Freshwater stream
CECT	7849	<i>Acinetobacter sp.</i>	Small stream
CECT	4911 ^T	<i>Aeromonas allosaccharophila</i>	Feces
CECT	4220	<i>Aeromonas allosaccharophila</i>	Feces
CECT	7289 ^T	<i>Aeromonas aquariorum</i>	Aquaria
CECT	4227 ^T	<i>Aeromonas bestiarum</i>	Fish
CECT	7451	<i>Aeromonas bestiarum</i>	Eel
CECT	5219	<i>Aeromonas bestiarum</i>	Cake
CECT	4342 ^T	<i>Aeromonas encheleia</i>	Eel
CECT	4985	<i>Aeromonas encheleia</i>	Hospital environment
CECT	5027	<i>Aeromonas encheleia</i>	Drinking water well
CECT	4224 ^T	<i>Aeromonas eucrenophila</i>	Freshwater fish
CECT	4855	<i>Aeromonas eucrenophila</i>	Water
CECT	4827	<i>Aeromonas eucrenophila</i>	Ascites of carp
CECT	7401 ^T	<i>Aeromonas fluviialis</i>	River
CECT	4221 ^T	<i>Aeromonas hydrophila</i> subsp. <i>anaerogenes</i>	Used oil-emulsions
CECT	4588	<i>Aeromonas hydrophila</i> subsp. <i>anaerogenes</i>	Feces
CECT	839T	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i>	Milk
CECT	4330	<i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i>	Eel tank water
CECT	4486 ^T	<i>Aeromonas ichthiosmia</i>	Surface water
CECT	4228 ^T	<i>Aeromonas jandaei</i>	Feces
CECT	4232 ^T	<i>Aeromonas media</i>	Fish farm effluent
CECT	7443 ^T	<i>Aeromonas piscicola</i>	Salmon fish
CECT	5176 ^T	<i>Aeromonas popoffi</i>	Drinking water treatment plant
CECT	5250	<i>Aeromonas popoffii</i>	Continental water
CECT	895 ^T	<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	Trout fish
CECT	4238	<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	Trout fish from river
CECT	4239	<i>Aeromonas salmonicida</i> subsp. <i>achromogenes</i>	Minnow fish from lake
CECT	896 ^T	<i>Aeromonas salmonicida</i> subsp. <i>masoucida</i>	Masou fish
CECT	5752 ^T	<i>Aeromonas salmonicida</i> subsp. <i>pectinolytica</i>	Polluted river

CECT	5753	<i>Aeromonas salmonicida</i> subsp. <i>pectinolytica</i>	Polluted river
CECT	894 ^T	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	Salmon fish
CECT	5209	<i>Aeromonas salmonicida</i> subsp. <i>salmonicida</i>	Sea water
CECT	5179 ^T	<i>Aeromonas salmonicida</i> subsp. <i>smithia</i>	Roach fish
CECT	7082 ^T	<i>Aeromonas tecta</i>	Feces
CECT	4257 ^T	<i>Aeromonas veronii</i>	Sputum of drowning victim
CECT	4819	<i>Aeromonas veronii</i>	Stool
CECT	4910	<i>Aeromonas veronii</i>	Feces
CECT	5761	<i>Aeromonas veronii</i>	Midgut of <i>Culex quinquefasciatus</i>
CECT	7060	<i>Aeromonas veronii</i>	Drinking water supply
CECT	5683 ^T	<i>Alcanivorax balearicus</i>	Subterranean saline lake
CECT	8580	<i>Ampullimonas aquatilis</i>	Bottled mineral water
CECT	8581 ^T	<i>Ampullimonas aquatilis</i>	Bottled mineral water
CECT	4166 ^T	<i>Ancylobacter aquaticus</i>	Lake water
CECT	7957	<i>Arcicella</i> sp.	Stream
CECT	7819	<i>Arcobacter</i> sp.	Pond
CECT	7825	<i>Arcobacter</i> sp.	Freshwater pond
CECT	5669	<i>Azoarcus</i> sp.	Contaminated aquifer
CECT	7823	<i>Azonexus</i> sp.	Freshwater pond
CECT	8551 ^T	<i>Belliella kenyensis</i>	Lake sediment
CECT	4273 ^T	<i>Blastomonas natatoria</i>	Swimming pool water
CECT	313	<i>Brevundimonas diminuta</i>	Contaminant in a culture of <i>Bacillus cereus</i>
CECT	317 ^T	<i>Brevundimonas diminuta</i>	Freshwater stream
CECT	7729 ^T	<i>Brevundimonas faecalis</i>	Sewage water from a hospital
CECT	834 ^T	<i>Brevundimonas intermedia</i>	Freshwater pond
CECT	327 ^T	<i>Brevundimonas vesicularis</i>	Urinary-bladder epithelium of <i>Hirudo</i> sp. (leech)
CECT	5640	<i>Caulobacter</i> sp.	Bottled mineral water
CECT	5648	<i>Caulobacter</i> sp.	River water
CECT	421	<i>Caulobacter vibrioides</i>	Well water

CECT	8008	<i>Chitinimonas sp.</i>	Stream
CECT	8101	<i>Chitinimonas sp.</i>	Stream
CECT	7703 ^T	<i>Chitinimonas viridis</i>	Pond
CECT	494 ^T	<i>Chromobacterium violaceum</i>	Freshwater
CECT	7302 ^T	<i>Chryseobacterium aquaticum</i>	Water reservoir
CECT	8497 ^T	<i>Chryseobacterium artocarpi</i>	Rhizosphere soil
CECT	7129 ^T	<i>Chryseobacterium hispanicum</i>	Drinking water distribution network
CECT	7794 ^T	<i>Chryseobacterium oncorhynchi</i>	Trout fish
CECT	7847	<i>Chryseobacterium oncorhynchi</i>	Trout fish
CECT	7357 ^T	<i>Chryseobacterium piscicola</i>	Salmon fish
CECT	7130	<i>Chryseobacterium sp.</i>	Drinking water
CECT	7795	<i>Chryseobacterium sp.</i>	Trout fish
CECT	8086	<i>Chryseobacterium sp.</i>	Wetland
CECT	7846	<i>Chryseobacterium sp.</i>	Trout fish
CECT	7797	<i>Chryseobacterium tructae</i>	Trout fish
CECT	7798 ^T	<i>Chryseobacterium tructae</i>	Trout fish
CECT	7793 ^T	<i>Chryseobacterium viscerum</i>	Trout fish
CECT	7845	<i>Chryseobacterium viscerum</i>	Trout fish
CECT	7547	<i>Chryseobacterium yonginense</i>	Artificial lake
CECT	7864 ^T	<i>Cronobacter universalis</i>	Freshwater
CECT	8087	<i>Duganella sp.</i>	Wetland
CECT	8858 ^T	<i>Emticicia aquatica</i>	Freshwater
CECT	9087 ^T	<i>Emticicia paludis</i>	Freshwater
CECT	8289 ^T	<i>Ferruginibacter yonginensis</i>	Mesotrophic artificial lake
CECT	8365 ^T	<i>Flavobacterium gyeonganense</i>	Stream
CECT	7955 ^T	<i>Flavobacterium jumunjinense</i>	Lake
CECT	7678 ^T	<i>Flavobacterium oncorhynchi</i>	Liver of a diseased rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7848	<i>Flavobacterium oncorhynchi</i>	Liver of a diseased rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7909	<i>Flavobacterium piscis</i>	Gills of rainbow trout, <i>Oncorhynchus mykiss</i>

CECT	7911 ^T	<i>Flavobacterium piscis</i>	Kidney of rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7844 ^T	<i>Flavobacterium plurextorum</i>	Eggs of rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7789	<i>Flavobacterium sp.</i>	Gills of rainbow trout
CECT	8006	<i>Flavobacterium sp.</i>	Lake
CECT	8038	<i>Flavobacterium sp.</i>	Wetland
CECT	8461	<i>Flavobacterium sp.</i>	Freshwater
CECT	7596	<i>Flavobacterium sp.</i>	Freshwater, pond
CECT	7791 ^T	<i>Flavobacterium tructae</i>	Liver of rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7905	<i>Flavobacterium tructae</i>	Gills of rainbow trout, <i>Oncorhynchus mykiss</i>
CECT	7595	<i>Flectobacillus sp.</i>	Freshwater, pond
CECT	7598	<i>Flectobacillus sp.</i>	Freshwater, pond
CECT	8121	<i>Heliimonas saccharivorans</i>	Deep mineral water aquifer
CECT	8122 ^T	<i>Heliimonas saccharivorans</i>	Deep mineral water aquifer
CECT	7143	<i>Herbaspirillum sp.</i>	Drinking water
CECT	9069	<i>Hymenobacter sp.</i>	Seawater
CECT	7546 ^T	<i>Hymenobacter yonginensis</i>	Freshwater from an artificial lake
CECT	4504	<i>Iodobacter fluviatilis</i>	Water of River Wey
CECT	4505 ^T	<i>Iodobacter fluviatilis</i>	Water of River Wey
CECT	7866	<i>Janthinobacterium sp.</i>	Mesotrophic artificial lake
CECT	7319 ^T	<i>Kinneretia asaccharophila</i>	Freshwater from a depth of 5 m in the center of Lake Kinneret
CECT	4519 ^T	<i>Klebsiella terrigena</i>	Drinking water
CECT	7826 ^T	<i>Lacihabitans soyangensis</i>	Lake Soyang
CECT	7355 ^T	<i>Lactobacillus aquaticus</i>	Eutrophic freshwater pond
CECT	7142 ^T	<i>Massilia aurea</i>	Drinking water

CECT	7753 ^T	<i>Massilia umbonata</i>	A mixture of soil collected near Guadalimar (Jaén, Spain) and sewage sludge compost from Granada (Spain)
CECT	8897 ^T	<i>Massilia violacea</i>	Wetland soil
CECT	7069 ^T	<i>Methylobacterium adhaesivum</i>	Drinking water
CECT	5998 ^T	<i>Methylobacterium aquaticum</i>	Drinking water
CECT	7810	<i>Methylobacterium aquaticum</i>	Fruit surface of <i>Arabidopsis thaliana</i>
CECT	5997 ^T	<i>Methylobacterium hispanicum</i>	Drinking water
CECT	7068 ^T	<i>Methylobacterium isbiliense</i>	Drinking water
CECT	7806 ^T	<i>Methylobacterium longum</i>	Phyllosphere of <i>Arabidopsis thaliana</i>
CECT	7180	<i>Methylobacterium sp.</i>	Drinking water
CECT	7814	<i>Methylobacterium sp.</i>	Phyllosphere of <i>Cardamine hirsuta</i>
CECT	7045 ^T	<i>Methylobacterium variabile</i>	Drinking water
CECT	7189	<i>Methylosinus sp.</i>	Drinking water
CECT	7190	<i>Methylosinus sp.</i>	Drinking water
CECT	8102 ^T	<i>Mucilaginibacter aquaedulcis</i>	Freshwater, lake
CECT	7660 ^T	<i>Mucilaginibacter dorajii</i>	Rhizosphere of <i>Platycodon grandiflorum</i>
CECT	7857 ^T	<i>Mucilaginibacter flavus</i>	Swamp
CECT	7550 ^T	<i>Mucilaginibacter myungsuensis</i>	Freshwater from a mesotrophic lake
CECT	7824 ^T	<i>Mucilaginibacter soyangensis</i>	Pond
CECT	7661	<i>Mucilaginibacter sp.</i>	Rhizosphere of <i>Codonopsis lanceolata</i>
CECT	7821	<i>Mucilaginibacter sp.</i>	Pond
CECT	8103	<i>Mucilaginibacter sp.</i>	Stream
CECT	8292	<i>Mucilaginibacter sp.</i>	Lake
CECT	8416	<i>Mucilaginibacter sp.</i>	Soil
CECT	3021 ^T	<i>Mycobacterium alvei</i>	Water of River Besos

CECT	3022 ^T	<i>Mycobacterium brumae</i>	Wastewater
CECT	7273 ^T	<i>Mycobacterium llatzerense</i>	A pure water distribution system of a haemodialysis unit in the Hospital Son Llàtzer
CECT	8638	<i>Mycobacterium sp.</i>	Mud, River Cardener
CECT	7897 ^T	<i>Nevskia aquatilis</i>	Deep mineral water aquifer
CECT	7898	<i>Nevskia aquatilis</i>	Deep mineral water aquifer
CECT	8290	<i>Novosphigibium sp.</i>	Lake
CECT	153	<i>Paenibacillus polymyxa</i>	Water
CECT	155 ^T	<i>Paenibacillus polymyxa</i>	Decomposing plants and soil
CECT	7179	<i>Paenibacillus sp.</i>	Drinking water
CECT	8010 ^T	<i>Parasediminibacterium paludis</i>	Swamp
CECT	7114 ^T	<i>Pedobacter aquatilis</i>	Drinking water
CECT	8725 ^T	<i>Pedobacter lignilitoris</i>	Wood of seashore
CECT	8669 ^T	<i>Pedobacter silvilitoris</i>	Wood of seashore
CECT	7115	<i>Pedobacter sp.</i>	Drinking water
CECT	7544	<i>Pedobacter sp.</i>	Freshwater pond
CECT	7865	<i>Pedobacter sp.</i>	Pond
CECT	8413	<i>Pedobacter sp.</i>	Freshwater
CECT	8689	<i>Pedobacter sp.</i>	Mesotrophic artificial lake
CECT	7234 ^T	<i>Pelomonas puraquae</i>	Haemodialysis water at the Hospital Son Llàtzer
CECT	7186	<i>Pheaeospirillum sp.</i>	Drinking water
CECT	7822	<i>Propionivibrio sp.</i>	Freshwater pond
CECT	7784 ^T	<i>Pseudomonas alcaligenes</i>	Swimming pool water
CECT	899 ^T	<i>Pseudomonas anguilliseptica</i>	Blood of pond-cultured eel, <i>Anguilla japonica</i>
CECT	378 ^T	<i>Pseudomonas fluorescens</i>	Pre-filter tanks, town water works
CECT	5881	<i>Psychrobacter fozii</i>	Water from Johnsons Dock cove

CECT	5889 ^T	<i>Psychrobacter fozii</i>	Soil from Johnsons Dock
CECT	5937 ^T	<i>Psychrobacter frigidicola</i>	Ornithogenic soil
CECT	5309 ^T	<i>Psychrobacter glacincola</i>	Anchor ice from the Amery Ice Shelf
CECT	4492 ^T	<i>Psychrobacter immobilis</i>	Unknown
CECT	5885 ^T	<i>Psychrobacter luti</i>	Muddy soil collected from the inlet Admiralty Bay
CECT	5882	<i>Psychrobacter sp.</i>	Glacier mud, Admiralty Bay
CECT	8623 ^T	<i>Rheinheimera arenilitoris</i>	Marine sand
CECT	3013 ^T	<i>Rhodococcus erythropolis</i>	Soil
CECT	3054	<i>Rhodococcus erythropolis</i>	Water of estuary
CECT	555 ^T	<i>Rhodococcus hoagii</i>	Lung abscess of foal
CECT	4443	<i>Rhodococcus hoagii</i>	Foal
CECT	7248 ^T	<i>Roseateles aquatilis</i>	Industrial water
CECT	7247 ^T	<i>Roseateles terrae</i>	Soil
CECT	7131 ^T	<i>Roseomonas aquatica</i>	Drinking water
CECT	8978	<i>Runella sp.</i>	Wetland freshwater
CECT	7725	<i>Runella sp.</i>	Freshwater, pond
CECT	8100 ^T	<i>Sediminibacterium goheungense</i>	Freshwater reservoir
CECT	7938 ^T	<i>Sphingobacterium detergens</i>	Soil sample taken near the crater lake Lagoa do Fogo
CECT	8383 ^T	<i>Sphingomonas lacus</i>	Soil near a pond
CECT	5832 ^T	<i>Sphingomonas phyllosphaerae</i>	Phyllosphere of a leguminous tree, <i>Acacia caven</i>
CECT	599 ^T	<i>Sphingomonas paucimobilis</i>	Hospital respirator
CECT	7182	<i>Sphingomonas sp.</i>	Drinking water
CECT	7183	<i>Sphingomonas sp.</i>	Drinking water
CECT	8804 ^T	<i>Sphingomonas vulcanisoli</i>	Soil of the Gotjawal lava forest

CECT	8016 ^T	<i>Sphingopyxis italica</i>	Tuff walls of the Roman catacombs of Saint Callixtus
CECT	8531 ^T	<i>Sphingorhabdus arenilitoris</i>	Marine sand
CECT	8860	<i>Tabrizicola sp.</i>	Freshwater
CECT	8198 ^T	<i>Taeseokella kangwonensis</i>	Freshwater
CECT	7724	<i>Undibacterium sp.</i>	Freshwater, pond
CECT	552	<i>Vibrio cholerae</i>	Pilgrim to Mecca
CECT	514 ^T	<i>Vibrio cholerae</i>	Human feces
CECT	4218 ^T	<i>Vibrio mimicus</i>	Human ear, 35 year-old female
CECT	4314 ^T	<i>Yersinia aldovae</i>	Drinking water
DSM	50342 ^T	<i>Pseudomonas alcaligenes</i>	Swimming-pool water
DSM	12111 ^T	<i>Pseudomonas anguilliseptica</i>	Pond-cultured eels, <i>Anguilla japonica</i>
DSM	50090 ^T	<i>Pseudomonas fluorescens</i>	Pre-filter tanks
DSM	50091	<i>Pseudomonas fluorescens</i>	Tap water
DSM	50108	<i>Pseudomonas fluorescens</i>	Enrichment with naphthalene

^T: Type strains

TABLE S2. Drinking water isolates identified by 16S rRNA sequence analysis corresponding to MSPs included in the Drinking Water Library. Distribution Network (DN), process water (PW), spring water (SW) and bottled natural mineral water (BW)

CECT No.	Isolate name	Isolated from	16S rRNA Identification*	Species identity (%)	Phylum*
9348	170207 BCN 21	DN	<i>Acidovorax delafieldii</i>	99.5	P
9495	171027 A14MA 26	BW	<i>Acidovorax</i> sp.	99.2 <i>A. radicans</i> , 99.1 <i>A. facilis</i>	P
9347	170207 BCN 18	DN	<i>Acinetobacter lwoffii</i>	99.7	P
9252	161013 MAM 34	SW	<i>Aeromicrobium</i> sp.	99.2 <i>A. chonae</i> , 98.8 <i>A. tamlense</i>	A
9468	170921 A12MA 49	BW	<i>Aquabacterium commune</i>	99.8	P
9469	170921 A12MC 25	BW	<i>Aquabacterium</i> sp.	98.6 <i>A. parvum</i> , 97.5 <i>A. commune</i>	P
9606	180319 A20MB 6	BW	<i>Aquabacterium</i> sp.	97.5 <i>A. commune</i> , 96.9 <i>A. citratiphilum</i>	P
9247	160118 FC 3	PW	<i>Bacillus simplex</i>	100	F
9349	170207 BCN 04 A	DN	<i>Bacillus</i> sp.	99.0 <i>B. halmapalus</i>	F
9350	170207 BCN 04 B	DN	<i>Bacillus</i> sp.	99.0 <i>B. halmapalus</i>	F
9360	170307 D3 2	DN	<i>Bacillus</i> sp.	97.9 <i>B. cavernae</i>	F
9362	170512 A4D 12	DN	<i>Blastomonas</i> sp.	99.4 <i>B. natatoria</i> , 99.3 <i>B. ursincola</i>	P
9352	170119 RLA 01 orange	DN	<i>Blastomonas</i> sp.	99.0 <i>B. aquatica</i> , 99.0 <i>B. aquatica</i> , 98.9 <i>B. ursincola</i>	P
9351	170119 RLA 01 yellow	DN	<i>Blastomonas</i> sp.	99.2 <i>B. natatoria</i> , 99.0 <i>B. aquatica</i> , 99.1 <i>B. ursincola</i>	P
9427	170614 PSA 03	DN	<i>Blastomonas</i> sp.	99.2 <i>B. aquatica</i> , 98.7 <i>B. natatoria</i>	P
9506	20160502 RIB 087	BW	<i>Bosea eneeae</i>	99.9	P
9325	170327 A5M 12	BW	<i>Bosea</i> sp.	99.2 <i>B. massiliensis</i> , 98.9 <i>B. lupini</i> , 98.8 <i>B. lathyri</i>	P
9449	130214 A10M 36	BW	<i>Bosea</i> sp.	99.6 <i>B. lupini</i> , 99.5 <i>B. vaviloniae</i> , 99.3 <i>B. vestrisii</i>	P
9600	180219 CSF 01	DN	<i>Bosea</i> sp.	99.8 <i>B. lupini</i> , 99.7 <i>B. vestrisii</i>	P

9290	130122 OI 19	PW	<i>Bosea</i> sp.	99.3 <i>B. robiniae</i> , 99.3 <i>B. minatitlanensis</i>	P
9605	121125 A10M 23	BW	<i>Bradyrhizobium</i> sp.	98.4 <i>B. kavangense</i> , 98.3 <i>B. daqingense</i>	
9364	170518 A2D 16	DN	<i>Brevundimonas albiqilva</i>	100	P
9277	160427 MSN 3	DN	<i>Brevundimonas bullata</i>	99.9	P
9317	170320 A1M 38	BW	<i>Brevundimonas kwangchunensis</i>	99.5	P
9466	160201 AB 04	DN	<i>Brevundimonas</i> sp.	98.8 <i>B. lenta</i> , 98.7 <i>B. subvibrioides</i>	P
9326	170327 A5M 43	BW	<i>Caulobacter</i> sp.	99.5 <i>C. segnis</i> , 99.3 <i>C. flavus</i> , 99.3 <i>C. vibrioides</i>	P
9549	180122 PS 09	DN	<i>Caulobacter</i> sp.	99.4 <i>C. segnis</i> , 99.2 <i>C. flavus</i>	P
9391	170515 CSA 01	DN	<i>Chryseobacterium daeguense</i>	100	B
9399	170531 PT 15	PW	<i>Chryseobacterium indoltheticum</i>	99.9	B
9270	20160502 FVL 110	BW	<i>Chryseobacterium</i> sp.	98.4 <i>C. lactis</i> , 98.3 <i>C. rhizoplanae</i>	B
9393	170515 CSA 11	DN	<i>Chryseobacterium</i> sp.	96.4 <i>C. pallidum</i>	B
9390	170515 CSA 02	DN	<i>Chryseobacterium</i> sp.	97.5 <i>C. aquaticum</i> , 97.2 <i>C. gambrini</i>	B
9321	170217 CSF 06	PW	<i>Chryseobacterium</i> sp.	98.3 <i>C. vrystaatense</i> , 97.3 <i>C. bernardetii</i>	B
9293	130820 UF 26	PW	<i>Chryseobacterium</i> sp.	96.0 <i>C. daeguense</i> , 96.1 <i>C. pallidum</i> , 95.9 <i>C. daecheongense</i>	B
9294	130820 OI 09	PW	<i>Cloacibacterium normanense</i>	100	B
9276	160111 SFS 20	PW	<i>Deinococcus seoulensis</i>	99.8	DT
9548	180122 BCN 04	DN	<i>Deinococcus soli</i>	100	DT
9426	170613 D3 03	DN	<i>Domibacillus</i> sp.	99.4 <i>D. robiginosus</i> , 99.3 <i>D. enclensis</i> , 99.2 <i>D. indicus</i>	
9275	151221 SFS 2	PW	<i>Dyadobacter</i> sp.	96.5 <i>D. sediminis</i> , 96.4 <i>D. jaingsuensis</i>	B
9623	180222 TN 02	DN	<i>Dyadobacter</i> sp.	97.5 <i>D. crusticola</i>	B

9511	171031 CRA 03	DN	<i>Ensifer</i> sp.	99.9 <i>E. adhaerens</i> , 99.8 <i>E. sesbaniae</i>	P
9303	160411 SFS 14	PW	<i>Flavobacterium aquatile</i>	99.2	B
9258	161024 MVT 172	SW	<i>Flavobacterium</i> sp.	98.5 <i>F. chungangense</i> , 98.4 <i>F. chilense</i>	B
9288	130122 CR 30	PW	<i>Flavobacterium</i> sp.	97.8 <i>F. succinicans</i> , 97.6 <i>F. granuli</i>	B
9356	160411 SFS 10	PW	<i>Flavobacterium succinicans</i>	99.2	B
9260	161013 MAM 21	SW	<i>Herminiimonas glaciei</i>	100	P
9259	161021 MVT 37	SW	<i>Herminiimonas</i> sp.	99.3 <i>H. fonticola</i> , 99.2 <i>H. aquatilis</i>	P
9471	170921 A12MB 15	BW	<i>Hydrogenophaga</i> sp.	99.8 <i>H. palleronii</i> , 99.0 <i>H. defluvii</i>	P
9291	130820 CR 17	PW	<i>Hydrogenophaga</i> sp.	98.3 <i>H. taeniospiralis</i>	P
9357	160411 SFS 6	PW	<i>Ideonella</i> sp.	98.6	P
9524	20160502 BRO 010	BW	<i>Janibacter hoylei</i>	100	A
9513	171110 A16MB 30	BW	<i>Limnobacter thiooxidans</i>	99.7	P
9437	130820 CR 18	PW	<i>Massilia</i> sp.	99.0 <i>M. yuzhufengensis</i>	P
9622	180319 A19MB 7	BW	<i>Methylibium petroleiphilum</i>	100	P
9607	121125 A10M 19	BW	<i>Methylobacterium</i> <i>marchantiae</i>	99.9	P
9284	170119 CRA 02	DN	<i>Methylobacterium</i> <i>persicinum</i>	100	P
9358	170419 OI 07	PW	<i>Microbacterium</i> <i>azadirachtae</i>	99.4	A
9560	160220 RIB 062	BW	<i>Microbacterium</i> sp.	99.2 <i>M. paraoxydans</i> , 98.5 <i>M. testaceum</i>	A
9550	180123 RPL 01	DN	<i>Micrococcus</i> sp.	99.8 <i>M. luteus</i> , 99.8 <i>M. yunnanensis</i>	A
9255	161021 MVT 49	SW	<i>Mycobacterium</i> <i>frederiksbergense</i>	99.9	A
9257	161021 MVT 46	SW	<i>Mycobacterium</i> <i>frederiksbergense</i>	99.7	A
9287	170119 PSA 04	DN	<i>Mycobacterium hodleri</i>	99.9	A
9254	161024 MVT 101	SW	<i>Nocardia cummidelens</i>	99.4	A
9279	160606 SFS 40	PW	<i>Nocardioides</i> sp.	98.7 <i>N. cavernae</i>	A
9400	170531 PT 14	PW	<i>Novospirillum itersonii</i>	100	P
9250	161114 CRA 3	DN	<i>Novosphingobium</i> sp.	98.8 <i>N. naphthalenivorans</i> , 98.8 <i>N. panipatense</i>	P
9392	170515 CSA 10	DN	<i>Novosphingobium</i> sp.	99.0 <i>N. lindaniclasticum</i>	P

9465	160427 PMD 08	DN	<i>Novosphingobium</i> sp.	97.9 <i>N. ginsenosidimutans</i> , 97.4 <i>N. lentum</i>	P
9599	180219 VDC 02	DN	<i>Novosphingobium</i> sp.	97.7 <i>N. lentum</i> , 97.5 <i>N. hassiacum</i>	P
9262	161021 MVT 59	SW	<i>Paenibacillus borealis</i>	99.4	F
9249	161107 CRN 1	DN	<i>Paenibacillus</i> sp.	94.2 <i>P. nanensis</i> , 94.1 <i>P. yoguinensis</i>	F
9547	160408 FVL 124	BW	<i>Paenibacillus woosongensis</i>	99.9	F
9322	170327 A4M 32	BW	<i>Pararhizobium</i> sp.	99.0 <i>P. herbae</i> , 98.9 <i>P. giardinii</i>	
9246	160111 SFS 10	PW	<i>Pedobacter ruber</i>	99.2	B
9496	171027 A14MB 31	BW	<i>Pedobacter</i> sp.	99.2	B
9470	171004 A13MA 11	BW	<i>Phyllobacterium</i> <i>myrsinacearum</i>	99.6	P
9528	171127 A18MC 8	BW	<i>Polaromonas</i> <i>eurypsychrophila</i>	99.8	P
9603	180319 A19MA 34	BW	<i>Polaromonas ginsengisoli</i>	100	P
9289	20160502 NAQ 285	BW	<i>Polaromonas</i> sp.	97.7 <i>P. vacuolata</i>	P
9519	161114 CSA 06	DN	<i>Porphyrobacter</i> sp.	99.7 <i>P. donghaensis</i> , 99.7 <i>P. colymbi</i>	P
9363	170518 A3D 05	DN	<i>Pseudomonas peli</i>	100	P
9261	161013 MAM 08	SW	<i>Pseudomonas</i> sp.	98.6 <i>P. cuatrocienegasensis</i> , 98.0 <i>P. composti</i>	P
9319	170207 BCN 22	DN	<i>Pseudoxanthomonas</i> <i>japonensis</i>	99.7 <i>P. japonensis</i>	P
9601	180219 CSF 03	DN	<i>Pseudoxanthomonas</i> <i>japonensis</i>	99.7 <i>P. japonensis</i>	P
9561	161924 MVT 41	SW	<i>Psychrobacillus</i> sp.	98.5 <i>P. psychrodurans</i> , 97.9 <i>P. psychrotolerans</i>	F
9559	160410 BRO 039	BW	<i>Psychrobacter pulmonis</i>	100.0	P
9320	170217 CSF 01	PW	<i>Rheinheimera chironomi</i>	99.4	P
9395	170531 FC 08	PW	<i>Rheinheimera</i> sp.	98.7 <i>R. arenilitoris</i> , 98.1 <i>R. chironomi</i>	P
9429	170703 FC 10	PW	<i>Rheinheimera aquatica</i>	99.2	P
9324	170327A5M9	BW	<i>Rhizobium</i> sp.	97.9 <i>R. ipomoeae</i> , 97.5 <i>R. daejeonense</i>	P
9285	170119 CSA 01	DN	<i>Rhodococcus</i> <i>corynebacterioides</i>	99.9	A

9397	170531 PT 08	PW	<i>Rhodoferax</i> sp.	98.8 <i>R. saidenbachensis</i> , 98.1 <i>Albidiferax ferrireducens</i>	P
9278	160606 SFS 39	PW	<i>Roseomonas</i> sp.	97.8 <i>R. lacus</i>	P
9512	170131 PSA 08	DN	<i>Shinella curvata</i>	100	P
9359	170206 PSA 01	DN	<i>Sphingobium limneticum</i>	99.8	P
9510	170613 PSN 09	DN	<i>Sphingobium limneticum</i>	99.9	P
9464	160427 PMD 04	DN	<i>Sphingobium</i> sp.	98.5 <i>S. hydrophobicum</i> , 98.5 <i>S. xenophagum</i>	P
9598	180219 VDC 04	DN	<i>Sphingobium</i> sp.	99.1 <i>S. lactosutens</i> , 98.8 <i>S. abikonensis</i>	P
9396	170531 PT 17	PW	<i>Sphingobium</i> sp.	99.1 <i>S. hydrophobicum</i> , 99.1 <i>S. xenophagum</i>	P
9361	170512 A1D 47	DN	<i>Sphingobium</i> sp.	97.9 <i>S. boeckii</i> , 97.3 <i>Rhizorhapis suberifaciens</i>	P
9551	180123 RPL02	DN	<i>Sphingomonas echinoides</i>	100	P
9447	110320 A10M 3	BW	<i>Sphingomonas</i> sp.	98.7 <i>S. mucosissima</i>	P
9448	110320 A10M 25	BW	<i>Sphingomonas</i> sp.	98.8 <i>S. melonis</i> , 98.8 <i>S. aquatilis</i>	P
9283	170119 CRA 01	DN	<i>Sphingomonas</i> sp.	98.6 <i>S. melonis</i> , 98.6 <i>S. aquatilis</i>	P
9286	170119 CSA 02	DN	<i>Sphingomonas</i> sp.	98.7 <i>S. mucosissima</i>	P
9323	170327A5M3	BW	<i>Sphingomonas</i> sp.	98.2 <i>S. leidy</i> , 97.0 <i>S. melonis</i>	P
9248	160201 MG 2	DN	<i>Sphingomonas yantingensis</i>	99.5	P
9604	180319 A19MA 18	BW	<i>Sphingopyxis bauzanensis</i>	99.8 <i>S. bauzanensis</i>	P
9597	180220 PLF 02	DN	<i>Sphingopyxis fribergensis</i>	99.5	P
9271	20160502 RIB 263	BW	<i>Sphingopyxis italica</i>	99.2	P
9499	20160502 FVL 226	BW	<i>Sphingopyxis</i> sp.	99.5 <i>S. panaciterrae</i> , 99.0 <i>S. chilensis</i>	P
9505	20160502 FVL 133	BW	<i>Sphingopyxis</i> sp.	99.5 <i>S. panaciterrae</i> , 99.5 <i>S. chilensis</i>	P
9467	160127 UF 05	PW	<i>Sphingopyxis soli</i>		P
9428	170613 PSB 04	DN	<i>Sphingopyxis</i> sp.	99.5 <i>S. chilensis</i> , 99.2 <i>S. panaciterrae</i>	P
9520	170515 CSB 12	DN	<i>Sphingopyxis</i> sp.	99.8 <i>S. chilensis</i> , 99.4 <i>S. panaciterrae</i>	P
9302	160606 SFS 30	PW	<i>Sphingorhabdus profundilacus</i>	100	P
9318	170320 A3M 1	BW	<i>Variovorax boronicumulans</i>	99.5	P

9497	171027 A14MC 44	BW	<i>Variovorax</i> sp.	99.2 <i>V. boronicumulans</i> , 98.9 <i>V. paradoxus</i>	P
9253	161021 MVT 27	SW	<i>Williamsia maris</i>	100	A
9558	161024 MVT 64	SW	<i>Xanthomonas</i> sp.	100 <i>X. translucens</i> , 100 <i>X. melonis</i>	P

*Cut-off values according to phyla: 99.0% Actinobacteria (A), 98.8% Firmicutes (F), 98.7% Proteobacteria (P), 98.5% Deinococcus-Thermus (DT) and 98.2% Bacteroidetes (B) (Meier-Kolthoff et al. 2013). When two or more species are above the respective cut-off, the identification is considered only to the genus level.

TABLE S3: Bacterial reference strains selected for internal validation with the BDAL+DWL database and identification score for the different sample preparation methods: direct transfer (Direct), formic acid extended direct transfer (Formic acid) and acid/acetonitrile extraction (Extraction). Score values corresponded to secure genus and probable species (score 2.000-2.299), probable genus (score 1.700-1.999) and unreliable results (score ≤ 1.699).

CECT strains selected for internal validation	Identification score values according to sample preparation method		
	Direct	Formic acid	Extraction
317 <i>Bevundimonas diminuta</i>	2.182	2.327	2.282
7301 <i>Chryseobacterium aquaticum</i>	1.907	1.879	1.938
7791 <i>Flavobacterium tructae</i>	2.131	2.043	2.125
5998 <i>Methylobacterium aquaticum</i>	1.898	1.95	2.23
7550 <i>Mucilaginibacter myungsuensis</i>	2.186	2.327	2.261
7273 <i>Mycobacterium llatzerense</i>	1.319	1.586	1.729
153 <i>Paenibacillus polymyxa</i>	1.613	1.492	1.563
7114 <i>Pedobacter aquatilis</i>	2.093	1.889	2.095
8016 <i>Sphingopyxis funeris</i>	2.038	2.049	2.145

TABLE S4. DWL bacterial strains selected for external validation with the BDAL+DWL database. Results shows the number of research centers participating in the four-round proficiency testing and the identification score values reported for secure genus and probable species (score 2.000-2.299), probable genus (score 1.700-1.999) and unreliable results (score ≤ 1.699). Successful measurement rates were calculated by counting all the identification measurements at any level and dividing them by the total number of measurements among all the research centers.

CECT strains selected for external validation	Identification score			No. Centers
	>2.000	1.700-1.900	<1.699	
9348 <i>Acidovorax delafieldii</i>	6	-	-	6
9495 <i>Acidovorax</i> sp.	6	-	-	6
9252 <i>Aeromicrobium choanae</i>	2	2	2	6
9469 <i>Aquabacterium</i> sp.	4	2	-	6
9466 <i>Brevundimonas</i> sp.	5	1	-	6
9326 <i>Caulobacter</i> sp.	5	-	1	6
9270 <i>Chryseobacterium lactis</i>	3	-	-	3
9293 <i>Chryseobacterium</i> sp.	4	2	-	6
9294 <i>Cloacibacterium normanense</i>	2	2	2	6
9276 <i>Deinococcus seoulensis</i>	3	-	-	3
9275 <i>Dyadobacter</i> sp.	6	-	-	6
9258 <i>Flavobacterium</i> sp.	6	-	-	6
9288 <i>Flavobacterium</i> sp.	4	2	-	6
9260 <i>Herminiimonas glaciei</i>	6	-	-	6
9291 <i>Hydrogenophaga</i> sp.	4	1	-	5
9471 <i>Hydrogenophaga</i> sp.	6	-	-	6
9284 <i>Methylobacterium persicinum</i>	6	-	-	6
9358 <i>Microbacterium azadirachtae</i>	3	-	-	3
9254 <i>Nocardia</i> sp.	3	-	-	3
9246 <i>Pedobacter ruber</i> .	4	1	1	6
9496 <i>Pedobacter</i> sp.	5	1	-	6
9470 <i>Phyllobacterium myrsinacearum</i>	5	-	1	6
9261 <i>Pseudomonas</i> sp.	6	-	-	6
9320 <i>Rheinheimera chironomi</i>	2	2	2	6
9510 <i>Sphingobium limneticum</i>	5	1	-	6
9283 <i>Sphingomonas</i> sp.	6	-	-	6

9271	<i>Sphingopyxis</i> sp.	4	2	-	6
9497	<i>Variovorax</i> sp.	4	2	-	6
9253	<i>Williamsia maris</i>	3	-	-	3
Total		128	21	9	158
Successful measurement rates		81%	13%	6%	

2. CAPITOL 2

Article 2: Monitorització de la dinàmica de les comunitats bacterianes en una planta de tractament d'aigua potable: un enfocament integrador mitjançant metabarcoding i indicadors microbians en grans volums d'aigua

“Monitoring bacterial community dynamics in a drinking water treatment plant: an integrative approach using metabarcoding and microbial indicators in large water volumes”

Anna Pinar-Méndez, Owen S. Wangensteen, Kim Præbel, Javier Méndez, Belén Galofré, Anicet R. Blanch, Cristina Garcia-Aljaro.

Water (2022); 14 (9), 1435, DOI: 10.3390/w14091435

Els recursos hídrics d'aigua dolça són sistemes especialment vulnerables a canvis climatològics. Actualment estem experimentant esdeveniments climàtics més extrems i amb més freqüència, amb pics de temperatura que s'estan accentuant en climes com el Mediterrani, i amb períodes més prolongats de sequera i pluges torrencials, agreujant l'estrès hídric. Un augment de temperatura pot promoure l'eutrofització de les aigües alterant l'equilibri de les comunitats bacterianes, per exemple afavorint la presència de patògens i oportunistes. Aquest escenari, podria impactar en la disponibilitat i qualitat dels recursos hídrics, fet que suposarà un repte per a la producció d'aigua potable. Així doncs, l'ús d'eines de control de qualitat complementàries a les convencionals (legislades) que permetin aprofundir en la totalitat de les comunitats que componen els sistemes d'aigua, són importants per a detectar possibles desequilibris en el sistema que afectin a la bioestabilitat de l'aigua, com per exemple recreixements no desitjats que afectin les propietats organolèptiques o que posin en risc la salut de les persones o la qualitat de l'aigua.

Un paràmetre crític del mostreig és obtenir una mostra significativa per a concentrar el màxim possible la microbiota present a l'aigua, sobre tot la que es podria trobar en concentracions molt baixes. L'ús de filtres DEUF com el Rexeed permeten la concentració de grans volums d'aigua (fins a 1100L) millorant la representativitat de la mostra i recuperant simultàniament bacteris, virus i protozous.

L'objectiu d'aquest estudi ha estat caracteritzar les poblacions de la microbiota de la planta de l'ETAP de Sant Joan Despí en condicions d'operativitat normals de la planta, optimitzant una metodologia per l'anàlisi de mostres amb baixa concentració microbiana, combinant l'ús de

l'anàlisi convencional d'indicadors microbiològics de la qualitat de l'aigua basat en la directiva europea EU 2020/2184, i metabarcoding del gen 16S ARNr per investigar la diversitat, variacions espacio-temporals, identificant quins tractaments influeixen en la composició bacteriana i quins grups resisteixen els processos de tractament a l'aigua final.

Es van prendre un total de 72 mostres de 9 processos de tractament de l'ETAP SJD (i 2 mostres extres d'una eventualitat a la planta) van ser preses al llarg d'un any cobrint un període amb temperatures de 11.5 a 29.5 °C: captació (pous i riu), pretractament (decantació i filtre de sorra), tractament convencional (ozonització i filtre de carbó), tractament avançat (osmosi inversa) i aigua final (cambra de mescla i aigua potable clorada). Les mostres es van concentrar mitjançant la filtració per membrana (riu i pretractament; fins a 2L) o per Rexeed (tractament convencional, avançat i aigua final; de 100-1.100 L). Per a totes les mostres es va procedir a analitzar els paràmetres de qualitat microbiològics legislat (bacteris aerobis heterotròfics (HPC) i els indicadors coliforms totals, *E. coli*, *Clostridium perfringens*, enterococs intestinals i colífags somàtics) i es va extreure el DNA per a l'anàlisi per metabarcoding. Addicionalment, es van enregistrar dades climatològiques i alguns paràmetres fisicoquímics i d'operativitat de la planta. Totes les mostres complien els estàndards de qualitat pel que respecta als resultats microbiològics. Els indicadors van mostrar diferències significatives al llarg dels tractaments, amb una reducció progressiva des de la captació fins a l'aigua final. A la captació del riu es van observar les concentracions més elevades de tot el sistema per a tots els indicadors (HPC: 4,5 \log_{10} CFU/mL; indicadors: 3-4,7 \log_{10} CFU, PFU ó NMP/100mL), i molt superiors a l'aigua subterrània (HPC: 0,04 \log_{10} CFU/mL; indicadors: negatiu). Només amb el pretractament, i respecte l'aigua del riu, es van reduir 3 \log_{10} els HPC, coliforms totals i *E. coli* i 4 log els enterococs i colífags somàtics. La línia de tractament convencional va reduir fins a 4 logaritmes els indicadors, mentre que la línia de el tractament avançat va eliminar completament la presència dels indicadors aconseguint una reducció d'entre 5 i 6 \log_{10} i reduint els valors de HPC al voltant de 1,4 \log_{10} CFU/mL. L'aigua final pre-clorada a la cambra de mescla, va presentar una concentració d'aproximadament 2 \log_{10} CFU/mL per HPC i valors negatius per a tots els indicadors, excepte en algunes mostres on es van detectar valors molt baixos de coliforms totals (0,34 \log_{10} NMP/100mL) i a una mostra enterococs (-0,7 \log_{10} CFU/100mL).

L'ús del mètode de concentració Rexeed va permetre analitzar volums de fins a 1.100 L, permetent la detecció d'indicadors en mostres on històricament havien estat considerades negatives ja que la normativa contempla únicament l'anàlisi d'un volum de 100 mL. Això va permetre estudiar l'eficiència dels diferents tractaments d'una manera més exhaustiva. També va permetre una caracterització més profunda i representativa de les comunitats bacterianes,

ja que es van poder mostrejar grans volums d'aigua, especialment rellevant en mostres amb una baixa biomassa microbiana com l'aigua potable clorada.

En l'anàlisi d'indicadors microbiològics no es van detectar grans canvis significatius a cada tractament, mentre si es van observar canvis en la diversitat i distribució de les comunitats bacterianes per metabarcoding. El total de 9.995.538 *reads* obtinguts es van poder agrupar en 10.039 *Amplicon Sequence Variants* (ASVs), i el seu anàlisi va permetre observar que les diferents etapes de potabilització van exercir una pressió selectiva sobre les comunitats bacterianes, on els processos amb més diversitat corresponien a la captació de pous, la cambra de mescla, filtració per carbó i el riu, i la menys diversa l'aigua potable clorada. Les etapes amb processos de desinfecció (decantació, ozonització i cloració) van mostrar un descens en la diversitat. Per exemple, en comparació amb el riu, l'etapa de decantació va reflectir una pèrdua de riquesa després de la desinfecció amb diòxid de clor, seguit d'un augment de la diversitat després de la filtració per sorra, fet que podria indicar un possible recreixement bacterià en els filtres. Els procediments de filtració van tenir una forta influència amb una major diversitat observada a la filtració per sorra i filtració per carbó, que podria indicar un microbioma establert associat amb els procediments de filtració com s'ha vist en altres estudis. Tanmateix, la desinfecció va tenir un impacte més gran, ja que els desinfectants residuals redueixen l'abundància i la diversitat de comunitats bacterianes, que esdevenen dominades per gèneres resistents a l'aigua posterior a la cloració.

La diversitat beta va indicar que les mostres de rius i pous eren clarament diferents presentant ASVs característics, així com l'aigua potable, i aquestes tres també s'allunyaven de la resta de tractaments, on es compartien més ASVs. No obstant, l'estudi de les seqüències o taxons compartits entre les diferents mostres (*core*) van determinar que no hi havia ASVs compartits a totes les mostres de totes les etapes de tractament.

La classificació taxonòmica dels ASVs va permetre observar que cada etapa presentava una microbiota específica. Els fílums més abundants al llarg de tots els tractaments van correspondre a Proteobacteria (57-73%), seguit de Bacteroidota (0.8-27%) i Bdellovibrionota (1-12%), excepte després de la cloració on el fílum majoritari passa a ser Cyanobacteria (56%) a l'aigua potable clorada. L'entorn afòtic dels processos de tractament suggereix que els cianobacteris detectats podrien ser no fotosintètics; tampoc es van detectar ni estancament ni microcistines (toxines cianobacterianes) al riu ni a l'aigua potable. Alguns dels gèneres més abundants en les diferents etapes van ser *Flavobacterium*, *Limnohabitans* i família Obscuribacteraceae.

Es va observar una influència de la temperatura ambiental respecte la distribució bacteriana a les diferents etapes del tractament (excepte en els pous on no es va observar diferència), amb una major diversitat en temperatures altes, que van ser significatives a decantació i filtració per

carbó, i moderada a la resta del tractament. Una temperatura més alta pot provocar canvis en les comunitats bacterianes ja que facilita el potencial de creixement bacterià, i podria derivar en l'aparició de certs patògens bacterians transmesos per l'aigua, patògens oportunistes i grups formadors de biofilm que poden comprometre la qualitat de l'aigua o suposar un risc per a la salut pública.

Les seqüències obtingudes es van analitzar per buscar a totes les mostres la presència de possibles patògens de transmissió hídrics així com els indicadors. Tot i representar abundàncies relatives baixes (0,002-8%), es va detectar la presència d'alguns gèneres patògens oportunistes com *Aeromonas*, *Arcobacter*, *Legionella*, *Pseudomonas* o *Mycobacterium*, no obstant es van reduir o eliminar a l'aigua potable final excepte *Aeromonas* i *Pseudomonas*. Els indicadors van presentar abundàncies relatives molt més baixes en comparació amb els resultats obtinguts als paràmetres microbiològics cultivables, però van ser detectats a etapes més avançades on els resultats de cultivables eren negatius. No obstant, per la sensibilitat de la tècnica, només es van poder analitzar els gèneres, i per tant cal esmentar que no totes les seves espècies són patògens animals, humans o vegetals, i poden correspondre a cèl·lules no viables o mortes després del tractament de potabilització, o a ADN lliure. Tanmateix, les dades d'abundància relativa s'han de tractar amb precaució, ja que la seqüenciació d'alt rendiment no està dissenyada amb finalitats quantitatives, i pot conduir a avaluacions inexactes de la qualitat de l'aigua. No obstant això, permet el seguiment de cultivables i no cultivables, patògens, oportunistes i indicadors microbians, observant les seves fluctuacions entre tractaments.

D'altra banda, una fuga de salmorra riu amunt (derivat de l'activitat de les mines de sal) va produir una eventualitat a la planta que va canviar temporalment el procediment habitual del tractament de l'aigua, aturant la captació superficial i funcionant exclusivament amb aigua de pous, filtre de carbó i cloració. L'estudi d'una mostra de pous i filtre de carbó, que presentaven valors normals dels paràmetres microbiològics, van revelar una estabilitat en les comunitats bacterianes dels pous però un augment de determinades poblacions bacterianes en el filtre de carbó en comparació amb els mostreigs programats de la planta, que es podria atribuir al creixement de les comunitats presents en els filtres o el seu despreniment temporal de les superfícies de carboni pel canvi sobtat de la font captació d'aigua.

Si bé els indicadors microbians i l'HPC ofereixen protecció contra patògens garantint la qualitat de l'aigua i poden detectar microorganismes viables, només s'observa la fracció cultivable i els resultats poden estar esbiaixats per les condicions de creixement. En canvi, amb metabarcoding es disposa d'una anàlisi més profunda de les comunitats microbianes (cultivables, no cultivables, autòtrofs, heteròtrofs, patògens, etc.), però com a inconvenients, no diferencia viabilitat, amb el que dificulta l'avaluació de riscos i la obtenció dels resultats és molt més lenta.

Com a conclusions principals, l'anàlisi dels paràmetres microbiològics per a l'avaluació de la qualitat l'aigua han mostrat que l'aigua potable reunia els estàndards de qualitat. No obstant, aquests anàlisis podrien no ser suficients per detectar variacions en les comunitats, i és que la normativa exigeix que no hi hagi canvis anòmals en la concentració d'HPC i que en aquest estudi s'han observat canvis en diversitat de les poblacions tot i no haver canvis significatius en els HPC, posant de manifest la limitació dels HPC com a paràmetre indicador de la qualitat de l'aigua. La diversitat observada suggereix que cada etapa de tractament té una microbiota específica, influenciada pel tipus de tractament aplicat (especialment procediments de filtració i cloració) i la temperatura ambient. L'ETAP presenta a Proteobacteria com a fílum dominant durant tot el procediment, amb un gran canvi després de l'etapa de cloració, desplaçat per Cyanobacteria. Per tant, l'ús combinat de paràmetres estandarditzats amb *metabarcoding* del gen 16S ARNr proporciona la imatge més completa de les comunitats de l'ETAP, que podria ser útil per a la detecció de variacions operatives i el seu impacte en la qualitat de l'aigua.



Article

Monitoring Bacterial Community Dynamics in a Drinking Water Treatment Plant: An Integrative Approach Using Metabarcoding and Microbial Indicators in Large Water Volumes

Anna Pinar-Méndez ^{1,2,3,*}, Owen S. Wangensteen ⁴, Kim Præbel ⁴, Belén Galofré ¹, Javier Méndez ², Anicet R. Blanch ^{2,3} and Cristina García-Aljaro ^{2,3,*}

¹ Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, 08028 Barcelona, Spain; bgalofre@aiguesdebarcelona.cat

² Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, 08028 Barcelona, Spain; jmendez@ub.edu (J.M.); ablanch@ub.edu (A.R.B.)

³ The Water Research Institute, University of Barcelona, 08001 Barcelona, Spain

⁴ Norwegian College of Fishery Science, Faculty of Biosciences, Fisheries and Economics, UiT The Arctic University of Norway, N-9037 Tromsø, Norway; owen.wangensteen@uit.no (O.S.W.); kim.praebel@uit.no (K.P.)

* Correspondence: apinar@aiguesdebarcelona.cat (A.P.-M.); crgarcia@ub.edu (C.G.-A.); Tel.: +34-93-247-98-12 (A.P.-M.); +34-93-402-14-84 (C.G.-A.)



Citation: Pinar-Méndez, A.; Wangensteen, O.S.; Præbel, K.; Galofré, B.; Méndez, J.; Blanch, A.R.; García-Aljaro, C. Monitoring Bacterial Community Dynamics in a Drinking Water Treatment Plant: An Integrative Approach Using Metabarcoding and Microbial Indicators in Large Water Volumes. *Water* **2022**, *14*, 1435. <https://doi.org/10.3390/w14091435>

Academic Editor: Anas Ghadouani

Received: 17 February 2022

Accepted: 28 April 2022

Published: 30 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Monitoring bacterial communities in a drinking water treatment plant (DWTP) may help to understand their regular operations. Bacterial community dynamics in an advanced full-scale DWTP were analyzed by 16S rRNA metabarcoding, and microbial water quality indicators were determined at nine different stages of potabilization: river water and groundwater intake, decantation, sand filtration, ozonation, carbon filtration, reverse osmosis, mixing chamber and post-chlorination drinking water. The microbial content of large water volumes (up to 1100 L) was concentrated by hollow fiber ultrafiltration. Around 10 million reads were obtained and grouped into 10,039 amplicon sequence variants. Metabarcoding analysis showed high bacterial diversity at all treatment stages and above all in groundwater intake, followed by carbon filtration and mixing chamber samples. Shifts in bacterial communities occurred downstream of ozonation, carbon filtration, and, more drastically, chlorination. Proteobacteria and Bacteroidota predominated in river water and throughout the process, but in the final drinking water, the strong selective pressure of chlorination reduced diversity and was clearly dominated by Cyanobacteria. Significant seasonal variation in species distribution was observed in decantation and carbon filtration samples. Some amplicon sequence variants related to potentially pathogenic genera were found in the DWTP. However, they were either not detected in the final water or in very low abundance (<2%), and all EU Directive quality standards were fully met. A combination of culture and high-throughput sequencing techniques may help DWTP managers to detect shifts in microbiome, allowing for a more in-depth assessment of operational performance.

Keywords: water quality monitoring; drinking water treatment plant; drinking water; 16S rRNA metabarcoding; amplicon sequencing; hollow fiber ultrafiltration; bacterial diversity

1. Introduction

Freshwater resources are particularly vulnerable to the consequences of climate change. A general increase in temperature can promote eutrophication of surface waters [1,2], whereas more frequent extreme weather events, such as heavy rainstorms and longer drought periods, may impact the availability of safe drinking water, especially in water-stressed regions such as the Mediterranean. Temperature is a crucial determinant of water quality, since it influences microbial dynamics such as uncontrolled proliferation that can

facilitate biofilm formation [3,4]. In a global warming scenario where water temperature is expected to rise, a potential change in microbial biodiversity can occur and can involve the emergence of some opportunistic bacteria in drinking water [3].

Pressures on water resources and quality are further exacerbated by the growing concentration of human populations in metropolitan areas and intensification of industrial activities [5]. In such a scenario, the provision of sufficient safe water for all citizens is an increasing challenge for water production management at drinking water treatment plants (DWTP).

The aim of a DWTP is to bring source water up to established drinking water quality standards by removing organic materials and pathogens and by reducing the total bacteria, with or without residual disinfectants. Current regulations to guarantee the safety of water for human consumption are still focused on monitoring culturable fecal indicators and heterotrophic bacteria plate counts (HPC). Although the European Directive (EU) 2020/2184 states that drinking water must be free of the fecal indicators *Escherichia coli* and intestinal enterococci, it does not stipulate upper limits for HPC but rather only the absence of abnormal changes; thus, each member state is responsible for defining a threshold for what constitutes an acceptable level of change. That minimum change in microbial water quality is defined as biological stability [6], and the aim is to provide drinking water to consumers with the same quality achieved in DWTPs. However, unwanted changes may occur during operational procedures or in distribution networks, and studies have shown that, during the potabilization process, water with a stable HPC may have a variable microbial composition and abundance that is undetectable by conventional methods [7–9]. Moreover, the water microbiome may contain opportunistic pathogens and other underexplored diversity. Therefore, alternative approaches are needed for a better understanding of microbiomes in water systems to improve water management from source to tap.

High-throughput sequencing techniques, such as 16S rRNA metabarcoding, have the potential to provide in-depth information that complements standard bacterial quality parameters, and they can help to generate a more accurate picture of microbial communities at different water treatment stages. Several water microbiome studies based on 16S rRNA amplicon sequencing have described microbial composition in DWTPs and distribution networks, which differs according to the water source (e.g., river [10–14], lake [15], groundwater [13,16] or seawater [17]).

A critical parameter in water quality monitoring is the sampling procedure, which needs to ensure the integrity and representativeness of the sampled water. Analyzing large volumes can be an effective approach to detect low-concentration targets. In recent years, dead-end ultrafiltration (DEUF) using hollow fiber filters has emerged as an effective sampling technique based on size exclusion, allowing for good simultaneous recoveries of bacteria, viruses and protozoa from large volumes of water (up to 1100 L) [18–20]. Advantages of using DEUF methods include an all-in-one concentrating procedure for microbes; an improved sample representativity (in comparison with conventional small-volume concentration methods); and the ability to detect low numbers of microbes, including biothreat agents [21], waterborne pathogens [19,22,23] or indicators [24] in large-volume water samples without filter clogging. However, very few studies have applied this method to study microbial biodiversity in the field of water, and they have been mainly focused on distribution networks [8,25,26]. To the best of our knowledge, this is the first study of a full-scale DWTP using 16S rRNA metabarcoding with hollow fiber ultrafiltration methods.

The main goal of this work is to assess the effects of different treatment stages on bacterial communities in a full-scale DWTP. By combining conventional microbial water quality analysis based on the European Directive (EU) 2020/2184 with 16S rRNA metabarcoding, we hope to achieve a more in-depth description of the microbial baseline in the DWTP to evaluate process performance and for the detection of process instability.

2. Materials and Methods

2.1. Sample Collection

The study was carried out in a full-scale DWTP in Sant Joan Despí, Barcelona (Catalonia, northeastern Spain) between 2018 and 2019. This large supply system, located in the lower Llobregat River Basin, has an average production capacity of $\sim 345,000 \text{ m}^3$ a day, which provides drinking water to more than 3.2 million inhabitants in the Barcelona metropolitan area. The complexity of the treatment applied is determined by distinctive features, such as scarcity of water resources ($22 \text{ m}^3/\text{s}$ average flow rate), irregular Mediterranean-type rainfall (from drought to flood) and the influence of industrial and agriculture activities upstream, including potash mining and wastewater treatment plants, on raw water quality (Figure 1). In summary, the plant has two sources of water (river and groundwater). Surface water from the Llobregat River is pumped into the DWTP and cleaned of sand and gravel, before undergoing sequential treatments. The pretreatment stages consist of coagulation/flocculation of organic matter and an initial disinfection with chlorine dioxide, followed by decantation to eliminate flocculants. Subsequently, sand filtration is performed to retain smaller particles held in suspension. The pretreated water then converges with groundwater intake, and both are pumped into two independent treatment lines: one based on conventional processes of ozonization and filtration by granular activated carbon, and the other based on advanced treatments, namely, ultrafiltration, reverse osmosis and remineralization. The water from both lines is mixed in a chamber and stabilized, and then it is treated with chlorine ($0.5\text{--}1.5 \text{ mg/L}$ residual chlorine) in a separated tank before being pumped into the distribution network.

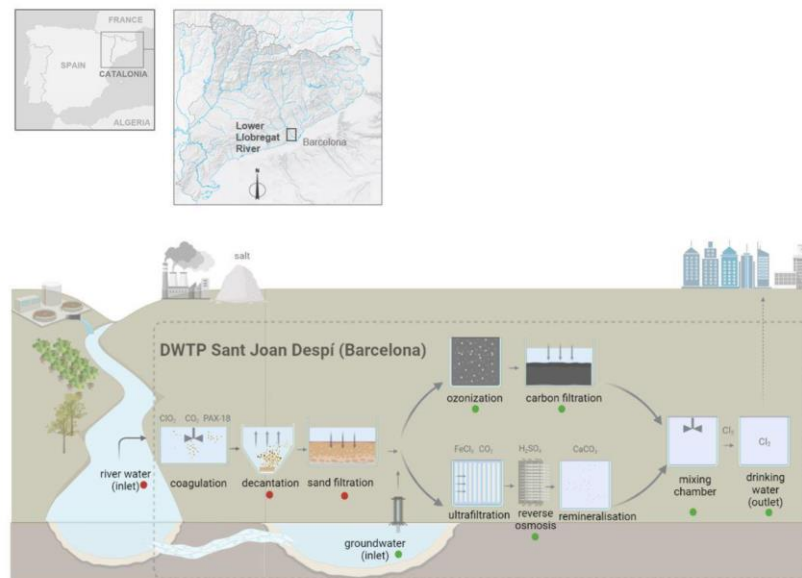


Figure 1. Study site in the lower Llobregat River and schematic diagram of DWTP stages in Sant Joan Despí (Barcelona Spain). Sites of sample collection for this study are indicated with a red spot (high microbial load) and green spot (lower microbial load). Physical map of Catalan hydrographic basins was made by Institut Cartogràfic and Geològic de Catalunya (ICGC).

Eight sampling campaigns were carried out over one year, covering seasons of high and low temperatures and avoiding big rainfall episodes, as the aim was to determine bacterial populations under usual DWTP operative conditions.

A total of 72 water samples were collected at nine different stages of treatment of the DWTP (Figure 1): river water (RW, n = 8) and groundwater (GW, n = 8) intake, decantation (DEC, n = 8), sand filtration (SF, n = 8), ozonization (OZ, n = 8), carbon filtration (CF, n = 8), reverse osmosis (RO, n = 8), mixed chamber (MIX, n = 8) and post-chlorinated water (DW, n = 8). Sampling sites were classified into two groups according to microbial load: river water and pretreatment stages (RW, DEC, SF) (high microbial load), and groundwater and conventional/advanced treatments (GW, OZ, CF, RO, MIX, DW) (lower microbial load). Sampling and processing conditions were different for each group, as described below.

For the first group (pretreatment stages), 2 L of water was collected in polyethylene sterile bottles with sodium thiosulfate (24 mg/L). For the second group (treatment stages), higher volumes were sampled (from 100 to 1100 L/sample) by dead-end hollow fiber ultrafiltration (DEUF) [18–20] using Rexeed™ 25-A filters (Asahi Kasei Medical Co, Tokyo, Japan) of 30 kDa size exclusion. The sampling procedure took several hours (2 to 8.5 h). Chlorinated samples were pretreated with sodium thiosulfate (24 mg/L). All samples were transported to the laboratory at 4 °C for further analyses within 24 h.

Surface Water Contamination Episode and Changes in Operational Procedures

One additional sampling campaign was carried out during a contamination episode affecting the Llobregat River, which forced the DWTP to change its routine operational conditions. The contamination was due to the leakage of brine from a 120 km-long collector that runs parallel to the river and discharges brine produced by potash mining upstream into the sea (for more information about the collector, see Martin-Alonso [27]). To prevent this pollution spill from entering the DWTP, the RW intake was stopped, and the DWTP continued operating only with GW, which does not undergo any of the pre-treatments. The water was treated with CF and chlorination but not with OZ, which was stopped to reduce subproduct bromate formation, nor with RO, as this process produces brine and is discharged into the same collector, which needed to be empty for the required repair. Sampling was carried out at two sites, GW (n = 1) and CF (n = 1), four days after the DWTP had changed to these operating conditions. Microbial water quality parameters were also analyzed.

2.2. Sample Processing and Nucleic Acid Extraction

For the first group (pretreatment stages), different volumes of water were filtered through 0.22 µm polycarbonate membranes (Millipore, Molsheim, Alsace, France): RW (250 mL), DEC (500 mL–750 mL) and SF (750 mL–1 L). Each filter was placed face down in a sterile flask containing 5 mL of Ringer's solution 1:4 and was eluted by sonication (40 kHz) for 4 min in a Branson Ultrasonics CPX 3800 h ultrasonic water bath (Fischer Scientific, Madrid, Spain), keeping the sample on ice to avoid overheating, and then the sample was vortexed for 1 min. The resulting eluate (ca. 5 mL), containing biomass detached from the membrane, was used for genomic DNA extraction. The remaining original sample was used to determine culturable microbial indicators.

For the second group (treatment stages), water samples were concentrated by Rexeed filters and were eluted in the backflush mode using a peristaltic pump at 1 L/min and 500 mL of phosphate-buffered saline solution (0.5 mL of 1% Antifoam (Sigma-Aldrich, Saint Louis, MO, USA), 10% Tween 80 (Scharlab, Sentmenat, Barcelona, Spain) and 0.5 mL of 10% NaPP (Sigma-Aldrich, Saint Louis, MO, USA)). The Rexeed eluate volume of 500–650 mL was divided to be used for different purposes: a total of 120–170 mL was kept for the culture-based analysis, and the remaining volume (or until filter clogging) was further reduced by centrifugation at 3500 × g using Centricon®70 plus (Merck, Darmstadt Germany) to a 2–4 mL concentrate for nucleic acid extraction.

DNA from samples was extracted using the automated MagNA Pure LC DNA Isolation Kit III (Bacteria, Fungi) (Roche Diagnostics GmbH, Mannheim, Germany), as described below. To obtain a maximum DNA yield, pre-isolation and external lysis steps were applied prior to extraction to reduce eluate volumes. Sample eluates (2 mL) were centrifuged for 10 min at $8000\times g$ at $4\text{ }^{\circ}\text{C}$, and the supernatants were discarded. After the addition of a premixed lysis buffer (130 μL) (Roche Diagnostics GmbH, Germany) and proteinase K (20 μL) (Roche Diagnostics GmbH, Germany), the sample was incubated for 10 min at $65\text{ }^{\circ}\text{C}$. To enhance cell lysis and inactivate potentially pathogenic organisms in the sample, a boiling step was performed, in which the samples were heated at $95\text{ }^{\circ}\text{C}$ for 10 min and then cooled on ice for 1 min.

For sample eluates with high viscosity or large pellets, additional steps were carried out, as recommended by the manufacturer. In summary, mechanical homogenization was performed to enhance lysis of bacterial cells, in which up to 0.40 g of 0.5 mm-diameter silica glass beads (Sigma-Aldrich, Saint Louis, MO, USA) were transferred to the sample tubes. Samples were homogenized by vortexing for 30 s and cooled for 1 min in a cooling block. The samples were then centrifuged for 5 min at $17,000\times g$, and the supernatant was placed in an Eppendorf tube to undergo the same pre-isolation steps as the other samples described above, except for the first centrifugation.

After the pre-isolation protocol, samples were transferred into the sample cartridge for automatic nucleic acid extraction. DNA quality was verified by visualization on 0.8% agarose electrophoresis gel, and the concentration was measured with a Qubit dsDNA HS Assay Kit on a Qubit 3.0 Fluorometer (Invitrogen, Waltham, MA, USA).

2.3. Microbial Water Quality Analysis

All samples were processed within 12 h of collection and were processed for microbial water quality according to the European Directive (EU) 2020/2184 [28], which is based on the analysis of culturable bacterial indicators and heterotrophic bacteria.

The heterotrophic bacteria were cultured by mass inoculation according to the ISO Standard 6222:1999 [29]. Different sample volumes (0.001–1 mL) were inoculated in duplicate in ISO Water Plate Count Agar (WPCA) (Oxoid, Basingstoke, UK) at $22\pm 2\text{ }^{\circ}\text{C}$ for $72\text{ h}\pm 3\text{ h}$.

The enumeration of culturable total coliform bacteria and *E. coli* based on the ISO Standard 9308-2:2012 [30] was performed simultaneously using a Colilert 18[®] test kit and a Quanti-Tray[®]/2000 (Idexx, Westbrook, ME, USA) with 100 mL sample volumes. Only for river water, samples were diluted 1/10 and 1/100. After being transferred to the Quanti-Tray/2000 blister-pack, the samples were incubated at $36\pm 2\text{ }^{\circ}\text{C}$ for $18\pm 4\text{ h}$.

Vegetative cells and spores of *Clostridium perfringens* were enumerated using membrane filtration and an internal validation method based on UNE-EN ISO 7937 (2005) [31] and UNE-EN 26461-2 (1995) [32]. Different sample volumes (100, 10, 1 and 0.1 mL) were filtered through a 0.22 μm -pore-size mixed cellulose ester membrane filter (Millipore, Molsheim, France). The filter was placed face up on tryptose sulfite cycloserine agar (Oxoid, Basingstoke, UK), with D-Cycloserine selective supplement and 4-methylumbelliferylphosphate (Merck, Darmstadt, Germany), and it was anaerobically incubated at $44\pm 1\text{ }^{\circ}\text{C}$ for $21\pm 3\text{ h}$ inside a jar with anaerobic gas generating sachets Genbox anaer (BioMérieux, Marcy l'Etoile, France) and an Anaerotest[™] anaerobic indicator strip (Merk, Darmstadt, Germany). To enumerate *C. perfringens*, plates were examined under long-wave UV light, and those presenting fluorescence were counted as positive results.

The detection and enumeration of intestinal enterococci were carried out using a membrane filtration method based on the ISO Standard 7899-2:2000 [33]. Serial dilutions of water samples (100, 10, 1 and 0.1 mL) were filtered through a 0.45 μm -pore-size mixed cellulose ester filter (Millipore, Darmstadt, Germany). The filter was placed on Slanetz and Bartley agar (Merck, Darmstadt, Germany) and incubated at $36\text{ }^{\circ}\text{C}\pm 2\text{ }^{\circ}\text{C}$ for $44\pm 4\text{ h}$. Colonies with red pigmentation were considered presumptive and were confirmed as

intestinal enterococci if they turned brown or black after incubation at 44 °C for 4 h on bile esculin azide agar (Merck, Darmstadt, Germany).

Somatic coliphages (SOMCPH) were concentrated and enumerated by a double-layer agar method adapted from ISO 10705-2:2000 [34] and ISO 10705-3:2003 [35,36]. Samples were processed by direct inoculation or membrane filtration depending on the matrix. For RW samples, 0.1 mL and 1 mL were analyzed by direct inoculation. For the first group of samples (pretreatment stages), membrane filtration of 500–1000 mL was needed, whereas for the second group (GW and treatment stages), volumes ranging from 6 to 10 mL were processed by direct inoculation, as they corresponded to eluates concentrated by REXEED.

2.4. Physicochemical and Climatological Data

A range of physicochemical data was measured *in situ* during sampling: sample temperature (°C), conductivity ($\mu\text{S}/\text{cm}$), turbidity (NTU), pH, ammonium ($\text{mg NH}_4^+/\text{L}$), total organic carbon (TOC: $\text{mg C}/\text{L}$), percentages of RW and GW intake, river flow rate (m^3/s) and residual amounts of the disinfectants chlorine dioxide, ozone and chlorine (ppm).

Climatological data were also recorded in the Sant Joan Despí DWTP meteorological station and retrieved from the Meteorological Service of Catalonia, including temperature (°C), relative humidity (%) and rainfall ($\text{mm}/24 \text{ h}$).

2.5. Metabarcoding by 16S rRNA Gene Amplicon Sequencing

To characterize bacterial diversity in the water ecosystem, metabarcoding was performed by paired-end amplicon sequencing. Two-run metabarcoding was applied to a total of 104 samples: 72 DWTP samples, 2 event samples, 8 additional samples used to check reproducibility between the different runs, 16 whole-procedure blanks (from each sampling to sequencing), 4 negative controls (from PCR to sequencing), and 2 positive controls (ZymoBIOMICS™ microbial community DNA standard (mock community) (Zymo research, Irvine, CA, USA)). A total of 6 DWTP samples were removed from the dataset, as they presented low reads and were sequenced in a third run.

2.5.1. PCR Amplification of DNA

The 16S rRNA V4 genomic region was amplified by PCR using barcoded primers 515F and 806R [37]. A detailed list containing the sequences of primers, sample barcodes and leading Ns can be found in Table S1. For all samples (5–10 ng total DNA), PCR was performed using 10 μL of AmpliTaq gold 360 Master Mix (Applied Biosystems, Bedford, MA, USA), 0.16 μL of bovine serum albumin, 1.84 or 5.84 μL of nuclease-free water, 2 μL of primer mix (4 μL of each 5 μM forward and reverse primer and 72 μL water) and 2 or 6 μL of sample-extracted DNA in a total volume of 20 μL per sample. PCR cycling conditions followed the protocol of Caporaso [37]. Briefly, the reaction started with a denaturation procedure held at 95 °C for 3 min. Then, 35 amplification cycles were set at 94 °C for 45 s, 50 °C for 60 s and 72 °C for 90 s. A final extension of 10 min at 72 °C was added. All samples were run with negative and positive controls of PCR blanks and mock communities, respectively. The PCR products were checked by visualization on 1% agarose gel containing purple 6x loading dye and by a reference 1 Kb DNA extension ladder marker (Invitrogen, Waltham, MA, USA).

2.5.2. Amplicon Multiplexing, Library Preparation and Illumina Sequencing

Multiplexing of samples was performed by pooling 10 μL of each barcoded amplicon. Library preparation consisted of PCR amplicon clean-up by the MinElute PCR Purification Kit (Qiagen, Hilden, Germany), and DNA quantification consisted of purified PCR products by the Qubit dsDNA BR Assay Kit (Invitrogen, Waltham, MA, USA). The DNA PCR-free protocol was used to prepare a multiplexed paired-end library for Illumina sequencing by the NEXTflex PCR-Free DNA Sequencing Kit (Bioo Scientific, Austin, TX, USA). The library was quantified by qPCR using the NEBNext Library Quant Kit (Illumina, San Diego, CA, USA), and the NEBioCalculator web tool was used for qPCR NGS library quantification

(<https://nebiocalculator.neb.com/>, accessed on 26 July 2019) to achieve optimal cluster densities and therefore an optimal sequence output. Multiplexed libraries with barcoded samples were spiked with 1% bacteriophage PhiX genome (PhiXControl Library; Illumina, San Diego, CA, USA) and were sequenced by the Illumina MiSeq platform using the 600-cycle 2 × 300 bp paired-end sequencing MiSeq Reagent Kit v3 (Illumina, San Diego, CA, USA) with a loading concentration of 20 pM.

2.5.3. Bioinformatic Analyses

The resulting FASTQ files were checked for quality (>90% of bases with a Phred quality score >Q30), processed and analyzed using Cutadapt [38], DADA 2 [39], QIIME2 [40] and Phyloseq [41] pipelines. Cutadapt was used to demultiplex samples (assign sequences to their sample) and to trim adapters, primers, barcodes and leading Ns from sequencing reads. A DADA 2 workflow was performed to denoise the sequences.

Firstly, quality filtering and the trimming of sequences was set to 200 bp (for forward reads) and 190 bp (for reverse reads) with a maximum number of expected errors allowed per read set at 1 (EE = 1). Then, samples with zero reads and low-quality sequences were deleted. Filtered sequences were dereplicated, the forward and reverse reads were aligned and merged, chimeras were removed and an amplicon sequence variant (ASV) table was obtained. Taxonomy was assigned to the resulting sequence variants with a pre-trained Naive Bayes classifier [42] using the SILVA SSU 138 reference database and was imported to the phyloseq R package for microbiome analyses.

To obtain more accurate profiling of microbial communities, the decontam R package [43] was used to remove sequences derived from contaminating DNA present in extraction or sequencing reagents that may interfere with samples.

2.6. Data Analyses

Statistical analyses were performed using the statistical software R version 4.0.5 (<https://www.r-project.org>, accessed on November 2021). For water quality, data were $\log_{10}(x + 1)$ transformed, and the Kruskal–Wallis test was used to assess differences in culturable and molecular parameters throughout the DWTP.

The phyloseq package [41] was used to analyze and visualize sequencing data.

Biodiversity was evaluated by analyzing the relative abundance of reads with the vegan package. Chao1 and Shannon indices were used to study alpha diversity. Beta diversity was also explored, representing sample groups (treatment stages) in a non-metric multidimensional scaling (NMDS) matrix by Bray–Curtis dissimilarities. Temperature impact was measured with both alpha (Pearson correlation) and beta diversity (PERMANOVA analysis). SIMPER analyses were also performed by the vegan package to identify which amplicon sequence variants contributed most to the differences between the DWTP stages. SourceTracker2 [44] was used to estimate the influence of the water source and treatment stages on the bacterial communities downstream.

3. Results

3.1. Sample Collection, and Physicochemical and Climatological Data

Climatological data as well as basic DWTP operating conditions are summarized in Table S2. Temperatures ranged from 11.5 to 29.5 °C, and humidity ranged from 44 to 71%. Episodes of rainfall 24 h before sampling were reported in 2 out of 8 samplings (N1 and N8), but the amounts were small (0.4 and 2.9 mm, respectively). When moderate or heavy rainfall occurred, sampling was postponed to avoid large weather-induced differences between samples. River flow rate ranged between 5.39 and 135 m³/s.

Some operational data were also measured: the percentage of RW or GW intake and the percentage of water processed by conventional or advanced treatments (Table S2). In 2 out of 8 samplings, RW accounted for almost 100% of water intake for the treatment procedure, whereas both RW and GW were used for the rest of the sampling.

Physicochemical data recorded for samples are synthesized in Table S3. Regarding temperature, RW was more variable than GW, averaging at 15 °C (5.6–23.4 °C). In contrast, GW presented a more stable temperature throughout the sampling campaign, with an average of 18 °C (16.1–20.2 °C). The average temperature of the final drinking water was 18 °C (8.5–25.3 °C).

3.2. Culturable Water Quality Indicators

The quality of all water samples taken from the DWTP was examined by determining culturable fecal indicators, as shown in Figure 2. The data obtained follow a non-normal distribution, and the water quality varied at each sampling point. The Kruskal–Wallis test was significant for each culturable indicator at all DWTP stages ($p < 0.00001$).

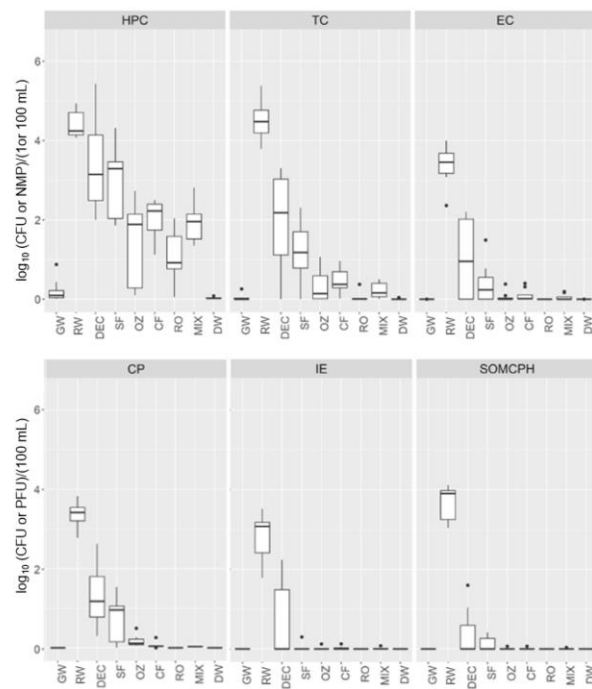


Figure 2. Boxplot of microbial analysis of water samples collected at different treatment stages in the DWTP: heterotrophic plate counts (HPC, CFU/mL), total coliforms (TC, NMP/100 mL), *E. coli* (EC, NMP/100 mL), *Clostridium perfringens* (CP, CFU/100 mL), intestinal enterococci (IE, CFU/100 mL) and somatic coliphages (SOMCPH, PFU/100 mL). Samples represented: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW).

RW samples had the highest concentration of bacterial indicators (average CFU/100 mL or NMP/100 mL: 3 \log_{10} intestinal enterococci (IE), 3.5 \log_{10} *Clostridium perfringens* (CP), 3.6 \log_{10} *E. coli* (EC) and 4.7 \log_{10} total coliforms (TC)), as well as the highest concentration of viral indicator somatic coliphages (SOMCPH; average 3.8 \log_{10} PFU/100 mL) and HPC (average 4.5 \log_{10} CFU/mL), whereas for GW, all water quality parameters were negative except for the HPC, with very few colonies (<7 CFU/mL) observed in some samples.

The dynamics of microbial indicators varied throughout DWTP processing. Samples from the pretreatment stages had a high HPC, especially at DEC (average 4.58 log₁₀ CFU/mL), being similar to those of the RW intake. After the SF stage, the HPC was reduced by an average of 1 log with respect to the DEC stage. Subsequently, a 2 and 3 log reduction in HPC was observed after the conventional and advanced treatments, respectively.

In addition, DEC and SF samples presented a 2 and 3 log reduction in TC, EC and CP, a 1.5 and 4 log reduction in IE and a 3 and 4 log reduction in SOMCPH, respectively. However, although TC decreased after DEC and SF and were further reduced by CF, their complete removal required RO. Thus, TC reduction was 4 log by conventional treatment and 5 log by advanced processing. The same trends were observed for the other indicators (EC, CP and IE), with 4 log reductions achieved by conventional treatment and 5 to 6 log removal by advanced treatment.

In the MIX samples, the HPC was similar to that of CF (~2 log₁₀ CFU/mL); TC levels were very low, ranging from 0.03 to 2.2 NMP/100 mL; and IE was detected in one sample (0.2 CFU/100 mL).

The DEUF method revealed the presence of microbial indicators below the detection limit of the standard ISO methods (<1 CFU/100 mL), thereby allowing a more accurate determination of the efficiency of the different treatments. Thus, positive results for *C. perfringens* were obtained until CF, whereas TC, EC, IE and SOMCPH were detected further downstream until the MIX, with negative results in the post-chlorination water. All microbial indicators after chlorination were below the limits established by European drinking water regulations, with the HPC revealing no or very few colonies (<0.2 CFU/mL).

3.3. Metabarcoding

An overview of the metabarcoding results according to sequencing depth and diversity at all treatment stages is shown in Table 1.

Table 1. Overview of the 16S DWTP sample dataset showing the number of raw reads, filtered reads, total amplicon sequence variants (ASVs), number of core ASVs and mean and standard deviation values (mean ± sd) of alpha diversity for richness (Chao1) and diversity (Shannon).

DWTP Stages	Sampling		Metabarcoding			α-Diversity	
	No. Samples	Raw Reads (Input)	Filtered Reads (Output)	Total ASVs	Core ASVs	Richness (Chao1)	Diversity (Shannon)
GW	8	1,658,112	831,634	2320	43	731 ± 143	5.17 ± 0.18
RW	8	1,891,711	1,048,230	1873	48	525 ± 241	4.65 ± 0.76
DEC	8	1,793,904	1,060,600	2244	3	448 ± 182	4.07 ± 0.46
SF	8	2,206,287	1,321,279	2359	31	561 ± 242	4.19 ± 0.50
OZ	8	1,646,557	953,977	1942	4	378 ± 133	3.66 ± 0.51
CF	8	2,212,847	1,335,303	2416	81	691 ± 258	4.87 ± 0.67
RO	8	1,501,907	920,979	1736	31	432 ± 169	4.32 ± 0.42
MIX	8	1,694,781	1,045,001	2146	51	583 ± 282	4.80 ± 0.75
DW	8	2,325,282	1,478,535	1417	20	314 ± 108	3.34 ± 0.68
Total	72	16,931,388	9,995,538	10,039			

After carrying out quality filtering, the 16S dataset of all DWTP samples comprised a total of 9,995,538 reads with an average of ~138,826 reads per sample (ranging from 40,959 to 274,999 reads). Negative controls ranged from 1627 to 6627 reads. Overall, the number of ASVs obtained by denoising was 10,039.

3.4. Diversity Analysis

The microbial diversity of the 72 water samples was investigated by ASV analysis. After rarefaction at the minimum number of reads obtained in the dataset (40,959), alpha and beta diversity were studied. The rarefaction curves show that all samples reached a plateau, suggesting good representation of bacterial communities (Figure S1).

3.4.1. Alpha Diversity Patterns

Species richness and diversity were examined by analyzing the average relative abundances of ASVs using the species richness estimate (Chao1) and the Shannon diversity index for each sample (Table 1 and Figure 3). Different trends in richness and diversity were observed according to the processing stage of the DWTP.

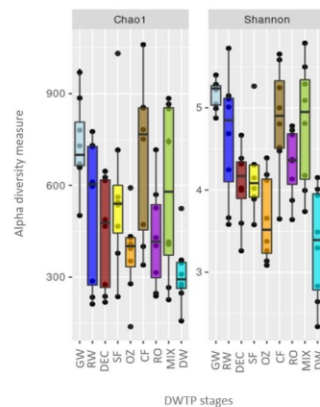


Figure 3. Boxplots representing alpha diversity indices of ASV richness (Chao1) and diversity (Shannon) for all DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW).

The Chao index was higher for GW (731) than for RW (525). In pretreatment, a decrease in richness was observed at DEC (448), followed by a slight increase after SF (561). During the conventional treatment, a decrease was observed after OZ (378), followed by a huge increase after the subsequent CF (691), which produced the second highest richness value overall. In the advanced treatment, RO led to a considerable reduction in richness, producing the third lowest value (432). In the mixing chamber, richness increased again (583), only to be significantly reduced by the final chlorination (314). Thus, the lowest number of ASVs were obtained after OZ and in the final drinking water, whereas the highest number of reads corresponding to different ASVs within samples were found in GW, in CF and in the MIX.

Shannon indices reveal high diversity in all DWTP samples, with the values ranging from 3.3 to 5.2, showing a similar trend to that of the Chao index, with only minor differences at certain stages. Although diversity was slightly higher in GW (5.17) than in RW (4.65), the difference is not statistically significant ($p = 0.67$). After DEC (4.07), diversity values significantly reduced ($p < 0.01$) and remained the same after SF (4.19). In the conventional treatment, OZ led to a small reduction in diversity (3.66), which then increased significantly ($p < 0.01$) after CF (4.87). Similar values were obtained between RO (4.32) and SF samples as well as between MIX (4.8) and CF samples. Regarding species richness, chlorination drastically reduced diversity (3.34, $p = 0.0001$). Thus, the lowest diversity values, indicating a more uniform species composition, were obtained after OZ and in the

final drinking water, whereas the GW samples, followed by CF and MIX samples, had the highest values.

Alpha diversity was also analyzed by the Kendal correlation between the relative abundance of each ASV, and no significant correlation was found.

3.4.2. Beta Diversity Patterns

To explore differences in the bacterial community between different steps of the potabilization process, non-metric multidimensional scaling (NMDS) plots were constructed using the Bray–Curtis coefficient of sequencing results. Samples are represented, and ellipses indicate standard deviation of within-sample variance (Figure 4). All the eight samplings corresponding to the same matrices cluster together, although with a certain dispersion. GW and RW intakes are clearly separated along the second axis, and both are also separated from DW along the first and second axis. One of the GW samples (N4) is separate from the others, corresponding to a period when the DWTP was operating with 100% RW (as shown in Table 1).

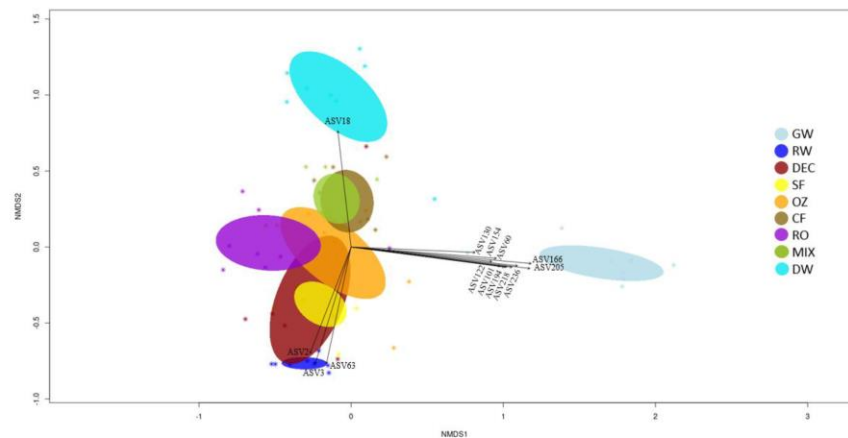


Figure 4. Bray–Curtis dissimilarity matrix (non-metric multidimensional scaling (NMDS)) of the 72 DWTP samples color-coded according to each sample type: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). Ellipses represent standard deviations of samples. The ten most abundant ASVs ($p = 0.001$) of samples from each treatment are illustrated with vectors that describe more precisely the differences within DWTP samples along both axes. GW (greyish blue) and DW (turquoise) show the least similarity with other sample types.

A partial overlap between ellipses appears for pretreatment stages and for conventional (OZ, CF) and advanced treatments (RO), and an overlap can be observed between SF and DEC samples and between FC and MIX samples.

The ten most abundant ASVs ($p < 0.001$) in DWTP samples at each treatment stage were vectorized into the Bray–Curtis nMDS representation (Figure 4). Samples of GW, RW and DW are clearly separated from these stages.

Significant vectors that separate GW samples along the first axis correspond to Planctomycetota [ASV60 (Candidatus Brocardia), ASV130 and ASV154 (GWA2-50-13 family)], Patiscibacteria [ASV218 (UBA9983 order) and ASV236 (Candidatus Giovannonibacteria order)], Verrucomicrobiota [ASV101, ASV122 and ASV194 (Candidatus Omnitrophus)] and unclassified others [ASV166 (unclassified bacteria) and ASV205 (unclassified Myxococota)]. Along the second axis, it can be seen that ASV18 was influential for the chlorinated

samples (DW), and ASV2, ASV3 and ASV63 best describe RW intake. RW was mainly represented by Proteobacteria (unclassified Comamonadaceae family and *Limnohabitans*,) and Bacteroidota (*Flavobacterium*), whereas Cyanobacteria (*Candidatus Obscuribacter*) particularly defined the chlorinated samples.

3.4.3. Bacterial Communities at Each Treatment Stage

The samples from water sources and treatments presented variable taxonomic profiles, although some consistencies in taxa were observed throughout the DWTP. The taxonomy results at the DWTP based on the relative abundance of reads on a phylum and class level are shown in Figure 5.

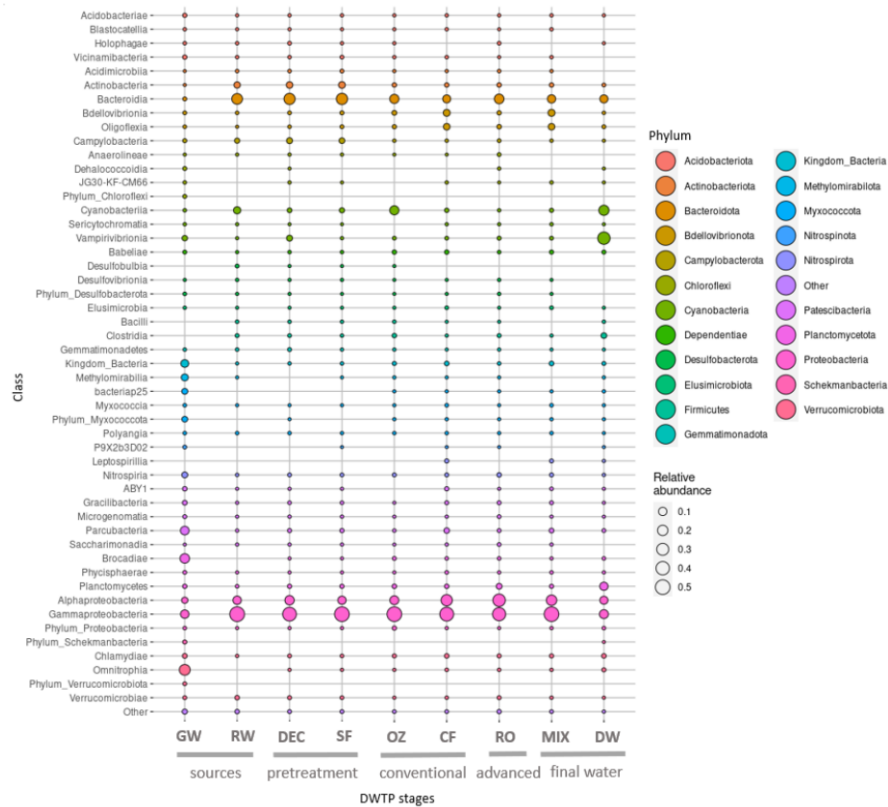


Figure 5. Bubble plot diagram showing the relative abundance (%) of ASVs grouped according to class for each DWTP stage: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). Bubble size indicates the relative abundance of each bacterial class. On the right, each class is grouped in the corresponding superior taxonomic classification (phyla).

The most abundant phylum throughout the DWTP was Proteobacteria (ranging from 57% to 73%), except in the final post-chlorination drinking water (21%). Bacteroidota

(0.8–27%) and Bdellovibrionota (1–12%) were the second and third most abundant phyla at all stages.

Among Proteobacteria, the class Gammaproteobacteria predominated throughout the DWTP, and Alphaproteobacteria increased after DEC, CF, RO and the MIX. Bacteroidia was the second most dominant class in RW and pretreatment samples, undergoing a reduction downstream. The most abundant Proteobacteria genus corresponded to *Limnohabitans*.

For a more detailed taxonomy for all samples at different hierarchy levels, see the Krona diagrams in Figure S2.

The top 50 ASVs correspond to approximately 41% of the total microbial composition in all the DWTP samples and belong to a total of 33 genera.

Taxonomy profiles of the two different water sources differed considerably in terms of phyla. Although the GW community showed higher phyla variability, ASVs were mainly affiliated to Verrucomicrobiota (25%), Planctomycetota (17%), Patescibacteria (15%) and Proteobacteria (14%). These were followed by unclassified bacteria (8%) and Methylophilum (5%), among others, and Candidatus Omnitrophus (23%) was the most abundant genus. In RW, Proteobacteria predominated (63%), followed by Bacteroidota (21%), Cyanobacteria (6%) and Actinobacteriota (5%), and the most abundant genera were *Limnohabitans* (12%) and *Flavobacterium* (11%) (Figure 6).

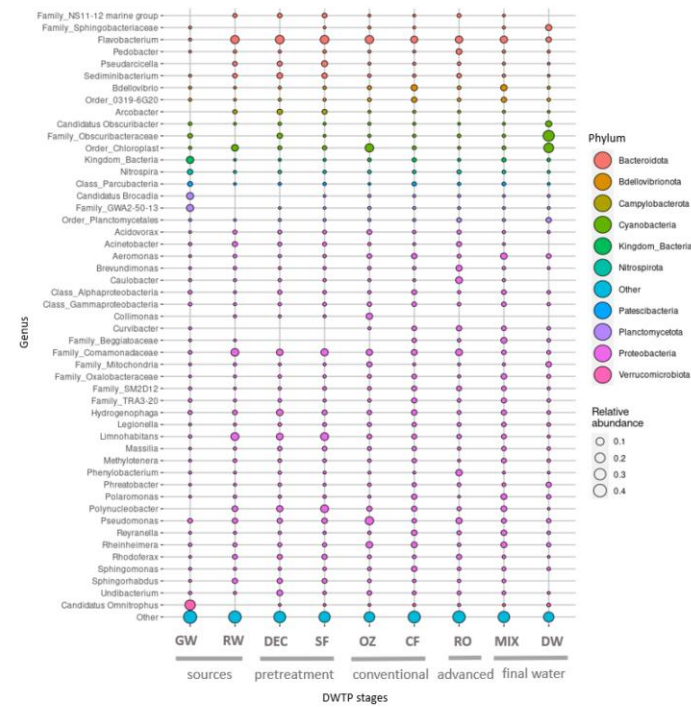


Figure 6. Bubble plot diagram showing the relative abundance (%) of ASVs grouped at the genus level for each DWTP stage: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). Bubble size indicates the relative abundance of each bacterial genus. On the right side, each class is grouped into their corresponding superior taxonomic classification (phyla).

In pretreatment, profiles of phylum, class and genus were similar between DEC and SF samples, and the most abundant genera were *Limnolobos* (7% and 10%, respectively) and *Flavobacterium* (14% and 15%, respectively). These genera were also dominant in RW.

In the conventional treatment, the first significant shift in bacterial distribution was observed after OZ, which resulted in a clear increase in Cyanobacteria (19%) and a slight decrease in Bacteroidota (11%), and the community remained dominated by Proteobacteria (63%). A reduction in Campylobacteriota (0.09%) and Actinobacteriota (0.8%) was also apparent. The most abundant genera were *Flavobacterium* (9%), *Collimonas* (8%) and *Pseudomonas* (8%), although a higher number of reads were affiliated to an unclassified Chloroplast order (18%). In the following CF step, Proteobacteria and Bacteroidota remained largely unchanged. However, Cyanobacteria (1%) abundance was lower, and Bdellovibrionota (11%) was higher. The reads corresponded, above all, to Proteobacteria (70%), Bdellovibrionota (10%) and Bacteroidota (9%). The genera *Flavobacterium* (7%), *Bdellovibrio* (5%), *Rheinheimera* (4%), *Aeromonas* (3%) and *Pseudomonas* (2%) were the most abundant.

After RO, the proportion of Proteobacteria (73%) was higher compared to SF and GW samples, whereas Bacteroidota abundance (16%) was higher than in GW but lower than in SF. Abundance of Patentedbacteria (1%), Planctomycetota (3%) and Verrucomicrobiota (1%) was lower than in GW but was similar to SF samples. *Flavobacterium* (8%), *Caulobacter* (7%), unculturable Comamonadaceae (7%) and *Pseudomonas* (5%) were the most common genera found.

In the MIX, where the conventional and advanced treatments converge, the taxonomy patterns were very similar to CF samples, with Proteobacteria (69%), Bacteroidota (12%) and Bdellovibrionota (12%) predominating. The most abundant genera were *Flavobacterium* (7%) and *Bdellovibrio* (6%), and *Legionella* was also found (2%).

Finally, the biggest shift in bacterial composition was observed in DW, where the predominant phylum was Cyanobacteria (56%). Notable proportions of Proteobacteria (21%) and Bacteroidota (8%) were also found, but the former was much less abundant than in the MIX. Firmicutes (3%) increased 100-fold, whereas the dominant genera belonged to the family Obscuribacteraceae (31%) and an unclassified Chloroplast order (20%).

3.5. Core Bacterial Communities at the DWTP

Taxa shared across the different samples were identified to determine the core communities in the DWTP and at each treatment stage. Although no ASVs were shared among all samples at all sampling sites, different core ASVs were found at each treatment stage, as shown in Table 1. Very few core ASVs were detected in samples of DEC and OZ (3 out of 2244 and 4 out of 1942, respectively), corresponding to the genera *Limnolobos*, *Porphyrobacter* and *Sediminibacterium* in DEC samples and *Aeromonas*, *Polynucleobacter*, *Porphyrobacter* and *Sphingorhabdus* in OZ samples.

Conversely, when samples were clustered according to ambient temperature (low ≤ 18 °C and high ≥ 22 °C), differences in core communities were observed. The samples from the pretreatment steps until the MIX had very few ASVs in common: ASV2 (*Limnolobos*) and ASV161 (*Reynanella*) at low temperatures and only ASV 16 (*Aeromonas*) at high temperatures. On the other hand, samples had more core ASVs at high than at low temperatures, except for GW and DEC. Venn diagrams showing the number of unique core ASVs and shared ASVs between high and low temperatures are provided in Figure 7.

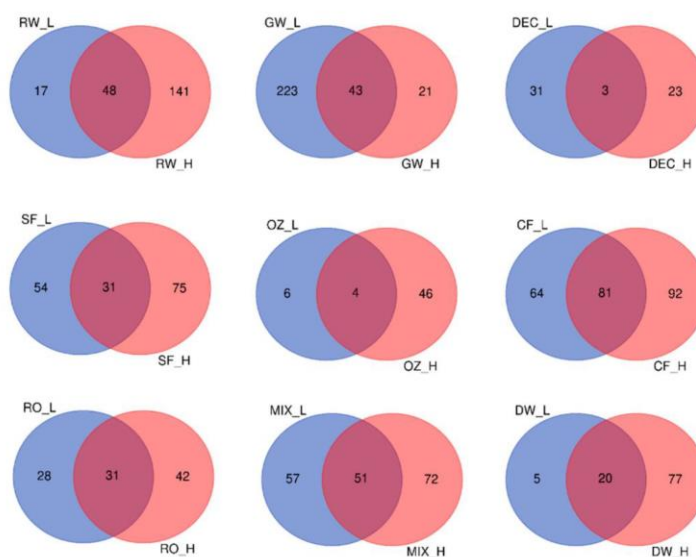


Figure 7. Venn diagrams showing the total number of shared or unique core ASVs for each treatment stage: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). Samples are separated by low (L, in blue) or high (H, in red) temperatures.

3.6. Impact of Temperature on Bacterial Communities

A moderate positive correlation between alpha diversity (Shannon index) and ambient temperature was found for some of the treatment stages, including DEC ($r = 0.73$) and CF ($r = 0.83$). A lower correlation was also found at other stages, but their values are not significant (RW $r = 0.57$, SF $r = -0.30$; OZ $r = 0.37$, RO $r = 0.51$, MIX $r = 0.31$, DW $r = 0.31$).

Beta diversity correlation with temperature was examined by permutational multivariate analysis of variance (PERMANOVA) of rarefied ASVs based on a Bray–Curtis dissimilarity matrix (Table S4). PERMANOVA analysis gave significant results for the temperature factor.

Regarding taxonomy profiles at high (HT) and low temperatures (LT) (Figure S3), differences in bacterial composition were found in some stages.

The bacterial composition in GW remained quite similar. In contrast, a moderate (non-significant) correlation with temperature was observed in RW, where Proteobacteria was always dominant (~70%). However, Bacteroidota was more abundant at LT (28%) than at HT (19%), and Actinobacteriota abundance was 10-fold greater at HT (8%) than at LT (0.8%). The main genera at LT were *Flavobacterium* (24%), *Limnohabitans* (22%) and unknown Comamonadaceae (14%). The abundance of the latter remained unchanged when temperatures increased, but a reduction was observed in *Flavobacterium* (7%).

Temperature was responsible for a significant shift in DEC samples; at LT, Bacteroidota was the major phylum (51%), followed by Proteobacteria (34%) and Campylobacterota (14%), whereas at HT, Proteobacteria dominated (84%), Bacteroidota abundance decreased (15%), and no reads of Campylobacterota were obtained. The predominant genus at LT was *Flavobacterium* (41%), and at HT, it was *Limnohabitans* (35%).

Although the composition in SF samples remained generally similar, Actinobacteriota was only detected at HT (9%), and the genus *Pseudomonas* (7%) was only detected at

LT. Moreover, *Flavobacterium* was more prevalent at LT (25%) than at HT (6%), when *Polynucleobacter* became predominant (20%).

At the OZ stage, temperature strongly affected bacterial composition; at LT, the major phyla were Proteobacteria (58%) and Cyanobacteria (42%), and at HT, Proteobacteria remained predominant (66%). However, Cyanobacteria reads were negligible (0.08%), being displaced by Bacteroidota (33%), which was not detected at LT. The most abundant genera at LT were *Pseudomonas* (46%) and unclassified genera from the Chloroplast order (42%), neither of which were observed at HT, when most reads were for *Flavobacterium* (33%) and *Rheinheimera* (19%).

Bacterial composition in CF samples was also significantly correlated with the season, with diversity increasing with temperature. Proteobacteria (83%) reads clearly dominated at LT and were lower at HT (71%), when Bacteroidota abundance increased (12%). *Flavobacterium* (10%) remained unchanged, whereas *Polynucleobacter* (10%) was detected only at HT, and *Sphingomonas* (8%) was only detected at LT.

The correlation with temperature at the RO stage was moderate (non-significant), with a clear dominance of Proteobacteria (94%) at HT. The reads decreased at LT (65%), when the next most abundant phylum was Bacteroidota (34%), which was scarcely detected at HT (0.4%). The most abundant genera at LT were *Flavobacterium* (19%) and *Pseudomonas* (13%), and at HT, they were *Caulobacter* (21%) and *Phenylobacterium* (18%).

In the MIX samples, temperature-associated differences at the genus and superior taxa level were statistically insignificant. Although the reads for *Flavobacterium* remained constant (~10%), *Polynucleobacter* (15%) was only detected at HT, followed by *Aeromonas* (14%), and at LT, *Polaromonas* (11%) was second in abundance.

In the final DW, a moderate but non-significant correlation with temperature was observed, with an obvious dominance of Cyanobacteria (87%) at LT, which was reduced somewhat at HT (63%) when Planctomycetota (12%) and Firmicutes (8%) were also detected. However, their levels were otherwise very low at LT (0.6% and 0.4%, respectively). The dominant genus in DW at LT was an unknown member of the Obscuribacteraceae family (58%), which was less abundant at HT (27%) when it accounted for half of the reads jointly, with unclassified reads belonging to the Chloroplast order (27%). The proportions of *Clostridium sensu stricto 1* (6%) and *Aeromonas* (5%) were greater at HT than at LT (0.4% and 0%, respectively). Regarding the ASV distribution according to temperature, the 50 most abundant core ASVs at HT and LT throughout the DWTP are illustrated in Figure 8.

3.7. Tracking the Origin of the ASVs along the DWTP

To investigate which upstream stages in the DWTP contribute most to the microbial communities downstream, samples corresponding to the final stages were analyzed by the SourceTracker algorithm. Intake water, pretreatment stages, as well as conventional and advanced treatments were treated as “source”, whereas the final stages, the MIX and post-chlorination drinking water (DW) were “sink” samples. The results of the analysis are shown in Table 2. Before the chlorination stage (Table 2), the reads from the MIX samples mainly matched those of CF (89.7%), with some also corresponding to SF (2.7%), RO (2%) and unknown sources (4.4%). However, after the chlorination stage (Table 2), most of the reads from the DW were from unknown sources (67.9%), followed by the OZ stage (17.4%) and the MIX (7.9%). GW and RO also contributed to the bacterial communities of the final chlorinated water (1.4% and 2%, respectively), whereas the contribution of CF was negligible (0.2%), despite being the main source of reads at the MIX stage.

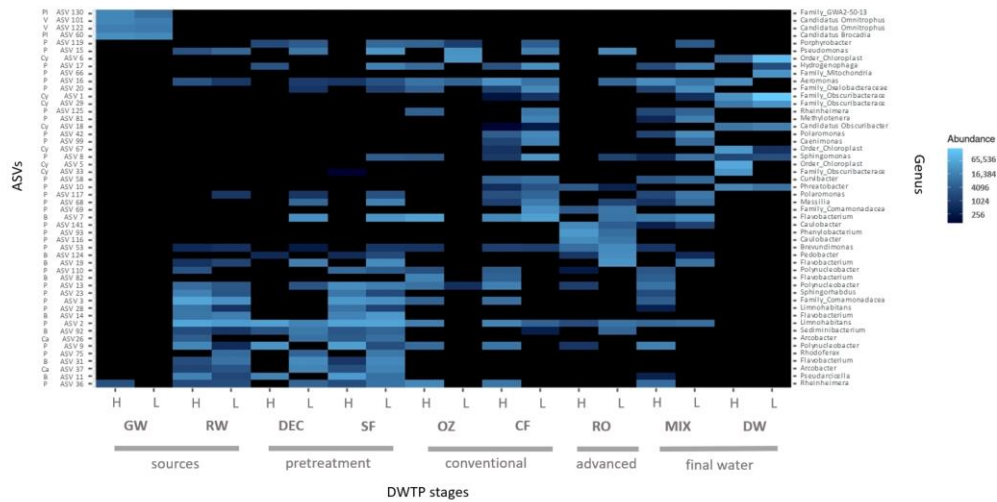


Figure 8. Heatmap of read abundances of the top 50 most abundant core ASVs in 72 DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). Samples are grouped according to the temperature of the sampling campaign, being either high (H) or low (L). Data are illustrated by a gradient color scale from black (not detected) to light blue (high relative abundance). On the left, the phyla classification for each ASV is indicated by one of the following codes: Bacteroidota (B), Campylobacterota (Ca), Cyanobacteria (Cy), Proteobacteria (P), Planctomycetota (Pl) or Verrucomicrobiota (V). ASV classification at the genus level is shown on the right.

Table 2. Average contribution of each source (DWTP stages) to bacterial communities in the MIX sink and DW sink. Samples correspond to: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW).

	GW	RW	DEC	SF	OZ	CF	RO	MIX	Unknown
Mean (MIX used as sink)	0.0%	0.2%	0.1%	2.7%	0.9%	89.7%	2.0%	-	4.4%
SD	0.0%	0.2%	0.2%	4.7%	1.3%	5.2%	2.5%	-	4.0%
Mean (DW used as sink)	1.4%	0.0%	0.1%	0.0%	17.4%	0.2%	2.0%	7.9%	67.9%
SD	3.9%	0.0%	0.2%	0.0%	25.2%	0.2%	5.5%	22.2%	39.9%

3.8. Surface Water Contamination Episode and Changes in Operational Procedures

Water quality values of GW and CF samples during the river contamination event were similar to those registered in the other sampling campaigns; in GW, all bacterial indicators were negative, whereas only 5 NMP/100 mL for TC and 1 NMP/100 mL for EC were detected in CF samples. HPC accounted for 0.48 CFU/mL and 104 CFU/mL in GW and CF, respectively.

The taxonomic profile in GW during the event was similar to that of the other GW samples, with no evident changes in water quality parameters and metabarcoding results. However, metabarcoding analysis revealed different diversity patterns in the CF water

in comparison with the samples taken under usual operating conditions (Figure S4). Although the phyla remained the same, there were some differences in relative abundance of Proteobacteria (30% lower), Bdellovibrionota, (22% lower), and Verrucomicrobiota (7.6% higher). Differences were also observed in ASV composition: 126 ASVs were shared with core CF samples (at low temperatures), and a total of 711 unique ASVs were found in CF samples during the event. At the genus level, *Bdellovibrio* was dominant (18%), followed by unclassified members of the Oligoflexia order 0319-6G20 (7%), *Reyranella* (7%) and *Candidatus Omnitrophus* (6%), whereas CF samples taken at low temperatures contained *Flavobacterium* (10%), *Sphingomonas* (8%), *Hydrogenophaga* (7%) and *Polaromonas* (6%), among others. Nonetheless, these results should be interpreted with caution, as only one sample was studied, and it may reflect occasional operational variation in the carbon filters.

The suspension of RW intake allowed us to track the effect of GW on CF communities, as none of the other pretreatment stages were in operation. As seen in Figure 9, CF effectively removed Acidobacteriota and Methylomirabilota from the GW and reduced Myxococcota, Patescibacteria, Planctomycetota and Verrucomicrobiota. However, an increase in Bdellovibrionota (from 0.2% to 26%) and Proteobacteria (from 4% to 53%) was observed in the CF samples.

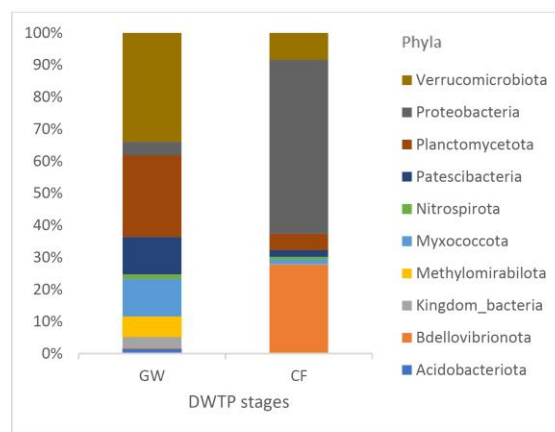


Figure 9. Bar charts of ASV relative abundance, indicating phyla in groundwater (GW) and carbon filtration (CF) samples during the river contamination episode.

3.9. Detection of Microbial Indicators and Potential Pathogens by 16S rRNA Sequencing

Sequences of bacterial groups relevant for water systems based on the Global Water Pathogens list [45] were found. A total of eight potentially pathogenic bacterial genera were identified in the DWTP with a very low relative abundance of reads (0.002 to 8%) (Table S5). *Pseudomonas* was detected at the OZ stage (8%) and in the DW (1%), *Aeromonas* and *Legionella* were detected in the MIX (4% and 2%, respectively) and DW (2% and 0.4%, respectively) and *Clostridium* was detected in OZ (1%) and DW (2%). Other potentially pathogenic genera, *Arcobacter*, *Campylobacteriales*, *Escherichia-Shigella* and *Mycobacterium*, were only detected in source water or at early treatment stages, with their abundance being very low (<1%) in later treatments or in the final DW. *Helicobacter*, *Leptospira* and *Vibrio* were not detected in any sample.

Regarding data on culturable water quality indicators (Section 3.2.), it was observed that culturable *E. coli*, an indicator of fecal contamination, accounted for an average of $3.5 \log_{10}$ CFU/100 mL in RW, whereas in metabarcoding analysis, it contributed only 0.003% of reads for *Escherichia-Shigella*. Additionally, *Clostridium sensu stricto* 1 was

found in DW in low abundance (2%), although it was higher than in the previous stages (0.002–0.02%), but no culturable *C. perfringens* was detected in the DW. Finally, *Enterococcus* was only detected in very low abundance (0.002%) in SF, OZ and CF.

4. Discussion

The performance of an advanced full-scale DWTP was assessed using culturable microbial indicators, as indicated in the current EU Directive on drinking water together with molecular methods (16S rRNA metabarcoding). Monitoring the dynamics of bacterial communities in a DWTP may help to better understand process performance, allowing a more in-depth assessment of operational variations that may modify water quality. For example, nutrient excess in source water may lead to regrowth of microorganisms or biofilm formation [3,46], which can affect the taste and odor of tap water and may even result in an increase in opportunistic pathogens, posing a risk to public health. Moreover, an increase in cyanobacterial blooms in surface water may produce clogging in filtration stages as well as being a health hazard, as some genera are toxin producers [2].

Although the quality of the analyzed drinking water complies with EU Directive regulations, the stipulated culture-based indicators were not able to detect changes in bacterial communities in the DWTP. For example, the HPC remained constant despite shifts in bacterial diversity.

The use of a DEUF concentration method allowed the reduction in microbial indicators to be monitored more precisely during water processing, especially in the advanced treatment stages, when indicator concentration is very low and not detected by conventional techniques. Thus, among the culturable indicators, *C. perfringens* was traced until carbon filtration, and total coliforms, *E. coli*, intestinal enterococci and somatic coliphages were detected until the mixing chamber, with no traces remaining in the final drinking water. Historical data from this DWTP (from 2011 to 2019, data not shown), based on the use of standard membrane filtration for routine water quality monitoring, indicate positive results for these indicators only at early stages (i.e., *E. coli* and somatic coliphages until sand filtration, or intestinal enterococci until carbon filtration). Previous studies using the DEUF method have reported recoveries of $45.5 \pm 24.0\%$ for fecal indicators and $22.4 \pm 9.3\%$ for bacteriophages [47], ranging from 35 to 95% for enteric bacteria and viruses [48] or 60–80% for bacteria, viruses and protozoa in drinking water [20], depending on the water matrix and degree of turbidity. Therefore, the reported concentrations for the different indicators may be even higher than in the current study.

Compared to other methods, DEUF allowed for more in-depth and representative characterization of the bacterial communities in the DWTP, as large water volumes could be sampled. Contaminants from DNA extraction kits or PCR reagents can confound the results of molecular analysis, especially in samples with a low microbial biomass, such as the final treated water [49]. In this context, large volume concentration methods are particularly useful. In our PCR procedure, sequence blanks clustered together, indicating that they did not significantly interfere with the characterization of the DWTP microbiome.

High diversity was observed at all DWTP stages, being the highest in GW, followed by the MIX and CF samples. Overall species abundance and richness were significantly reduced by the treatments, especially disinfection, with the lowest values being obtained in the OZ, DW and DEC samples. River water, the main intake source of the DWTP, had a lower species richness than those of the SF, FC and MIX samples. The higher diversity observed at SF and CF may indicate an established microbiome associated with filtration procedures, as has been reported [13]. Diversity at the DEC stage was lower than it was after the following step of SF, reflecting a loss in richness after the disinfection procedure with chlorine dioxide and possible bacterial regrowth in the filters. A similar trend was observed in the conventional treatment, with a decrease in diversity after OZ, followed by a high increase in CF samples and again with a decrease in the advanced treatment, where there was reduced species richness at the RO stage. Diversity in the mixing chamber was high, falling between the values of the CF and SF samples. Finally, chlorination significantly

reduced species diversity in the final water. Chlorine has a strong bactericidal effect, damaging bacterial cell structure, although certain groups resist the disinfection process. ASVs of bacterial communities characterized at disinfection stages may correspond to chlorine-resistant strains of *Aeromonas*, *Clostridium*, Cyanobacteria or *Pseudomonas*, among others (reviewed in Luo et al. [50]). Low diversity in RO samples can be explained by the low nutrient levels in permeate water, where only specially adapted microorganisms can thrive. Genera detected at this stage include *Caulobacter*, *Comamonas*, *Curvibacter*, *Flavobacterium* and *Pseudomonas*, which are considered potentially biofilm-forming.

Although Bdellovibrionota was detected at different treatment stages with $\leq 1\%$ abundance, it accounted for 10 and 12% of reads in the CF and MIX samples, respectively, mainly consisting of unculturable Oligoflexia and *Bdellovibrio*. The members of this phylum are described as obligate predators of bacteria (Gram-negative and Gram-positive) and have been proposed as biocontrol agents for their ability to prey on multi-drug-resistant bacteria [51]. Concerning core amplicon sequence variants, only one ASV was detected in the samples taken at high temperatures and two in the samples taken at low temperatures, which correspond to *Aeromonas* and to *Limnohabitans* and *Reyranella*, respectively. The detection of core ASVs under the standard operational conditions of the DWTP is of interest, as they can provide new monitoring tools for detecting impaired treatments, a potential application that still needs more research.

Over the short period when GW was the only source of intake and when the routine operational conditions of the DWTP changed, an increase in certain bacterial populations at the CF stage was detected, which can be attributed to the growth of communities present in biofilms in the filter or to their temporal detachment from carbon surfaces. Such conditions can be met after a sudden shift of water intake for the exclusive use of GW and can be attributed to differences in temperatures compared with other water sources [52]. In this study, in the winter season when this episode occurred, river water temperature gathered at 7.9 °C, whereas groundwater presented 17.1 °C on the day of sampling. Nevertheless, in the current study, the significance of these results may be limited by the unrepresentativeness of the samples. Different potabilization stages exerted a selective pressure on bacterial communities and shaped their composition downstream. Filtration procedures had a strong influence, which is in agreement with other studies on sand filtration [13] and carbon filtration [16,53]. However, disinfection had a greater impact, as reported previously, as residual disinfectants reduce the abundance and diversity of bacterial communities, which become dominated by resistant genera (including opportunistic pathogens) in post-chlorination water [14,15,54] and in distribution networks [54–56]. In the present study, the chlorination stage shifted the dominant bacterial population in the final drinking water from Proteobacteria to Cyanobacteria, a phylum that was poorly represented in source water (0–3%) and at other treatment stages (<6%), being slightly higher in OZ samples (19%). The aphotic environment of many of these processes (but not the pretreatment) suggests that the unclassified Cyanobacteria that were detected are non-photosynthetic. The dominant genus in the DW belongs to the unculturable Obscuribacteraceae, which are found in dark conditions, followed by an unclassified Chloroplast order. Most databases, such as SILVA, classify chloroplasts within the Cyanobacteria phylum, as they are considered a phylogenetic lineage evolved from endosymbiotic Cyanobacteria [57]. Thus, chloroplast reads are commonly found in aquatic environments, and have been reported in studies of river [58] and lake [59] water quality.

Little is known about the diversity and ecology of non-photosynthetic Cyanobacteria in water environments, and their classification is still under debate. Di Rienzi and colleagues [60] proposed their inclusion in a sibling phylum, Melainobacteria (named after “Melaina”, a nymph of dark waters in Greek mythology), whose sequenced genomes confirm an absence of a photosynthetic apparatus. Another study suggests that the orders Obscuribacterales and Vampirovibrionales, which are capable of aerobic and anaerobic respiration, should also be included in this new phylum [61]. Although the SILVA database does not phylogenetically differentiate the non-photosynthetic cyanobacterial

group, Cyanobacteria have been reported in aphotic environments, such as groundwater and animal guts [60], koala feces and activated sludge from a wastewater treatment plant [61] and continental subsurfaces 420–607 m in depth [62].

Although most studies characterizing bacterial communities in DWTP stages and drinking water have not detected Cyanobacteria and report Proteobacteria as the dominant phylum [11,13–15,25,55,63,64], other researchers have obtained a high abundance of Cyanobacteria reads in water from distribution networks [8], treatment stages [65] or sludge storage in drinking water processing [65,66]. These variable results may be accounted for by factors such as location, the type and quality of source water, and whether or not disinfectant procedures are used, all of which create a unique habitat in each DWTP or distribution network. Cyanobacteria cells have an outer peptidoglycan layer thicker than most Gram-negative bacteria and a cell wall mainly composed of lipids and carotenoids. Furthermore, the outer membrane has fewer conductance porins than most Gram-negative bacteria, which is an adaptation to low-nutrient environments that enhances resistance toward harmful chemicals [57]. This structure may confer mechanical stability when exposed to chlorine, with chlorine-resistant groups being found in the final water. No cyanobacterial blooms, stagnation or microcystin (cyanobacterial toxins) were detected in the river or drinking water, which was safe for consumption.

Changes in microbial communities related to ambient temperature were also observed in different treatment stages, with a correlation of higher diversity with higher temperatures, specially at DEC and CF. Such temporal variability has been widely described [10,13,14,67]. Higher temperatures can trigger changes in bacterial communities, as they facilitate bacterial growth of different genera. Implications of high temperatures in water systems have been documented, such as the occurrence of certain waterborne bacterial pathogens, opportunistic pathogens and biofilm-forming groups that may compromise water quality or pose a risk to public health. For instance, Calero Preciado et al. (2021) [3] observed a significant increase in relative abundant reads of *Pseudomonas* and *Fusarium* in chlorinated distribution networks favored by temperature, both with the capability to promote biofilm formation. Moreover, the presence of the opportunistic pathogen *Mycobacterium* spp. was also registered to rise together with temperature. Other genera that promote biofilm development in chlorinated distribution networks, such as *Aeromonas*, also show greater abundance at higher temperatures [68]. Similarly, in unchlorinated drinking water supplies, van der Wielen and van der Kooij (2013) [69] described a higher occurrence of *Legionella pneumophila* and *Stenotrophomonas maltophilia* at high temperatures. Both of them are likely to grow in biofilms and can be moderately to highly virulent, as they often cause pneumonia or bacteremia disease, respectively, in immunocompromised patients. These observed changes related to temperature may be of interest to drinking water managers in the context of the global warming scenario.

Very few 16S rRNA sequences (0.002–8%) retrieved from the source water and treatment stages were associated with genera of waterborne bacterial pathogens in the global priority list. Reads of *Aeromonas*, *Arcobacter*, the order Campylobacterales, *Legionella*, *Mycobacterium*, *Escherichia-Shigella* and *Pseudomonas* detected in the DWTP were reduced or eliminated in the DW, except for *Pseudomonas* and *Aeromonas* (1% and 2%). *Arcobacter* reads, which were low in RW and pretreatment stages (1–3%), were negligible in drinking water (0.007%), demonstrating the effectiveness of conventional and advanced treatments in removing *Arcobacter* species, as previously shown by Collado et al. (2010) [70]. *Helicobacter*, *Leptospira* and *Vibrio* were not detected at any stage of the process. It should be mentioned that not all the species of the targeted potentially pathogenic genera are animal, human or plant pathogens. Additionally, the reads may correspond to non-viable or dead cells after the potabilization treatment or to free DNA. The present analysis of drinking water samples confirms that the applied potabilization process provides safe drinking water, free of fecal indicators.

Discrepancies were observed in the results between conventional culture-based techniques and metabarcoding. In RW, culturable *E. coli* values were high (3.5 log₁₀ CFU/100 mL),

but the abundance of reads was very low (0.003%), perhaps because of a PCR bias against rare taxa in non-treated water samples. In DW, the opposite occurred with *C. perfringens*, which was detected by the molecular technique (3%) but not by culture. However, relative abundance data must be treated with caution, as high-throughput sequencing is not designed for quantitative purposes. Therefore, the analysis of water quality based only on sequencing methods can lead to inaccurate assessments of microbial hazards [54]. Nevertheless, a molecular approach allows for the simultaneous tracking of pathogens, opportunistic pathogens, microbial indicators and their fluctuations between treatments and within samples. The resulting taxonomic data can be used to develop quantitative methods such as qPCR to provide a more comprehensive understanding of microbial population dynamics.

Both approaches to water quality monitoring (culture- and molecular-based) have benefits and limitations. Although microbial indicators and HPC offer protection against pathogens and can detect viable microorganisms, the results can be biased by the medium or by growth conditions, and changes in microbial composition are not detected. In this respect, 16S rRNA metabarcoding provides a more in-depth analysis of microbial communities, including both culturable and non-culturable, autotrophic and heterotrophic, photic and aphotic and waterborne bacteria. However, metabarcoding also has shortcomings: obtaining results is more time-consuming; there is no standardized methodology; rare taxa can be neglected due to primer bias; and cell viability remains unconfirmed, which hinders risk assessment [71]. Although all drinking water samples in this study fully met the quality standards of the EU Directive, samples with a stable HPC were observed to have variable bacterial diversity. Therefore, the use of 16S rRNA metabarcoding in water quality management in combination with standardized techniques can provide useful information on the efficiency of the different stages of water treatment and may help to address the challenge of potabilization related to likely expected deterioration of surface water due to future climate threats that may compromise water quality.

5. Conclusions

The application of 16S rRNA metabarcoding to assess microbial communities in a full-scale DWTP in Barcelona (Spain), which uses two different water sources, a pretreatment stage and a combination of conventional and advanced treatments with final chlorination, provides new insights into compositional changes throughout the process.

Highly diverse microbial communities were observed, suggesting that each treatment stage has specific microbiota, including unculturable bacteria.

Differences in community structure can be explained by the water sources, the type of treatment applied at each stage (especially filtration and chlorination procedures), microhabitats or biofilm formation in carbon filtration and, to a lesser extent, seasonal variations in ambient temperature. Proteobacteria was the dominant phylum in source water and throughout the procedure, with a big shift after the chlorination stage when it was displaced by Cyanobacteria, which was previously found in low abundance, except at the ozonization stage.

The studied drinking water met all the quality standards stipulated by the EU Directive. However, molecular analysis revealed that bacterial diversity, even in samples with a stable HPC, differs throughout the process. Therefore, the use of 16S rRNA metabarcoding in combination with standardized parameters provides a more in-depth analysis of the microbiota in a DWTP, which can be useful for the detection of operational variations and for their impact on water quality.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/w14091435/s1>. Supplementary material associated with this article: Supplementary File S1 (Tables S1–S5 and Figure S1). Supplementary File S2 (interactive version of Figure S2): Figure S2. Krona plots of the relative abundance reads of bacteria detected by 16S metabarcoding at all DWTP stages: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Taxonomic profiles are simultaneously displayed

by hierarchy levels from kingdom to genus by selecting taxonomic depths: 1: Kingdom 2: Phylum 3: Class 4: Order 5: Family 6: Genus. Supplementary File S3 (interactive version of Figure S3): Figure S3. Krona charts of the core community bacteria detected by 16S metabarcoding at all DWTP stages grouped by sampling time at high (H) or low temperatures (L): groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Taxonomic profiles are simultaneously displayed by hierarchy levels from kingdom to genus by selecting taxonomic depths: 1: Kingdom 2: Phylum 3: Class 4: Order 5: Family 6: Genus. Supplementary File S4 (interactive version of Figure S4): Figure S4. Krona charts of relative abundant core ASVs in groundwater (GW_core_L) and CF (FC_core_L) taken at low temperatures under usual operating conditions, and GW (GW_Cont_Episode) and CF (CF_Cont_Episode) during the river contamination episode.

Author Contributions: A.P.-M. performed sampling and water quality analysis; A.P.-M. and O.S.W. carried out the metabarcoding experiments; A.P.-M., O.S.W., J.M. and C.G.-A. processed the data analysis; C.G.-A., A.R.B., B.G., K.P. and O.S.W. conceived and supervised the study; A.P.-M. wrote the manuscript. All authors contributed intellectual content and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Pla de Doctorats Industrials of the Catalan Government grant number 2016 DI 083.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original DNA sequences are available upon request.

Acknowledgments: This study was supported by Pla de Doctorats Industrials of the Catalan Government (2016 DI 083) and the Water Research Institute. We thank Jordi Martin, Marta Ganzer and Sergio Montes for their support and expertise in DWTP procedures. The graphical abstract and Figure 1 were created with BioRender.com, accessed on 15 February 2022.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Whitehead, P.G.; Wilby, R.L.; Battarbee, R.W.; Kernan, M.; Wade, A.J. A review of the potential impacts of climate change on surface water quality. *Hydrol. Sci. J.* **2009**, *54*, 101–123. [[CrossRef](#)]
- O’Neil, J.M.; Davis, T.W.; Burford, M.A.; Gobler, C.J. The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change. *Harmful Algae* **2012**, *14*, 313–334. [[CrossRef](#)]
- Calero Preciado, C.; Boxall, J.; Soria-Carrasco, V.; Martínez, S.; Douterelo, I. Implications of Climate Change: How Does Increased Water Temperature Influence Biofilm and Water Quality of Chlorinated Drinking Water Distribution Systems? *Front. Microbiol.* **2021**, *12*, 658927. [[CrossRef](#)] [[PubMed](#)]
- Li, W.; Zhang, J.; Wang, F.; Qian, L.; Zhou, Y.; Qi, W.; Chen, J. Effect of disinfectant residual on the interaction between bacterial growth and assimilable organic carbon in a drinking water distribution system. *Chemosphere* **2018**, *202*, 586–597. [[CrossRef](#)]
- Favere, J.; Barbosa, R.G.; Sleutels, T.; Verstraete, W.; De Gussem, B.; Boon, N. Safeguarding the microbial water quality from source to tap. *NPJ Clean Water* **2021**, *4*, 28. [[CrossRef](#)]
- Prest, E.I.; Hammes, F.; van Loosdrecht, M.C.M.; Vrouwenvelder, J.S. Biological stability of drinking water: Controlling factors, methods, and challenges. *Front. Microbiol.* **2016**, *7*, 45. [[CrossRef](#)]
- Prest, E.I.; El-Chakhtoura, J.; Hammes, F.; Saikaly, P.E.; van Loosdrecht, M.C.M.; Vrouwenvelder, J.S. Combining flow cytometry and 16S rRNA gene pyrosequencing: A promising approach for drinking water monitoring and characterization. *Water Res.* **2014**, *63*, 179–189. [[CrossRef](#)]
- Liu, T.; Kong, W.; Chen, N.; Zhu, J.; Wang, J.; He, X.; Jin, Y. Bacterial characterization of Beijing drinking water by flow cytometry and MiSeq sequencing of the 16S rRNA gene. *Ecol. Evol.* **2016**, *6*, 923–934. [[CrossRef](#)]
- Lautenschlager, K.; Hwang, C.; Liu, W.T.; Boon, N.; Köster, O.; Vrouwenvelder, H.; Egli, T.; Hammes, F. A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Res.* **2013**, *47*, 3015–3025. [[CrossRef](#)]
- Romero, P.E.; Calla-Quispe, E.; Castillo-Vilcahuaman, C.; Yokoo, M.; Fuentes-Rivera, H.L.; Ramirez, J.L.; Ampuero, A.; Ibáñez, A.J.; Wong, P. From the Andes to the desert: 16S rRNA metabarcoding characterization of aquatic bacterial communities in the Rimac river, the main source of water for Lima, Peru. *PLoS ONE* **2021**, *16*, e0250401. [[CrossRef](#)]
- Chao, Y.; Ma, L.; Yang, Y.; Ju, F.; Zhang, X.X.; Wu, W.M.; Zhang, T. Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Sci. Rep.* **2013**, *3*, 3550. [[CrossRef](#)]

12. Shaw, J.L.A.; Monis, P.; Weyrich, L.S.; Sawade, E.; Drikas, M.; Cooper, A.J. Using amplicon sequencing to characterize and monitor bacterial diversity in drinking water distribution systems. *Appl. Environ. Microbiol.* **2015**, *81*, 6463–6473. [CrossRef] [PubMed]
13. Pinto, A.J.; Xi, C.; Raskin, L. Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environ. Sci. Technol.* **2012**, *46*, 8851–8859. [CrossRef]
14. Hou, L.; Zhou, Q.; Wu, Q.; Gu, Q.; Sun, M.; Zhang, J. Spatiotemporal changes in bacterial community and microbial activity in a full-scale drinking water treatment plant. *Sci. Total Environ.* **2018**, *625*, 449–459. [CrossRef] [PubMed]
15. Li, Q.; Yu, S.; Li, L.; Liu, G.; Gu, Z.; Liu, M.; Liu, Z.; Ye, Y.; Xia, Q.; Ren, L. Microbial Communities Shaped by Treatment Processes in a Drinking Water Treatment Plant and Their Contribution and Threat to Drinking Water Safety. *Front. Microbiol.* **2017**, *8*, 2465. [CrossRef] [PubMed]
16. Bruno, A.; Sandionigi, A.; Bernasconi, M.; Panio, A.; Labra, M.; Casiraghi, M. Changes in the drinking water microbiome: Effects of water treatments along the flow of two drinking water treatment plants in a urbanized area, milan (Italy). *Front. Microbiol.* **2018**, *9*, 2557. [CrossRef] [PubMed]
17. Belila, A.; El-Chakhtoura, J.; Otaibi, N.; Muyzer, G.; Gonzalez-Gil, G.; Saikaly, P.E.; van Loosdrecht, M.C.M.; Vrouwenvelder, J.S. Bacterial community structure and variation in a full-scale seawater desalination plant for drinking water production. *Water Res.* **2016**, *94*, 62–72. [CrossRef] [PubMed]
18. Rhodes, E.R.; Hamilton, D.W.; See, M.J.; Wymer, L. Evaluation of hollow-fiber ultrafiltration primary concentration of pathogens and secondary concentration of viruses from water. *J. Virol. Methods* **2011**, *176*, 38–45. [CrossRef] [PubMed]
19. Hill, V.R.; Kahler, A.M.; Jothikumar, N.; Johnson, T.B.; Hahn, D.; Cromeans, T.L. Multistate Evaluation of an Ultrafiltration-Based Procedure for Simultaneous Recovery of Enteric Microbes in 100-Liter Tap Water Samples. *Appl. Environ. Microbiol.* **2007**, *73*, 4218–4225. [CrossRef]
20. Gunnarsdóttir, M.J.; Gardarsson, S.M.; Figueras, M.J.; Puigdomènech, C.; Juárez, R.; Saucedo, G.; Arnedo, M.J.; Santos, R.; Monteiro, S.; Avery, L.; et al. Water safety plan enhancements with improved drinking water quality detection techniques. *Sci. Total Environ.* **2020**, *698*, 134185. [CrossRef]
21. EPA; CDC. *Comparison of Ultrafiltration Techniques for Recovering Biothreat Agents in Water*; EPA/600/R-11/103; U.S. Environmental Protection Agency: Washington, DC, USA, 2011.
22. Ferrari, S.; Frosth, S.; Svensson, L.; Fernström, L.L.; Skarin, H.; Hansson, I. Detection of *Campylobacter* spp. in water by dead-end ultrafiltration and application at farm level. *J. Appl. Microbiol.* **2019**, *127*, 1270–1279. [CrossRef] [PubMed]
23. Cuevas-Ferrando, E.; Randazzo, W.; Pérez-Cataluña, A.; Sánchez, G. HEV Occurrence in Waste and Drinking Water Treatment Plants. *Front. Microbiol.* **2020**, *10*, 2937. [CrossRef] [PubMed]
24. Francy, D.S.; Stelzer, E.A.; Brady, A.M.G.; Huitger, C.; Bushon, R.N.; Ip, H.S.; Ware, M.W.; Villegas, E.N.; Gallardo, V.; Lindquist, H.D.A. Comparison of filters for concentrating microbial indicators and pathogens in lake water samples. *Appl. Environ. Microbiol.* **2013**, *79*, 1342–1352. [CrossRef]
25. Wang, F.; Li, W.; Li, Y.; Zhang, J.; Chen, J.; Zhang, W.; Wu, X. Molecular analysis of bacterial community in the tap water with different water ages of a drinking water distribution system. *Front. Environ. Sci. Eng.* **2018**, *12*. [CrossRef]
26. Revetta, R.P.; Pemberton, A.; Lamendella, R.; Iker, B.; Santo Domingo, J.W. Identification of bacterial populations in drinking water using 16S rRNA-based sequence analyses. *Water Res.* **2010**, *44*, 1353–1360. [CrossRef]
27. Martin-Alonso, J. Barcelona's water supply improvement: The brine collector of the Llobregat river. *Water Sci. Technol.* **1994**, *30*, 221–227. [CrossRef]
28. Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the Quality of Water Intended for Human Consumption. Available online: <https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32020I284>. (accessed on 23 January 2022).
29. *International Standard ISO 6222:1999*; Water Quality—Enumeration of Culturable Micro-Organisms—Colony Count by Inoculation in a Nutrient Agar Culture Medium. International Organization for Standardization: Geneva, Switzerland, 1999.
30. *International Standard ISO 9308-2:2012*; International Standard Water Quality—Enumeration of *Escherichia coli* and Coliform bacteria—Part 2: Most Probable Number Method. International Organization for Standardization: Geneva, Switzerland, 2012.
31. *UNE-EN ISO 7937:2005*; Microbiology of Food and Animal Feeding Stuffs—Horizontal Method for the Enumeration of *Clostridium Perfringens*—Colony-Count Technique. International Organization for Standardization: Geneva, Switzerland, 2005.
32. *UNE-EN ISO 26461-2:1995*; Water Quality—Detection and Enumeration of the Spores of Sulfite-Reducing Anaerobes (*Clostridia*): Part 2: Method by Membrane Filtration. International Organization for Standardization: Geneva, Switzerland, 1995.
33. *International Standard ISO 7899-2:2000*; Water Quality—Detection and Enumeration of Intestinal Enterococci—Part 2: Membrane Filtration Method. International Organization for Standardization: Geneva, Switzerland, 2000.
34. *International Standard ISO 10705-2:2000*; Water Quality—Detection and Enumeration of Bacteriophages—Part 2: Enumeration of Somatic Coliphages. International Organization for Standardization: Geneva, Switzerland, 2000.
35. Méndez, J.; Audicana, A.; Isern, A.; Llana, J.; Moreno, B.; Tarancón, M.L.; Jofre, J.; Lucena, F. Standardised evaluation of the performance of a simple membrane filtration-elution method to concentrate bacteriophages from drinking water. *J. Virol. Methods* **2004**, *117*, 19–25. [CrossRef]
36. *International Standard ISO 10705-3:2003*; Water Quality—Detection and Enumeration of Bacteriophages—Part 3: Validation of Methods for Concentration of Bacteriophages from Water. International Organization for Standardization: Geneva, Switzerland, 2003.

37. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4516–4522. [[CrossRef](#)]
38. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMB J.* **2011**, *17*, 10. [[CrossRef](#)]
39. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [[CrossRef](#)]
40. Bolyen, E.; Rideout, J.R.; Dillon, M.R.; Bokulich, N.A.; Abnet, C.C.; Al-Ghalith, G.A.; Alexander, H.; Alm, E.J.; Arumugam, M.; Asnicar, F.; et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852–857. [[CrossRef](#)] [[PubMed](#)]
41. McMurdie, P.J.; Holmes, S. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [[CrossRef](#)] [[PubMed](#)]
42. Bokulich, N.A.; Kaehler, B.D.; Rideout, J.R.; Dillon, M.; Bolyen, E.; Knight, R.; Huttley, G.A.; Gregory Caporaso, J. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **2018**, *6*, 90. [[CrossRef](#)] [[PubMed](#)]
43. Davis, N.M.; Proctor, D.M.; Holmes, S.P.; Relman, D.A.; Callahan, B.J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* **2018**, *226*. [[CrossRef](#)] [[PubMed](#)]
44. Knights, D.; Kuczynski, J.; Charlson, E.S.; Zaneveld, J.; Mozer, M.C.; Collman, R.G.; Bushman, F.D.; Knight, R.; Kelley, S.T. Bayesian community-wide culture-independent microbial source tracking. *Nat. Methods* **2011**, *8*, 761–763. [[CrossRef](#)]
45. Pruden, A.; Ashbolt, N.; Miller, J. Overview of issues for water bacterial pathogens. In *Water and Sanitation for the 21st Century: Health and Microbiological Aspects of Excreta and Wastewater Management*; Rose, J.B., Jiménez-Cisneros, B., Eds.; Global Water Pathogen Project; Part 3: Specific Excreted Pathogens: Environmental and Epidemiology Aspects—Section 2: Bacteria; Michigan State University: East Lansing, MI, USA; UNESCO: New York, NY, USA. [[CrossRef](#)]
46. Husband, S.; Fish, K.E.; Douterelo, I.; Boxall, J. Linking discoloration modelling and biofilm behaviour within drinking water distribution systems. *Water Sci. Technol. Water Supply* **2016**, *16*, 942–950. [[CrossRef](#)]
47. Pascual-Benito, M.; Emiliano, P.; Casas-Mangas, R.; Dacal-Rodríguez, C.; Gracenea, M.; Araujo, R.; Valero, F.; García-Aljaro, C.; Lucena, F. Assessment of dead-end ultrafiltration for the detection and quantification of microbial indicators and pathogens in the drinking water treatment processes. *Int. J. Hyg. Environ. Health* **2020**, *230*. [[CrossRef](#)]
48. Kahler, A.; Johnson, T.; Hahn, D.; Narayanan, J.; Derado, G.; Hill, V. Evaluation of an Ultrafiltration-Based Procedure for Simultaneous Recovery of Diverse Microbes in Source Waters. *Water* **2015**, *7*, 1202–1216. [[CrossRef](#)]
49. Salter, S.J.; Cox, M.J.; Turek, E.M.; Calus, S.T.; Cookson, W.O.; Moffatt, M.F.; Turner, P.; Parkhill, J.; Loman, N.J.; Walker, A.W. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **2014**, *12*, 87. [[CrossRef](#)]
50. Luo, L.W.; Wu, Y.H.; Yu, T.; Wang, Y.H.; Chen, G.Q.; Tong, X.; Bai, Y.; Xu, C.; Wang, H.B.; Ikuno, N.; et al. Evaluating method and potential risks of chlorine-resistant bacteria (CRB): A review. *Water Res.* **2021**, *188*, 116474. [[CrossRef](#)]
51. Kadouri, D.E.; To, K.; Shanks, R.M.Q.; Doi, Y. Predatory bacteria: A potential ally against multidrug-resistant Gram-negative pathogens. *PLoS ONE* **2013**, *8*. [[CrossRef](#)] [[PubMed](#)]
52. Moona, N.; Wunsch, U.J.; Bondelind, M.; Bergstedt, O.; Sapmaz, T.; Pettersson, T.J.R.; Murphy, K.R. Temperature-dependent mechanisms of DOM removal by biological activated carbon filters. *Environ. Sci. Water Res. Technol.* **2019**, *5*, 2232–2241. [[CrossRef](#)]
53. Oh, S.; Hammes, F.; Liu, W.T. Metagenomic characterization of biofilter microbial communities in a full-scale drinking water treatment plant. *Water Res.* **2017**, *128*, 278–285. [[CrossRef](#)] [[PubMed](#)]
54. Acharya, K.; Halla, F.F.; Massawa, S.M.; Mgana, S.M.; Komar, T.; Davenport, R.J.; Werner, D. Chlorination effects on DNA based characterization of water microbiomes and implications for the interpretation of data from disinfected systems. *J. Environ. Manag.* **2020**, *276*, 111319. [[CrossRef](#)]
55. Atnafu, B.; Desta, A.; Assefa, F. Microbial Community Structure and Diversity in Drinking Water Supply, Distribution Systems as well as Household Point of Use Sites in Addis Ababa City, Ethiopia. *Microb. Ecol.* **2021**. [[CrossRef](#)]
56. Bertelli, C.; Courtois, S.; Rosikiewicz, M.; Piriou, P.; Aebly, S.; Robert, S.; Loret, J.F.; Greub, G. Reduced chlorine in drinking water distribution systems impacts bacterial biodiversity in biofilms. *Front. Microbiol.* **2018**, *9*, 2520. [[CrossRef](#)]
57. Hoiczky, E.; Hansel, A. Cyanobacterial Cell Walls: News from an Unusual Prokaryotic Envelope. *J. Bacteriol.* **2000**, *182*, 1191–1199. [[CrossRef](#)]
58. Beale, D.J.; Karpe, A.V.; Ahmed, W.; Cook, S.; Morrison, P.D.; Staley, C.; Sadowsky, M.J.; Palombo, E.A. A community multi-omics approach towards the assessment of surface water quality in an urban river system. *Int. J. Environ. Res. Public Health* **2017**, *14*. [[CrossRef](#)]
59. Zhu, B.; Cao, H.; Li, G.; Du, W.; Xu, G.; Domingo, J.S.; Gu, H.; Xu, N.; Duan, S.; Lu, J. Biodiversity and dynamics of cyanobacterial communities during blooms in temperate lake (Harsha Lake, Ohio, USA). *Harmful Algae* **2019**, *82*, 9–18. [[CrossRef](#)]
60. Di Rienzi, S.C.; Sharon, I.; Wrighton, K.C.; Koren, O.; Hug, L.A.; Thomas, B.C.; Goodrich, J.K.; Bell, J.T.; Spector, T.D.; Banfield, J.F.; et al. The human gut and groundwater harbor non-photosynthetic bacteria belonging to a new candidate phylum sibling to Cyanobacteria. *eLife* **2013**, *2*, e01102. [[CrossRef](#)]
61. Soo, R.M.; Skennerton, C.T.; Sekiguchi, Y.; Imelfort, M.; Paech, S.J.; Dennis, P.G.; Steen, J.A.; Parks, D.H.; Tyson, G.W.; Hugenholtz, P. An expanded genomic representation of the phylum Cyanobacteria. *Genome Biol. Evol.* **2014**, *6*, 1031–1045. [[CrossRef](#)] [[PubMed](#)]

62. Puente-Sánchez, F.; Arce-Rodríguez, A.; Oggerin, M.; García-Villadangos, M.; Moreno-Paz, M.; Blanco, Y.; Rodríguez, N.; Bird, L.; Lincoln, S.A.; Tornos, F.; et al. Viable cyanobacteria in the deep continental subsurface. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 10702–10707. [[CrossRef](#)] [[PubMed](#)]
63. Sala-Comorera, L.; Blanch, A.R.; Casanovas-Massana, A.; Monleón-Getino, A.; García-Aljaro, C. Traceability of different brands of bottled mineral water during shelf life, using PCR-DGGE and next generation sequencing techniques. *Food Microbiol.* **2019**, *82*, 1–10. [[CrossRef](#)] [[PubMed](#)]
64. Poitelon, J.B.; Joyeux, M.; Welté, B.; Duguet, J.P.; Prestel, E.; Dubow, M.S. Variations of bacterial 16S rDNA phylotypes prior to and after chlorination for drinking water production from two surface water treatment plants. *J. Ind. Microbiol. Biotechnol.* **2010**, *37*, 117–128. [[CrossRef](#)] [[PubMed](#)]
65. Jalili, F.; Trigui, H.; Guerra Maldonado, J.F.; Dorner, S.; Zamyadi, A.; Shapiro, B.J.; Terrat, Y.; Fortin, N.; Sauvé, S.; Prévost, M. Can Cyanobacterial Diversity in the Source Predict the Diversity in Sludge and the Risk of Toxin Release in a Drinking Water Treatment Plant? *Toxins* **2021**, *13*, 25. [[CrossRef](#)] [[PubMed](#)]
66. Pei, H.; Xu, H.; Wang, J.; Jin, Y.; Xiao, H.; Ma, C.; Sun, J.; Li, H. 16S rRNA Gene Amplicon Sequencing Reveals Significant Changes in Microbial Compositions during Cyanobacteria-Laden Drinking Water Sludge Storage. *Environ. Sci. Technol.* **2017**, *51*, 12774–12783. [[CrossRef](#)] [[PubMed](#)]
67. Douterelo, I.; Husband, S.; Loza, V.; Boxall, J. Dynamics of biofilm regrowth in drinking water distribution systems. *Appl. Environ. Microbiol.* **2016**, *82*, 4155–4168. [[CrossRef](#)]
68. Egorov, A.I.; Birkenhauer Best, J.M.; Frebis, C.P.; Karapondo, M.S. Occurrence of *Aeromonas* spp. in a random sample of drinking water distribution systems in the USA. *J. Water Health* **2011**, *9*, 785–798. [[CrossRef](#)]
69. van der Wielen, P.W.J.J.; van der Kooij, D. Nontuberculous mycobacteria, fungi, and opportunistic pathogens in unchlorinated drinking water in the Netherlands. *Appl. Environ. Microbiol.* **2013**, *79*, 825–834. [[CrossRef](#)]
70. Collado, L.; Kasimir, G.; Perez, U.; Bosch, A.; Pinto, R.; Saucedo, G.; Huguet, J.M.; Jose, M. Occurrence and diversity of *Arcobacter* spp. along the Llobregat River catchment, at sewage effluents and in a drinking water treatment plant. *Water Res.* **2010**, *44*, 3696–3702. [[CrossRef](#)]
71. Boers, S.A.; Jansen, R.; Hays, J.P. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* **2019**, *38*, 1059–1070. [[CrossRef](#)] [[PubMed](#)]

Supplementary material**Monitoring bacterial community dynamics in a drinking water treatment plant: an integrative approach using metabarcoding and microbial indicators in large water volumes**

Anna Pinar-Méndez^{a,c,d*}, Owen S. Wangensteen^b, Kim Præbel^b, Belén Galofré^a, Javier Méndez^c, Anicet R. Blanch^{c,d}, Cristina Garcia-Aljaro^{c,d*}

^a Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, Barcelona, Spain

^b Norwegian College of Fishery Science, Faculty of Biosciences, Fisheries and Economics, UiT The Arctic University of Norway, Tromsø, Norway

^c Department of Genetics, Microbiology and Statistics, Faculty of Biology, University of Barcelona, Barcelona, Spain

^d The Water Research Institute, University of Barcelona, Barcelona, Spain

Table S1. The full sequence list of 96 barcodes and leading Ns used for 16S rRNA metabarcoding

Oligo name	Leading Ns	8-base tag	Forward primer	Sequence	Well position
F515-T01	NN	AACAA GCC	GTGCCAGCMGCCGC GGTAA	NNAACAAGCCGTGCCAGCMG CCGCGGTAA	A01
F515-T02	NNN	GGAAT GAG	GTGCCAGCMGCCGC GGTAA	NNNGGAATGAGGTGCCAGCM GCCGCGGTAA	A02
F515-T03	NNNN	AATTG CCG	GTGCCAGCMGCCGC GGTAA	NNNNAATTGCCGGTGCCAGC MGCCGCGGTAA	A03
F515-T04	NN	CGACC ATA	GTGCCAGCMGCCGC GGTAA	NNCGACCATAGTGCCAGCMG CCGCGGTAA	A04
F515-T05	NNN	ATGCT GAC	GTGCCAGCMGCCGC GGTAA	NNNATGCTGACGTGCCAGCM GCCGCGGTAA	A05
F515-T06	NNNN	TGAGA CAG	GTGCCAGCMGCCGC GGTAA	NNNNTGAGACAGGTGCCAGC MGCCGCGGTAA	A06
F515-T07	NN	GAGCT TAC	GTGCCAGCMGCCGC GGTAA	NNGAGCTTACGTGCCAGCMGC CGCGGTAA	A07
F515-T08	NNN	TTACC AGG	GTGCCAGCMGCCGC GGTAA	NNNTTACCAGGGTGCCAGCM GCCGCGGTAA	A08
F515-T09	NNNN	TGAGA GCT	GTGCCAGCMGCCGC GGTAA	NNNNTGAGAGCTGTGCCAGC MGCCGCGGTAA	A09
F515-T10	NN	CTGAC CTT	GTGCCAGCMGCCGC GGTAA	NNCTGACCTTGTGCCAGCMGC CGCGGTAA	A10
F515-T11	NNN	ATGCT TGG	GTGCCAGCMGCCGC GGTAA	NNNATGCTTGGGTGCCAGCMG CCGCGGTAA	A11
F515-T12	NNNN	AACAC CGT	GTGCCAGCMGCCGC GGTAA	NNNNAACACCGTGTGCCAGC MGCCGCGGTAA	A12
F515-T13	NN	TTACC GCT	GTGCCAGCMGCCGC GGTAA	NNTACCCTGTGCCAGCMGC CGCGGTAA	B01
F515-T14	NNN	CCAGT ATG	GTGCCAGCMGCCGC GGTAA	NNNCCAGTATGGTGCCAGCM GCCGCGGTAA	B02
F515-T15	NNNN	TGAGA TGC	GTGCCAGCMGCCGC GGTAA	NNNNTGAGATGCGTGCCAGC MGCCGCGGTAA	B03
F515-T16	NN	GTGCA ACT	GTGCCAGCMGCCGC GGTAA	NNGTGCAACTGTGCCAGCMGC CGCGGTAA	B04
F515-T17	NNN	ACAAC CGA	GTGCCAGCMGCCGC GGTAA	NNNACAACCGAGTGCCAGCM GCCGCGGTAA	B05
F515-T18	NNNN	TGAGC CTA	GTGCCAGCMGCCGC GGTAA	NNNNTGAGCCTAGTGCCAGC MGCCGCGGTAA	B06
F515-T19	NN	ATGGA GGT	GTGCCAGCMGCCGC GGTAA	NNATGGAGGTGTGCCAGCMG CCGCGGTAA	B07
F515-T20	NNN	TCATA CGC	GTGCCAGCMGCCGC GGTAA	NNNTCATACCGTGCCAGCMG CCGCGGTAA	B08
F515-T21	NNNN	CTGAG TCT	GTGCCAGCMGCCGC GGTAA	NNNNCTGAGTCTGTGCCAGCM GCCGCGGTAA	B09
F515-T22	NN	GAGGT GAA	GTGCCAGCMGCCGC GGTAA	NNGAGGTGAAGTGCCAGCMG CCGCGGTAA	B10
F515-T23	NNN	GGCAT GTA	GTGCCAGCMGCCGC GGTAA	NNNGGCATGTAGTGCCAGCM GCCGCGGTAA	B11
F515-T24	NNNN	GTGCC ATA	GTGCCAGCMGCCGC GGTAA	NNNNGTGCCATAGTGCCAGC MGCCGCGGTAA	B12
F515-T25	NN	TTACG CCA	GTGCCAGCMGCCGC GGTAA	NNTTACGCCAGTGCCAGCMGC CGCGGTAA	C01
F515-T26	NNN	TCATA GCG	GTGCCAGCMGCCGC GGTAA	NNNTCATAGCGGTGCCAGCM GCCGCGGTAA	C02
F515-T27	NNNN	TGAGG ACA	GTGCCAGCMGCCGC GGTAA	NNNNTGAGGACAGTGCCAGC MGCCGCGGTAA	C03
F515-T28	NN	AACAG GAG	GTGCCAGCMGCCGC GGTAA	NNAACAGGAGGTGCCAGCMG CCGCGGTAA	C04
F515-T29	NNN	GAGTA ACC	GTGCCAGCMGCCGC GGTAA	NNNGAGTAACCGTGCCAGCM GCCGCGGTAA	C05
F515-T30	NNNN	CAGCT CAT	GTGCCAGCMGCCGC GGTAA	NNNNCAGCTCATGTGCCAGCM GCCGCGGTAA	C06

F515-T31	NN	TGCTC CAA	GTGCCAGCMGCCGC GGTAA	NNTGCTCCAAGTGCCAGCMGC CGCGGTAA	C07
F515-T32	NNN	TCAGT CGA	GTGCCAGCMGCCGC GGTAA	NNNTCAGTCGAGTGCCAGCM GCCGCGGTAA	C08
F515-T33	NNNN	GGAGA AGA	GTGCCAGCMGCCGC GGTAA	NNNNGGAGAAGAGTGCCAGC MGCCGCGGTAA	C09
F515-T34	NN	GTGCT CAA	GTGCCAGCMGCCGC GGTAA	NNGTGCTCAAGTGCCAGCMGC CGCGGTAA	C10
F515-T35	NNN	ACAAG ACC	GTGCCAGCMGCCGC GGTAA	NNNACAAGACCGTGCCAGCM GCCGCGGTAA	C11
F515-T36	NNNN	CAGGA ACA	GTGCCAGCMGCCGC GGTAA	NNNNCAGGAACAGTGCCAGC MGCCGCGGTAA	C12
F515-T37	NN	GTGAT CTC	GTGCCAGCMGCCGC GGTAA	NNGTGATCTCGTGCCAGCMGC CGCGGTAA	D01
F515-T38	NNN	ACTTG GCT	GTGCCAGCMGCCGC GGTAA	NNNACTTGGCTGTGCCAGCMG CCGCGGTAA	D02
F515-T39	NNNN	CGATA CAC	GTGCCAGCMGCCGC GGTAA	NNNNCGATACACGTGCCAGC MGCCGCGGTAA	D03
F515-T40	NN	TCATC CTG	GTGCCAGCMGCCGC GGTAA	NNTCATCTGGTGCCAGCMGC CGCGGTAA	D04
F515-T41	NNN	CAGGC TAA	GTGCCAGCMGCCGC GGTAA	NNNCAGGCTAAGTGCCAGCM GCCGCGGTAA	D05
F515-T42	NNNN	TCATC GGT	GTGCCAGCMGCCGC GGTAA	NNNNTCATCGGTGTGCCAGCM GCCGCGGTAA	D06
F515-T43	NN	AGTTC CAC	GTGCCAGCMGCCGC GGTAA	NNAGTTCACGTGCCAGCMGC CGCGGTAA	D07
F515-T44	NNN	AACCA ACG	GTGCCAGCMGCCGC GGTAA	NNNAACCAACGGTGCCAGCM GCCGCGGTAA	D08
F515-T45	NNNN	CTGCG AAT	GTGCCAGCMGCCGC GGTAA	NNNNCTGCGAATGTGCCAGC MGCCGCGGTAA	D09
F515-T46	NN	GTGGT TAG	GTGCCAGCMGCCGC GGTAA	NNGTGGTTAGGTGCCAGCMGC CGCGGTAA	D10
F515-T47	NNN	AGAAC GTG	GTGCCAGCMGCCGC GGTAA	NNNAGAACGTGGTGCCAGCM GCCGCGGTAA	D11
F515-T48	NNNN	TCATG CCT	GTGCCAGCMGCCGC GGTAA	NNNNTCATGCCTGTGCCAGCM GCCGCGGTAA	D12
F515-T49	NN	GGATA GCA	GTGCCAGCMGCCGC GGTAA	NNGGATAGCAGTGCCAGCMG CCGCGGTAA	E01
F515-T50	NNN	CAGTA GAC	GTGCCAGCMGCCGC GGTAA	NNNCAGTAGACGTGCCAGCM GCCGCGGTAA	E02
F515-T51	NNNN	TAGCC ACT	GTGCCAGCMGCCGC GGTAA	NNNNTAGCCACTGTGCCAGCM GCCGCGGTAA	E03
F515-T52	NN	GAGGA CTA	GTGCCAGCMGCCGC GGTAA	NNGAGGACTAGTGCCAGCMG CCGCGGTAA	E04
F515-T53	NNN	AGAAG AGG	GTGCCAGCMGCCGC GGTAA	NNNAGAAGAGGGTGCCAGCM GCCGCGGTAA	E05
F515-T54	NNNN	CGATG AGT	GTGCCAGCMGCCGC GGTAA	NNNNCGATGAGTGCCAGC MGCCGCGGTAA	E06
F515-T55	NN	GTGTA GTC	GTGCCAGCMGCCGC GGTAA	NNGTGTAGTCGTGCCAGCMGC CGCGGTAA	E07
F515-T56	NNN	GAGTT CCT	GTGCCAGCMGCCGC GGTAA	NNNGAGTTCCTGTGCCAGCMG CCGCGGTAA	E08
F515-T57	NNNN	ACACA CAG	GTGCCAGCMGCCGC GGTAA	NNNNACACACAGGTGCCAGC MGCCGCGGTAA	E09
F515-T58	NN	AACCT AGC	GTGCCAGCMGCCGC GGTAA	NNAACCTAGCGTGCCAGCMG CCGCGGTAA	E10
F515-T59	NNN	ACACA GGT	GTGCCAGCMGCCGC GGTAA	NNNACACAGGTGTGCCAGCM GCCGCGGTAA	E11
F515-T60	NNNN	TAGAG CTG	GTGCCAGCMGCCGC GGTAA	NNNNTAGAGCTGGTGCCAGC MGCCGCGGTAA	E12
F515-T61	NN	AACCT CAG	GTGCCAGCMGCCGC GGTAA	NNAACCTCAGGTGCCAGCMG CCGCGGTAA	F01
F515-T62	NNN	GGATG ATC	GTGCCAGCMGCCGC GGTAA	NNNGGATGATCGTGCCAGCM GCCGCGGTAA	F02
F515-T63	NNNN	TATCT GGC	GTGCCAGCMGCCGC GGTAA	NNNNTATCTGGCGTGCCAGCM GCCGCGGTAA	F03

F515-T64	NN	CTGGT TGA	GTGCCAGCMGCCGC GGTAA	NNCTGGTTGAGTGCCAGCMGC CGCGGTAA	F04
F515-T65	NNN	TCCAA CAC	GTGCCAGCMGCCGC GGTAA	NNNTCCAACACGTGCCAGCM GCCGCGGTAA	F05
F515-T66	NNNN	GTGTG AAG	GTGCCAGCMGCCGC GGTAA	NNNNGTGTGAAGGTGCCAGC MGCCGCGGTAA	F06
F515-T67	NN	AACCT GCT	GTGCCAGCMGCCGC GGTAA	NNAACCTGCTGTGCCAGCMGC CGCGGTAA	F07
F515-T68	NNN	GGATG TCT	GTGCCAGCMGCCGC GGTAA	NNNGGATGTCTGTGCCAGCMG CCGCGGTAA	F08
F515-T69	NNNN	TCCAA GCA	GTGCCAGCMGCCGC GGTAA	NNNNTCCAAGCAGTGCCAGC MGCCGCGGTAA	F09
F515-T70	NN	AGAAT GCC	GTGCCAGCMGCCGC GGTAA	NNAGAATGCCGTGCCAGCMG CCGCGGTAA	F10
F515-T71	NNN	GTGTG TGT	GTGCCAGCMGCCGC GGTAA	NNNNGTGTGTGTGTGCCAGCMG CCGCGGTAA	F11
F515-T72	NNNN	CACGG ATA	GTGCCAGCMGCCGC GGTAA	NNNNCACGGATAGTGCCAGC MGCCGCGGTAA	F12
F515-T73	NN	CAACC TCA	GTGCCAGCMGCCGC GGTAA	NNCAACCTCAGTGCCAGCMGC CGCGGTAA	G01
F515-T74	NNN	GGATT CGA	GTGCCAGCMGCCGC GGTAA	NNNGGATTTCGAGTGCCAGCM GCCGCGGTAA	G02
F515-T75	NNNN	TTAGG CAC	GTGCCAGCMGCCGC GGTAA	NNNNTTAGGCACGTGCCAGC MGCCGCGGTAA	G03
F515-T76	NN	CGCAT AGA	GTGCCAGCMGCCGC GGTAA	NNCGCATAGAGTGCCAGCMG CCGCGGTAA	G04
F515-T77	NNN	CAGTT CTC	GTGCCAGCMGCCGC GGTAA	NNNCAGTTCTCGTGCCAGCMG CCGCGGTAA	G05
F515-T78	NNNN	AACGA GTG	GTGCCAGCMGCCGC GGTAA	NNNNAACGAGTGGTGCCAGC MGCCGCGGTAA	G06
F515-T79	NN	CTGTC AAC	GTGCCAGCMGCCGC GGTAA	NNCTGTCAACGTGCCAGCMGC CGCGGTAA	G07
F515-T80	NNN	TCCAC CTA	GTGCCAGCMGCCGC GGTAA	NNNTCCACCTAGTGCCAGCMG CCGCGGTAA	G08
F515-T81	NNNN	CGCAT CTT	GTGCCAGCMGCCGC GGTAA	NNNNCGCATCTTGTGCCAGCM GCCGCGGTAA	G09
F515-T82	NN	GATAC GCT	GTGCCAGCMGCCGC GGTAA	NNGATACGCTGTGCCAGCMGC CGCGGTAA	G10
F515-T83	NNN	GTGTT GGA	GTGCCAGCMGCCGC GGTAA	NNNGTGTGGAGTGCCAGCM GCCGCGGTAA	G11
F515-T84	NNNN	ACACT AGG	GTGCCAGCMGCCGC GGTAA	NNNNACACTAGGGTGCCAGC MGCCGCGGTAA	G12
F515-T85	NN	CGCCA ATT	GTGCCAGCMGCCGC GGTAA	NNCGCCAATTGTGCCAGCMGC CGCGGTAA	H01
F515-T86	NNN	TAGCA AGG	GTGCCAGCMGCCGC GGTAA	NNNTAGCAAGGGTGCCAGCM GCCGCGGTAA	H02
F515-T87	NNNN	GTCAC AGA	GTGCCAGCMGCCGC GGTAA	NNNNGTCACAGAGTGCCAGC MGCCGCGGTAA	H03
F515-T88	NN	TCCAG ATC	GTGCCAGCMGCCGC GGTAA	NNTCCAGATCGTGCCAGCMGC CGCGGTAA	H04
F515-T89	NNN	ACACT CCT	GTGCCAGCMGCCGC GGTAA	NNNACACTCCTGTGCCAGCMG CCGCGGTAA	H05
F515-T90	NNNN	GTAA CGG	GTGCCAGCMGCCGC GGTAA	NNNNGTTAACGGGTGCCAGC MGCCGCGGTAA	H06
F515-T91	NN	AACGC GAT	GTGCCAGCMGCCGC GGTAA	NNAACGCGATGTGCCAGCMG CCGCGGTAA	H07
F515-T92	NNN	CGCGT AAT	GTGCCAGCMGCCGC GGTAA	NNNCGCGTAATGTGCCAGCM GCCGCGGTAA	H08
F515-T93	NNNN	CCGAT ACT	GTGCCAGCMGCCGC GGTAA	NNNNCCGATACTGTGCCAGCM GCCGCGGTAA	H09
F515-T94	NN	AGACA TCC	GTGCCAGCMGCCGC GGTAA	NNAGACATCCGTGCCAGCMG CCGCGGTAA	H10
F515-T95	NNN	GATAG AGG	GTGCCAGCMGCCGC GGTAA	NNNGATAGAGGGTGCCAGCM GCCGCGGTAA	H11
F515-T96	NNNN	CTGTT GTG	GTGCCAGCMGCCGC GGTAA	NNNNCTGTTGTGGTGCCAGCM GCCGCGGTAA	H12

Oligo name	Leading Ns	8-base tag	Forward primer	Sequence	Well position
R806-T01	NNN N	AACAAG CC	GGACTACHVGGGTW TCTAAT	NNNNAACAAGCCGGACTACHVGG GTWTCTAAT	A01
R806-T02	NNN	GGAATG AG	GGACTACHVGGGTW TCTAAT	NNNGGAATGAGGGACTACHVGGG TWTCTAAT	A02
R806-T03	NN	AATTGC CG	GGACTACHVGGGTW TCTAAT	NNAATTGCCGGGACTACHVGGGT WTCTAAT	A03
R806-T04	NNN N	CGACCA TA	GGACTACHVGGGTW TCTAAT	NNNNCGACCATAGGACTACHVGG GTWTCTAAT	A04
R806-T05	NNN	ATGCTG AC	GGACTACHVGGGTW TCTAAT	NNNATGCTGACGGACTACHVGGG TWTCTAAT	A05
R806-T06	NN	TGAGAC AG	GGACTACHVGGGTW TCTAAT	NNTGAGACAGGGACTACHVGGGT WTCTAAT	A06
R806-T07	NNN N	GAGCTT AC	GGACTACHVGGGTW TCTAAT	NNNNGAGCTTACGGACTACHVGG GTWTCTAAT	A07
R806-T08	NNN	TTACCA GG	GGACTACHVGGGTW TCTAAT	NNNTTACCAGGGGACTACHVGGG TWTCTAAT	A08
R806-T09	NN	TGAGAG CT	GGACTACHVGGGTW TCTAAT	NNTGAGAGCTGGACTACHVGGGT WTCTAAT	A09
R806-T10	NNN N	CTGACC TT	GGACTACHVGGGTW TCTAAT	NNNNCTGACCTGGACTACHVGG GTWTCTAAT	A10
R806-T11	NNN	ATGCTT GG	GGACTACHVGGGTW TCTAAT	NNNATGCTTGGGGACTACHVGGG TWTCTAAT	A11
R806-T12	NN	AACACC GT	GGACTACHVGGGTW TCTAAT	NNAACACCGTGGACTACHVGGGT WTCTAAT	A12
R806-T13	NNN N	TTACCG CT	GGACTACHVGGGTW TCTAAT	NNNNTTACCGCTGGACTACHVGG GTWTCTAAT	B01
R806-T14	NNN	CCAGTA TG	GGACTACHVGGGTW TCTAAT	NNNCCAGTATGGGACTACHVGGG TWTCTAAT	B02
R806-T15	NN	TGAGAT GC	GGACTACHVGGGTW TCTAAT	NNTGAGATGCGGACTACHVGGGT WTCTAAT	B03
R806-T16	NNN N	GTGCAA CT	GGACTACHVGGGTW TCTAAT	NNNNGTGCAACTGGACTACHVGG GTWTCTAAT	B04
R806-T17	NNN	ACAACC GA	GGACTACHVGGGTW TCTAAT	NNNACAACCGAGGACTACHVGGG TWTCTAAT	B05
R806-T18	NN	TGAGCC TA	GGACTACHVGGGTW TCTAAT	NNTGAGCCTAGGACTACHVGGGT WTCTAAT	B06
R806-T19	NNN N	ATGGAG GT	GGACTACHVGGGTW TCTAAT	NNNNATGGAGGTGGACTACHVGG GTWTCTAAT	B07
R806-T20	NNN	TCATAC GC	GGACTACHVGGGTW TCTAAT	NNNTCATACGCGGACTACHVGGGT WTCTAAT	B08
R806-T21	NN	CTGAGT CT	GGACTACHVGGGTW TCTAAT	NNCTGAGTCTGGACTACHVGGGT WTCTAAT	B09
R806-T22	NNN N	GAGGT GAA	GGACTACHVGGGTW TCTAAT	NNNNGAGGTGAAGGACTACHVGG GTWTCTAAT	B10
R806-T23	NNN	GGCATG TA	GGACTACHVGGGTW TCTAAT	NNNGGCATGTAGGACTACHVGGG TWTCTAAT	B11

R806 -T24	NN	GTGCCA TA	GGACTACHVGGGTW TCTAAT	NNGTGCCATAGGACTACHVGGGT WTCTAAT	B12
R806 -T25	NNN N	TTACGC CA	GGACTACHVGGGTW TCTAAT	NNNNTTACGCCAGGACTACHVGG GTWTCTAAT	C01
R806 -T26	NNN	TCATAG CG	GGACTACHVGGGTW TCTAAT	NNNTCATAGCGGGACTACHVGGG TWTCTAAT	C02
R806 -T27	NN	TGAGGA CA	GGACTACHVGGGTW TCTAAT	NNTGAGGACAGGACTACHVGGGT WTCTAAT	C03
R806 -T28	NNN N	AACAGG AG	GGACTACHVGGGTW TCTAAT	NNNNAACAGGAGGGACTACHVGG GTWTCTAAT	C04
R806 -T29	NNN	GAGTAA CC	GGACTACHVGGGTW TCTAAT	NNNGAGTAACCGGACTACHVGGG TWTCTAAT	C05
R806 -T30	NN	CAGCTC AT	GGACTACHVGGGTW TCTAAT	NNCAGCTCATGGACTACHVGGGT WTCTAAT	C06
R806 -T31	NNN N	TGCTCC AA	GGACTACHVGGGTW TCTAAT	NNNNTGCTCCAAGGACTACHVGG GTWTCTAAT	C07
R806 -T32	NNN	TCAGTC GA	GGACTACHVGGGTW TCTAAT	NNNTCAGTCGAGGACTACHVGGG TWTCTAAT	C08
R806 -T33	NN	GGAGA AGA	GGACTACHVGGGTW TCTAAT	NNGGAGAAGAGGACTACHVGGGT WTCTAAT	C09
R806 -T34	NNN N	GTGCTC AA	GGACTACHVGGGTW TCTAAT	NNNNGTGCTCAAGGACTACHVGG GTWTCTAAT	C10
R806 -T35	NNN	ACAAGA CC	GGACTACHVGGGTW TCTAAT	NNNACAAGACCGGACTACHVGGG TWTCTAAT	C11
R806 -T36	NN	CAGGAA CA	GGACTACHVGGGTW TCTAAT	NNCAGGAACAGGACTACHVGGGT WTCTAAT	C12
R806 -T37	NNN N	GTGATC TC	GGACTACHVGGGTW TCTAAT	NNNNGTGATCTCGGACTACHVGG GTWTCTAAT	D01
R806 -T38	NNN	ACTTGG CT	GGACTACHVGGGTW TCTAAT	NNNACTTGGCTGGACTACHVGGGT WTCTAAT	D02
R806 -T39	NN	CGATAC AC	GGACTACHVGGGTW TCTAAT	NNCGATACACGGACTACHVGGGT WTCTAAT	D03
R806 -T40	NNN N	TCATCC TG	GGACTACHVGGGTW TCTAAT	NNNNTCATCCTGGGACTACHVGG GTWTCTAAT	D04
R806 -T41	NNN	CAGGCT AA	GGACTACHVGGGTW TCTAAT	NNNCAGGCTAAGGACTACHVGGG TWTCTAAT	D05
R806 -T42	NN	TCATCG GT	GGACTACHVGGGTW TCTAAT	NNTCATCGGTGGACTACHVGGGT WTCTAAT	D06
R806 -T43	NNN N	AGTTCC AC	GGACTACHVGGGTW TCTAAT	NNNNAGTTCCACGGACTACHVGG GTWTCTAAT	D07
R806 -T44	NNN	AACCAA CG	GGACTACHVGGGTW TCTAAT	NNNAACCAACGGGACTACHVGGG TWTCTAAT	D08
R806 -T45	NN	CTGCGA AT	GGACTACHVGGGTW TCTAAT	NNCTGCGAATGGACTACHVGGGT WTCTAAT	D09
R806 -T46	NNN N	GTGGTT AG	GGACTACHVGGGTW TCTAAT	NNNNGTGGTTAGGACTACHVGG GTWTCTAAT	D10
R806 -T47	NNN	AGAACG TG	GGACTACHVGGGTW TCTAAT	NNNAGAACGTGGGACTACHVGGG TWTCTAAT	D11
R806 -T48	NN	TCATGC CT	GGACTACHVGGGTW TCTAAT	NNTCATGCCTGGACTACHVGGGT WTCTAAT	D12

R806 -T49	NNN N	GGATAG CA	GGACTACHVGGGTW TCTAAT	NNNNGGATAGCAGGACTACHVGG GTWTCTAAT	E01
R806 -T50	NNN	CAGTAG AC	GGACTACHVGGGTW TCTAAT	NNNCAGTAGACGGACTACHVGGG TWTCTAAT	E02
R806 -T51	NN	TAGCCA CT	GGACTACHVGGGTW TCTAAT	NNTAGCCACTGGACTACHVGGGT WTCTAAT	E03
R806 -T52	NNN N	GAGGA CTA	GGACTACHVGGGTW TCTAAT	NNNNGAGGACTAGGACTACHVGG GTWTCTAAT	E04
R806 -T53	NNN	AGAAGA GG	GGACTACHVGGGTW TCTAAT	NNNAGAAGAGGGGACTACHVGGG TWTCTAAT	E05
R806 -T54	NN	CGATGA GT	GGACTACHVGGGTW TCTAAT	NNCGATGAGTGGACTACHVGGGT WTCTAAT	E06
R806 -T55	NNN N	GTGTAG TC	GGACTACHVGGGTW TCTAAT	NNNNGTGTAGTCGGACTACHVGG GTWTCTAAT	E07
R806 -T56	NNN	GAGTTC CT	GGACTACHVGGGTW TCTAAT	NNNGAGTTCCTGGACTACHVGGGT WTCTAAT	E08
R806 -T57	NN	ACACAC AG	GGACTACHVGGGTW TCTAAT	NNACACACAGGGACTACHVGGGT WTCTAAT	E09
R806 -T58	NNN N	AACCTA GC	GGACTACHVGGGTW TCTAAT	NNNNAACCTAGCGGACTACHVGG GTWTCTAAT	E10
R806 -T59	NNN	ACACAG GT	GGACTACHVGGGTW TCTAAT	NNNACACAGGTGGACTACHVGGG TWTCTAAT	E11
R806 -T60	NN	TAGAGC TG	GGACTACHVGGGTW TCTAAT	NNTAGAGCTGGGACTACHVGGGT WTCTAAT	E12
R806 -T61	NNN N	AACCTC AG	GGACTACHVGGGTW TCTAAT	NNNNAACCTCAGGGACTACHVGG GTWTCTAAT	F01
R806 -T62	NNN	GGATGA TC	GGACTACHVGGGTW TCTAAT	NNNGGATGATCGGACTACHVGGG TWTCTAAT	F02
R806 -T63	NN	TATCTG GC	GGACTACHVGGGTW TCTAAT	NNTATCTGGCGGACTACHVGGGT WTCTAAT	F03
R806 -T64	NNN N	CTGGTT GA	GGACTACHVGGGTW TCTAAT	NNNNCTGGTTGAGGACTACHVGG GTWTCTAAT	F04
R806 -T65	NNN	TCCAAC AC	GGACTACHVGGGTW TCTAAT	NNNTCCAACACGGACTACHVGGGT WTCTAAT	F05
R806 -T66	NN	GTGTGA AG	GGACTACHVGGGTW TCTAAT	NNGTGTGAAGGGACTACHVGGGT WTCTAAT	F06
R806 -T67	NNN N	AACCTG CT	GGACTACHVGGGTW TCTAAT	NNNNAACCTGCTGGACTACHVGG GTWTCTAAT	F07
R806 -T68	NNN	GGATGT CT	GGACTACHVGGGTW TCTAAT	NNNGGATGTCTGGACTACHVGGG TWTCTAAT	F08
R806 -T69	NN	TCCAAG CA	GGACTACHVGGGTW TCTAAT	NNTCCAAGCAGGACTACHVGGGT WTCTAAT	F09
R806 -T70	NNN N	AGAATG CC	GGACTACHVGGGTW TCTAAT	NNNNAGAATGCCGGACTACHVGG GTWTCTAAT	F10
R806 -T71	NNN	GTGTGT GT	GGACTACHVGGGTW TCTAAT	NNNGTGTGTGTGGACTACHVGGG TWTCTAAT	F11
R806 -T72	NN	CACGGA TA	GGACTACHVGGGTW TCTAAT	NNCACGGATAGGACTACHVGGGT WTCTAAT	F12
R806 -T73	NNN N	CAACCT CA	GGACTACHVGGGTW TCTAAT	NNNNCAACCTCAGGACTACHVGG GTWTCTAAT	G01

R806 -T74	NNN	GGATTC GA	GGACTACHVGGGTW TCTAAT	NNNGGATTCGAGGACTACHVGGG TWTCTAAT	G02
R806 -T75	NN	TTAGGC AC	GGACTACHVGGGTW TCTAAT	NNTTAGGCACGGACTACHVGGGT WTCTAAT	G03
R806 -T76	NNN N	CGCATA GA	GGACTACHVGGGTW TCTAAT	NNNCGCATAGAGGACTACHVGG GTWTCTAAT	G04
R806 -T77	NNN	CAGTTC TC	GGACTACHVGGGTW TCTAAT	NNNCAGTTCTCGGACTACHVGGGT WTCTAAT	G05
R806 -T78	NN	AACGAG TG	GGACTACHVGGGTW TCTAAT	NNAACGAGTGGGACTACHVGGGT WTCTAAT	G06
R806 -T79	NNN N	CTGTCA AC	GGACTACHVGGGTW TCTAAT	NNNNCTGTCAACGGACTACHVGG GTWTCTAAT	G07
R806 -T80	NNN	TCCACC TA	GGACTACHVGGGTW TCTAAT	NNNTCCACCTAGGACTACHVGGGT WTCTAAT	G08
R806 -T81	NN	CGCATC TT	GGACTACHVGGGTW TCTAAT	NNCGCATCTGGACTACHVGGGT WTCTAAT	G09
R806 -T82	NNN N	GATACG CT	GGACTACHVGGGTW TCTAAT	NNNNGATACGCTGGACTACHVGG GTWTCTAAT	G10
R806 -T83	NNN	GTGTTG GA	GGACTACHVGGGTW TCTAAT	NNNGTGTGGAGGACTACHVGGG TWTCTAAT	G11
R806 -T84	NN	ACACTA GG	GGACTACHVGGGTW TCTAAT	NNACACTAGGGGACTACHVGGGT WTCTAAT	G12
R806 -T85	NNN N	CGCCAA TT	GGACTACHVGGGTW TCTAAT	NNNCGCCAATTGGACTACHVGG GTWTCTAAT	H01
R806 -T86	NNN	TAGCAA GG	GGACTACHVGGGTW TCTAAT	NNNTAGCAAGGGGACTACHVGGG TWTCTAAT	H02
R806 -T87	NN	GTCACA GA	GGACTACHVGGGTW TCTAAT	NGTCACAGAGGACTACHVGGGT WTCTAAT	H03
R806 -T88	NNN N	TCCAGA TC	GGACTACHVGGGTW TCTAAT	NNNNTCCAGATCGGACTACHVGG GTWTCTAAT	H04
R806 -T89	NNN	ACACTC CT	GGACTACHVGGGTW TCTAAT	NNNACACTCCTGGACTACHVGGGT WTCTAAT	H05
R806 -T90	NN	GTTAAC GG	GGACTACHVGGGTW TCTAAT	NGTTAACGGGGACTACHVGGGT WTCTAAT	H06
R806 -T91	NNN N	AACGCG AT	GGACTACHVGGGTW TCTAAT	NNNNAACGCGATGGACTACHVGG GTWTCTAAT	H07
R806 -T92	NNN	CGCGTA AT	GGACTACHVGGGTW TCTAAT	NNNCGCGTAATGGACTACHVGGG TWTCTAAT	H08
R806 -T93	NN	CCGATA CT	GGACTACHVGGGTW TCTAAT	NNCCGATACTGGACTACHVGGGT WTCTAAT	H09
R806 -T94	NNN N	AGACAT CC	GGACTACHVGGGTW TCTAAT	NNNNAGACATCCGGACTACHVGG GTWTCTAAT	H10
R806 -T95	NNN	GATAGA GG	GGACTACHVGGGTW TCTAAT	NNNGATAGAGGGGACTACHVGGG TWTCTAAT	H11
R806 -T96	NN	CTGTTG TG	GGACTACHVGGGTW TCTAAT	NNCTGTTGTGGGACTACHVGGGT WTCTAAT	H12

Table S2. Overview of sampling campaigns, climatological data, and some basic operating conditions (percentages of each source intake and percentage of water treated by conventional or advanced treatment) at the DWTP

Sampling ID	Collection date	Temperature (°C)	Humidity (%)	Rainfall (mm/24h)	Percentage intake sources (%)		Percentage water treated by (%)		River flow rate (m ³ /s)
					RW	GW	Conventional	Advanced	
N1	16/04/2018	16.6	66	0.4	53	47	76	24	135
N2	18/06/2018	26.3	52	0	88	12	70	30	29.61
N3	24/07/2018	29.5	64	0	88	12	66	34	10.37
N4	28/08/2018	28.4	71	0	100	0	75	25	13.29
N5	07/01/2019	11.5	44	0	85	15	80	20	25.5
N6	26/02/2019	14.3	63	0	86	14	80	20	5.39
N7	25/03/2019	18.8	51	0	83	17	64	36	6.92
N8	27/05/2019	21.9	59	2.9	99.75	0.25	63	37	9.4

Table S3. Overview of physicochemical data gathered at the different stages of DWTP during sampling campaigns: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post chlorination water (DW). Data are presented as mean and standard deviation values (mean±sd) or NA (not analyzed) where data was not monitored.

DWTP stage	Sample temperature (°C)	Conductivity (µS/cm)	Ammonium (mg NH ₄ ⁺ /L)	TOC (mg C/L)	pH	Turbidity (NTU)
GW	18 ± 1.5	1508 ± 67	0.05 ± 0.03	NA	NA	0.12 ± 0.13
RW	15 ± 5.24	1116 ± 301	0.17 ± 0.16	3.55 ± 0.54	8.11 ± 0.17	75.87±52.87
DEC	NA	NA	NA	NA	NA	1.56 ± 0.94
SF	NA	NA	NA	NA	NA	0.28 ± 0.17
OZ	NA	NA	NA	NA	NA	NA
CF	18 ± 5.1	1369 ± 282	NA	NA	7.2 ± 0.12	NA
RO	NA	165 ± 40	NA	NA	NA	NA
MIX	16 ± 5	1008 ± 190	0.05 ± 0.04	NA	7.42 ± 0.21	NA
DW	18 ± 6	1017 ± 251	0.05 ± 0.04	1.37 ± 0.32	7.4 ± 0.18	0.09 ± 0.03

Table S4. PERMANOVA analysis based on Bray–Curtis dissimilarities of ASVs abundances from DWTP for temperature variable

	Df	Sum Sq	Mean Sq	F	N.Perm	R2	Pr(>F)
Temperature	1	1.2898	1.28977	3.248	999	0.04434	0.001
Residuals	70	27.7966	0.39709			0.95566	
total	71	29.0864				1	

Table S5. List of potential pathogens genera and percentages of reads detected along DWTP.

Genus	Potential group risk	GW	RW	DEC	SF	OZ	CF	RO	MIX	DW
<i>Arcobacter</i>	Human/animal pathogen	ND	1%	3%	2.00%	0.07%	0.07%	0.07%	0.04%	0.007%
<i>Aeromonas</i>	Human/animal pathogen	0.02%	0.4%	0.2%	0.3%	2%	3%	0.8%	4%	2%
<i>Campylobacter</i> (Campylobacterales)	Human/animal pathogen	0.4%	5%	3%	3%	0.09%	0.08%	0.20%	0.05%	0.01%
<i>Clostridium</i>	Human/animal pathogen	ND	0.002%	0.03%	0.008%	1%	0.02%	0.2%	0.02%	2%
<i>Escherichia-Shigella</i>	Human/animal pathogen	ND	0.003%	0.007%	0.009%	ND	ND	ND	ND	ND
<i>Enterococcus</i>	Human/animal pathogen	ND	ND	ND	0.01	0.002%	0.002%	ND	ND	ND
<i>Helicobacter</i>	Human pathogen	ND	ND	ND	ND	ND	ND	ND	ND	ND
<i>Legionella</i>	Human pathogen	0.4%	0.05%	0.4%	0.6%	0.6%	1%	0.6%	2%	0.4%
<i>Leptospira</i>	Human/animal pathogen	ND	ND	ND	ND	ND	ND	ND	ND	ND
<i>Mycobacterium</i>	Human/animal pathogen	0.02%	ND	0.02%	0.5%	0.3%	0.02%	0.009%	0.003%	0.2%
<i>Pseudomonas</i>	Human pathogen	1%	3%	2%	3%	8%	2%	5%	3%	1%
<i>Vibrio</i>	Human/animal pathogen	ND	ND	ND	ND	ND	ND	ND	ND	ND

DWTP stages: Groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post chlorination water (DW). ND (not detected)

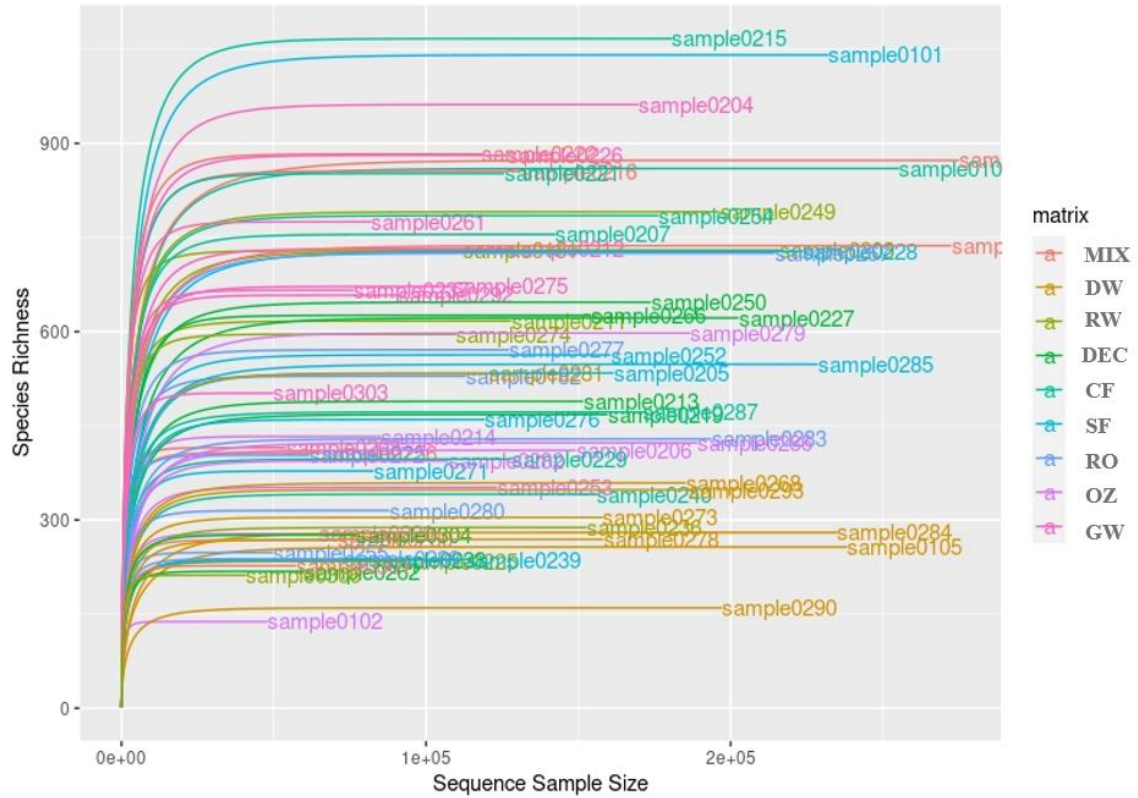


Figure S1. Rarefaction curves of all the samples analyzed in this study. All samples showed an initial exponential growth and then flatten curves indicating a good coverage of microbial communities. Groundwater (GW), river water (RW), Decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post chlorination water (DW).

For **interactive version of figures S2, S3 and S4**, the reader is referred to the **web version** of this article.
The following supporting information can be downloaded at :

<https://www.mdpi.com/article/10.3390/w14091435/s1>

3. CAPITOL 3

Article 3: Comparació de l'avaluació de la diversitat bacteriana mitjançant metabarcoding del gen 16S ARNr i MALDI-TOF MS a una estació de tractament d'aigua potable a gran escala

“Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant”

Anna Pinar-Méndez, Belén Galofré, Anicet R. Blanch, Cristina Garcia-Aljaro.

Water Research, under review (manuscript number WR 69746), sotmès el 19/5/2022

La potabilització impacta positivament en la qualitat de l'aigua reduint la càrrega de matèria orgànica i els microorganismes presents fent-la apta pel consum humà. Malgrat els avenços en la legislació vigent de la qualitat de l'aigua, es continua únicament amb el control d'indicadors bacterians de contaminació fecal i en el recompte de bacteris heteròtrofs (HPC), mentre que s'omet la composició del microbioma de l'aigua. Els bacteris heteròtrofs contribueixen de manera natural en el tractament de l'aigua reduint la matèria orgànica i els contaminants, no obstant també poden proliferar induint un deteriorament de l'aigua o fins i tot esdevenir un risc per a la salut. Amb el que resulta de gran importància conèixer la diversitat bacteriana en els processos de potabilització per tenir la imatge completa.

L'objectiu d'aquest estudi es va centrar en la caracterització de la diversitat bacteriana cultivable a l'ETAP de Sant Joan Despí, mitjançant la identificació d'HPC per MALDI-TOF MS, i la comparació amb els bacteris totals identificats per *metabarcoding* del gen 16S ARNr (mètode molecular no dependent de cultiu) obtinguda a l'article 2.

Es van prendre un total de 63 mostres al llarg d'un any corresponents a 9 processos de potabilització: captació (pous i riu), pretractament (decantació i filtre de sorra), tractament convencional (ozonització i filtre de carbó), tractament avançat (osmosi inversa) i aigua final (cambra de mescla i aigua potable clorada).

Les mostres d'aigua es van processar de manera diferent i amb diferents volums d'acord amb la seva càrrega microbiana: per filtració de membrana (riu i pretractament; fins a 2 L) o per ultrafiltració amb Rexeed (tractament convencional, avançat i aigua final; de 100-1.100 L), i es van analitzar els paràmetres legiscats (coliforms totals, *E. coli*, *C. perfringens*, enterococs intestinals i colífags somàtics) per comprovar la qualitat microbiològica de totes les mostres. Amb les plaques de HPC, a més de fer el recompte, es van aïllar un total de 1.807 soques (~30

per mostra) i es van identificar per MALDI-TOF MS fent ús de la llibreria DWL (Pinar-Méndez et al., 2021). Mentre que, per a la comparació amb les dades obtingudes per metabarcoding, es van re-analitzar les seqüències obtingudes d'un treball previ corresponent a les mateixes mostres (Pinar-Méndez et al., 2022).

Els recomptes d'HPC van mostrar resultats positius al llarg de tots els tractaments, amb un rang de valors de mitjana des de 4,62 log₁₀ CFU/mL a la captació de riu, fins a una forta i significativa disminució a 0,027 log₁₀ CFU/mL a l'aigua potable clorada. Es va observar una variabilitat estacional en relació a la temperatura ambiental durant del mostreig, que es va traduir amb valors significativament més alts al riu, ozonització i aigua potable clorada en temperatures elevades (>22°C).

La caracterització per MALDI-TOF MS va mostrar un predomini de Proteobacteria i Bacteroidota a tots els processos de tractament amb excepció de l'aigua potable clorada on el fílum majoritari va ser Firmicutes. A nivell de gènere es van identificar 57 gèneres diferents a tota l'ETAP amb diferent distribució a cada etapa. *Aeromonas* i *Pseudomonas* van ser els gèneres més abundants a pous i riu respectivament, mentre que el pretractament indueix canvis essent *Flavobacterium* el gènere majoritari observat a decantació i filtre per sorra, el qual està poc representat a la captació. L'ozonització també produeix canvis importants amb un increment de *Bacillus*, mentre que un nou canvi es va observar rere la filtració per carbó amb una dominància de *Chryseobacterium*. El tractament avançat va presentar un perfil diferent al tractament convencional amb *Comamonas* com el gènere més abundant. La cambra de mescla va presentar una composició de gèneres similars al filtre de carbó, no obstant rere la cloració es va detectar el canvi més significatiu on *Bacillus* va dominar a l'aigua potable clorada. Es va poder observar una variació estacional relacionada amb temperatura ambiental elevada. La temperatura té un efecte modulador en la dinàmica microbiana que pot alterar la composició de les comunitats microbianes. En aquest estudi, l'abundància de *Chryseobacterium* al filtre de carbó a altes temperatures era molt més elevat que a baixes temperatures.

Malgrat l'ús combinat de la llibreria original de l'equip Bruker BDAL més la DWL, un 38% de les soques (686) no es van poder identificar. Els espectres sense classificar es van re-processar per analitzar les seves similituds amb finalitats d'agrupament (clústers) amb una resolució màxima a nivell de gènere. Això va permetre agrupar les 565 soques desconegudes en un total de 91 clústers, mentre que la resta (121) corresponien a soques amb un perfil proteic únic. L'anàlisi de les soques no identificades va permetre desglossar aquesta fracció observant la seva abundància, fluctuació i variacions estacionals d'alguns grups al llarg de la planta. Addicionalment, va permetre calcular la diversitat bacteriana nivell de gènere per a avaluar conjuntament amb la fracció de soques identificades. Es va observar una elevada diversitat

variable al llarg del tractament, des d'una gran diversitat a riu i pous, fins a una disminució estadísticament significativa rere la cloració a l'aigua potable. La influència de la temperatura ambiental va mostrar un petit augment de la diversitat (no estadísticament significatiu) en temperatures elevades per a les etapes de decantació, filtració per sorra, ozonització, filtració per carbó i cambra de mescla, mentre que per a riu, osmosi i aigua potable es va registrar un lleuger augment en temperatures baixes, i la diversitat a les mostres de pous es van mantenir estable envers la temperatura ambiental.

Els resultats de l'estudi de diversitat bacteriana d'heteròtrofs cultivables per MALDI-TOF es van comparar amb l'estudi previ de caracterització de comunitats bacterianes per metabarcoding. Ambdós van observar diferències en taxonomia i diversitat tot i presentar recomptes de HPC sense canvis significatius a les mostres d'aigua.

Es va observar una elevada diversitat bacteriana per ambdós mètodes no obstant, per MALDI es va observar un índex de diversitat més baix (interval d'índex de Shannon: 0,38 - 2,88) en comparació amb metabarcoding (interval d'índex de Shannon: 3,3 - 5,2) probablement degut a la diferència en les poblacions dianes analitzades, bacteris cultivables (MALDI) en comparació amb els bacteris totals (*metabarcoding*). Les tendències de diversitat entre ambdues tècniques coincidien en una major diversitat a riu i pous i baixa diversitat a aigua potable, però presentaven algunes diferències, per exemple: MALDI va mostrar un augment en la diversitat a l'ozonització respecte el filtre de sorra, en canvi disminuïa segons *metabarcoding*.

A nivell taxonòmic, les dues tècniques van mostrar un perfil de fílums similars, amb domini de Proteobacteria al llarg dels tractaments excepte a l'aigua potable, on la cloració marca una forta pressió selectiva que redueix la diversitat i produeix un canvi en el fílum majoritari de l'aigua final, essent Cyanobacteria (segons metabarcoding) o Firmicutes (segons MALDI). Per a avaluar la capacitat d'ambdós tècniques a un nivell taxonòmic més baix (gènere), es va comparar la detecció i fluctuació d'aquests al llarg dels diferents tractaments i es van observar similituds i diferències. En alguns casos, determinats gèneres només van ser detectats per MALDI-TOF o per metabarcoding, amb excepció dels gèneres més abundants de cada etapa segons MALDI-TOF, que van ser també detectats per metabarcoding, però amb abundàncies diferents. Per exemple, *Bacillus* va presentar un 58% d'abundància a l'aigua potable clorada segons MALDI-TOF, però només 0,05% segons metabarcoding, ja que segons aquesta tècnica el gènere majoritari corresponia a *Obscuribacteraceae* (31%), no detectat per MALDI-TOF. Tanmateix, atenent a un criteri d'absència/presència, es va comparar la robustesa de les dues tècniques en la detecció dels 57 gèneres diferents a la planta segons MALDI-TOF. Es va observar que en 41 ocasions la tècnica MALDI-TOF MS va detectar la presència d'alguns gèneres no detectats per metabarcoding, mentre que en 152 ocasions, va succeir el contrari. Gràcies a l'anàlisi per

metabarcoding es va poder detectar una major varietat de gèneres a l'aigua potable agrupats dins de Gammaproteobacteria (grup potencial formador de biofilms) i no detectats per MALDI. Tanmateix, per a conèixer les localitzacions o punts calents on es troben els 6 gèneres majoritaris en gran abundància (>5%) segons cada tècnica, es va traçar la seva localització a l'ETAP. Així, per exemple, es va poder comprovar que ambdós aproximacions coincidien en que *Flavobacterium* i *Pseudomonas* són gèneres amb elevada abundància a la planta, però els punts calents diferien en algunes etapes. Per exemple, segons metabarcoding, *Flavobacterium* es trobava en gran abundància en l'ozonització i la filtració per carbó, però no hi és segons MALDI, i *Pseudomonas* només es va detectar per metabarcoding a l'aigua clorada, però no per MALDI-TOF. La resta de gèneres amb gran abundància a tota la planta diferien per ambdós tècniques, essent també *Bacillus*, *Aeromonas*, *Chryseobacterium* i *Acidovorax* les majoritàries per MALDI-TOF, i *Limnohabitans*, i gèneres no descrits agrupats dins de Chloroplast i Obscuribacteraceae per metabarcoding. Aquest últim va presentar una gran abundància a l'aigua potable segons metabarcoding, mentre que segons MALDI el gènere dominant va ser *Bacillus*, però cal tenir en compte que Cyanobacteria presenta gairebé un 50% de soques no cultivables fins el moment, per tant podria ser que no es detectin mitjançant tècniques basades en cultius com ara MALDI-TOF.

L'estudi de les comunitats bacterianes heteròtrofes cultivables mitjançant l'ús de MALDI-TOF MS suposa una avantatge en la determinació ràpida i fiable de la presència de certs gèneres. L'anàlisi requereix bacteris viables, i per tant, aporta una informació valuosa ja que aquests tenen capacitat per créixer i colonitzar o alterar diferents sistemes d'aigües, però el seu poder d'identificació depèn directament de la base de dades. En aquest estudi, derivat de l'esforç de la creació de la DWL, s'han obtingut bons resultats d'identificació a totes les matrius, en especial a l'aigua potable, presentant els valors més baixos de soques no resoltes. El seu poder de resolució ha permès identificar el 30% dels aïllats a nivell d'espècie i el 32% a nivell de gènere, mentre que per metabarcoding només el 36% de les seqüències es van poder assignar a nivell de gènere, i la resta a taxons superiors, a causa de limitació de la petita mida del fragment d'ADN seqüenciat. No obstant, la creació d'una base de dades per MALDI-TOF és laboriosa, i requereix l'ús d'altres tècniques d'identificació per a caracteritzar les soques. A banda, els bacteris cultivables tenen l'inconvenient de representar una fracció petita de la totalitat de la microbiota de l'aigua, amb el que, el mètode 16S rRNA metabarcoding permet superar aquesta limitació proporcionant una visió més detallada de la totalitat de les comunitats bacterianes, però els seus resultats poden estar influenciats pels protocols d'extracció i processament bioinformàtic, i no és adequada per diferenciar viabilitat, amb el que dificulta interpretar per si sola els perills derivats de la presència de determinats taxons.

Com a conclusions principals, les tècniques de MALDI-TOF MS i metabarcoding han permès observar fluctuacions entre etapes, estacionalitat segons la temperatura ambiental, i una gran diversitat a l'ETAP, sent més elevada per *metabarcoding* en comparació amb MALDI. Ambdós metodologies van permetre descriure els diferents gèneres de cada tractament observant similituds i diferències, on alguns gèneres només es van detectar mitjançant una metodologia o les dues. No obstant, ambdues tècniques coincideixen amb un domini de Proteobacteria en les etapes de tractament i es diferencien principalment en la composició de l'aigua potable final (Firmicutes segons MALDI-TOF, i Cyanobacteria segons metabarcoding), així com en la identificació de taxons inferiors a fílum.

En general, la composició microbiana als sistemes d'aigua és complexa d'interpretar amb un sol enfocament, o només basant-se en els indicadors microbians regulats. Per tant, és de gran importància combinar eines addicionals a la Directiva d'aigües de consum, com MALDI-TOF MS i metabarcoding del gen 16S ARNr, que aporten informació complementària i en profunditat de les comunitats bacterianes que poden influir en la qualitat de l'aigua més enllà de la informació d'absència o presència d'indicadors fecals o recompte de HPC, amb el que una visió holística pot ajudar a millorar la gestió de la qualitat de l'aigua.

Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant

Anna Pinar-Méndez^{a,b,c}, Belén Galofré^a, Anicet R. Blanch^{b,c}, Cristina García-Aljaro^{b,c}

^a Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, Barcelona, Spain

^b Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

^c Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona (UB), Barcelona, Spain

***Corresponding authors:**

Dr. Cristina García Aljaro

Mailing address: crgarcia@ub.edu

Avinguda Diagonal, 643, 08028 Barcelona, Spain

Phone: (+34) 934 02 14 84

Ms. Anna Pinar Méndez

Mailing address: apinar@aiguesdebarcelona.cat

General Batet 1-7, 08028 Barcelona, Spain

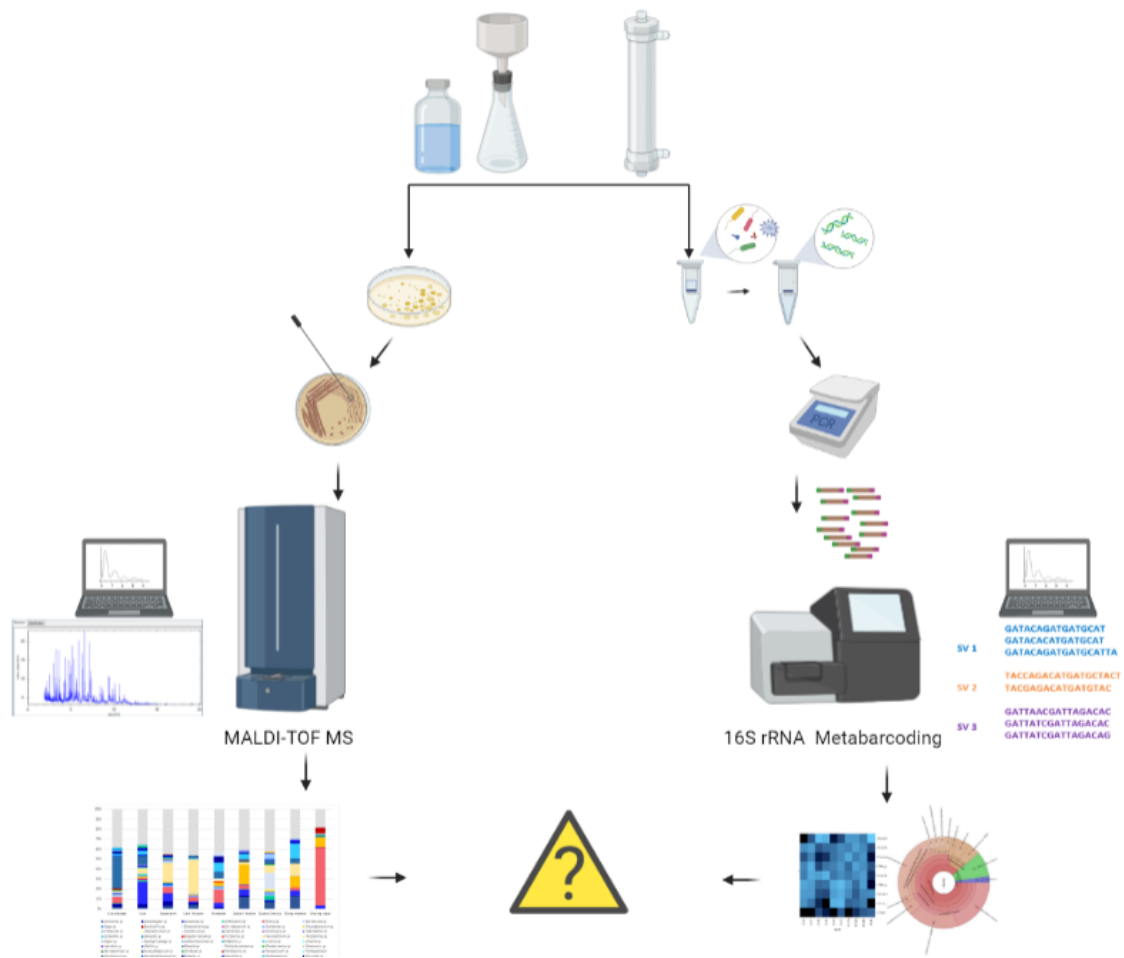
Phone: (+34) 93 247 98 12

Abstract

Bacterial communities in a full-scale drinking water treatment plant (DWTP) were characterized using matrix-assisted laser desorption/ionization time of flight mass-spectrometry (MALDI-TOF MS) to identify HPC isolates and the obtained results were compared to 16S rRNA (V4) metabarcoding data acquired in a previous study. Sixty-three samples were collected at nine stages of the potabilization process: river water and groundwater intake, decantation, sand filtration, ozonization, carbon filtration, reverse osmosis, the mixing chamber and post-chlorination drinking water. In total, 1,807 bacterial colonies were isolated, 32% of which were successfully identified to at least the genus level by MALDI-TOF MS using our previously developed Drinking Water Library. Trends in diversity were similar by both approaches, but differences were observed in the detection of taxa, especially at lower hierarchy levels. High bacterial diversity was observed in river and groundwater, where Proteobacteria predominated. The diversity decreased significantly after the chlorination step, where *Bacillus* sp. (Firmicutes) and an unknown genus of Obscuribacteraceae (Cyanobacteria) were the most prevalent genera according to MALDI-TOF MS and metabarcoding, respectively. The two approaches gave similar results for the decantation, sand filtration and mixing chamber steps, where the most abundant taxon was *Flavobacterium*. The combined use of these culture-based and culture-independent methods to characterize microbial populations may help to better understand the role of bacteria in water treatment and quality, which will be of value for DWTP management.

Keywords: Water quality monitoring; drinking water treatment plant; MALDI-TOF MS; 16S rRNA metabarcoding; bacterial diversity

Graphical abstract



Highlights:

- MALDI-TOF MS and metabarcoding analysis yielded different taxonomic results
- Proteobacteria predominated in the DWTP except in the final drinking water
- Chlorination strongly reduced bacterial diversity and shaped its taxonomy
- Temperature affected bacterial diversity in decantation and carbon filtration
- Culture and molecular methods are complementary in water quality management

1. Introduction

Among the Sustainable Development Goals of the United Nations, Goal 6 (Clean Water and Sanitation) calls for universal access to sufficient, safe, and affordable drinking water by 2030 (United Nations, 2020). By reducing levels of organic matter and microorganisms in water, conventional or advanced treatments can greatly improve its quality from source to tap and provide safe drinking water. Although each drinking water treatment plant (DWTP) has its own variations, the most common treatment processes employ a multibarrier approach consisting of coagulation, flocculation, decantation, filtration (e.g., sand or granular/biological activated carbon), with (or without) a final disinfection step (e.g., using UV, chlorine or, chloramine) (Gitis and Hankins, 2018; Betancourt and Rose, 2004).

The recently updated Drinking Water Directive of the European Union (EU 2020/2184) has incorporated the monitoring of viral indicators of fecal pollution (coliphages) (Anonymous, 2020). However, despite such advances, water quality regulation still relies on monitoring a few culturable microorganisms that represent only a minor fraction of the water microbiome.

Heterotrophic bacteria can improve the quality of non-treated water by reducing organic matter and pollutants (Li et al., 2017; Proctor and Hammes, 2015) (Benner et al., 2013; (Skjevrak et al., 2004; (Liu et al., 2016; Wingender and Flemming, 2011)). However, if they proliferate in water systems, they may cause water deterioration or even pose a health risk (Liu et al., 2017, 2016). Moreover, high numbers of heterotrophic bacteria can interfere with coliform detection. Therefore, water quality monitoring has traditionally focused not only on the detection of culturable bacterial indicators of fecal contamination but also on heterotrophic bacteria counts (Bartram et al., 2003). Therefore, obtaining information about overall microbial diversity could shed new light on the functioning of treatment processes.

The wide range of techniques available for the study of microbial communities in DWTPs fall within two broad strategies: culture-dependent (involving the controlled growth of specific microorganisms on selected media) and culture-independent (based on the analysis of nucleic acids previously extracted from samples).

Different culture-based methods for bacterial identification are available. For instance, biochemical phenotyping down to the species level can be performed using API galleries (bioMérieux, France), the Phene Plate System (PhP-Plate Microplate Techniques AB,

Sweden) or Biolog (Biolog, US). Widely used, this approach has achieved highly discriminatory bacterial characterization in water samples (Blanch et al., 2007; Hou et al., 2018; Sala-Comorera et al., 2016a, 2016b). However, phenotyping is laborious and time-consuming, especially in assessments of biodiversity, as each identification requires at least 24h incubation (and up to 120h) and its application is costly. Recently, the use of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) has been proposed as a rapid and robust technology for bacterial identification, involving the analysis of mass spectra of ribosomal proteins extracted from a whole bacterial cell previously isolated and grown. This system has been progressively implemented in routine monitoring, for example, of fermentation processes in the food industry (Angelakis et al., 2011; Kim et al., 2021). It has also been applied to facilitate taxonomic assignment of meiofauna (Rossel et al., 2019) or to identify heterotrophic bacteria present in water during treatment (Sala-Comorera et al., 2017) and in drinking water (Pinar-Méndez et al., 2021), producing results within a few minutes. However, as MALDI-TOF MS was originally designed for clinical diagnostics, mass spectra of target strains may not be available in the database, which limits successful identification.

On the other hand, the use of culture-independent methods, such as 16S rRNA metabarcoding, which allows the study of the total bacteria (culturable and non-culturable), has increased dramatically, thanks to the availability of high-throughput platforms that can simultaneously sequence millions of DNA fragments at a reasonable cost. Although this approach allows an in-depth analysis of the water microbiome, its implementation is still time-consuming. Moreover, the results can be biased by the nucleic extraction protocol or choice of primer, rare groups are not differentiated, and cell viability is not detected (Boers et al., 2019), all of which can complicate the interpretation of data for risk assessment.

Among one of the most challenging issues currently facing drinking water production and water quality management are the uncertainties related to climate change and its impact on freshwater resources. In this context, a consistent monitoring of bacterial communities in DWTPs combined with regulated water quality analysis based on the detection of microbial indicators may help to evaluate the current state of the art in water treatment and distribution systems and their resilience to water stress.

The aim of this work was to combine and compare two different approaches (culture-dependent and -independent) to characterize the bacterial communities in a full-scale DWTP. The culture-based strategy consisted of heterotrophic plate counts (HPC)

(European Directive (EU) 2020/2184) and identification by MALDI-TOF MS using a previously developed in-house database (Pinar-Méndez et al., 2021). The culture-independent analysis was based on 16S rRNA metabarcoding (V4 region) data obtained using an Illumina MiSeq platform in a previous study (Pinar-Méndez et al., 2022).

2. Material and methods

2.1. Study site and sampling

Samples were collected from a full-scale DWTP in Sant Joan Despí, 6 km south of Barcelona (Catalonia, North-East of Spain), in the lower basin of the Llobregat River. The anthropogenic impact in the area is high, with widespread industrial and agricultural activities, and the quality of river water is affected by wastewater effluents and industrial discharges. The DWTP has two sources of water: river and groundwater, as described previously (Pinar-Méndez et al., 2022). The water is subjected to sequential treatments as shown in **Figure 1**. The water from two treatment lines is mixed in a chamber and treated with chlorine (0.5-1.5 mg/L residual chlorine) in a separate tank before being pumped into the distribution systems.

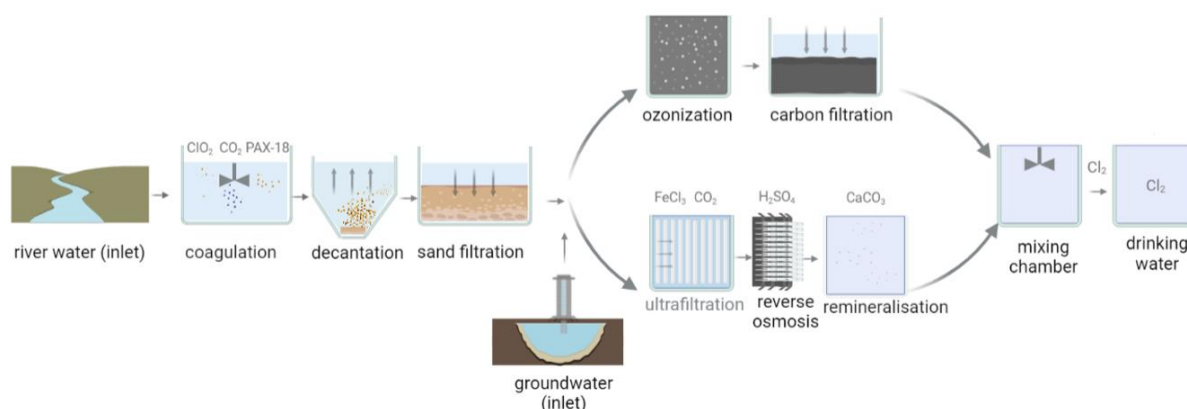


Figure 1. Schematic diagram of the stages in a drinking water treatment plant in Sant Joan Despí (Barcelona, Spain).

Seven sampling campaigns were carried out over one year, including the winter of 2018 and summer of 2019, when temperatures ranged from 11.5°C to 29.5°C. A total of 63 water samples were collected at nine stages of treatment in the DWTP: river water (RW, $n=7$) and groundwater (GW, $n=7$) intake, decantation (DEC, $n=7$), sand filtration (SF, $n=7$), ozonization (OZ, $n=7$), carbon filtration (CF, $n=7$), reverse osmosis (RO, $n=7$), mixed chamber (MIX, $n=7$) and post-chlorination drinking water (DW, $n=7$).

Sampling sites were classified into two categories according to the microbial load: high or low. For high microbial load samples, corresponding to river water and pretreatment stages (RW, DEC, SF), small volumes of water (2L) were collected in polyethylene sterile bottles containing sodium thiosulfate (24 mg/L). For samples of the groundwater and conventional/advanced treatments (GW, OZ, CF, RO, MIX, DW), which had a lower microbial load, higher volumes (from 100 to 1100 L/sample) were collected using Rexeed™ 25-A filters (Asahi Kasei Medical Co, Japan) according to the previously described dead-end hollow fiber ultrafiltration (DEUF) method (Gunnarsdottir et al., 2020; Hill et al., 2007; Rhodes et al., 2011). All samples were transported to the laboratory at 4°C for analysis.

2.2. Sample processing

All samples were processed within 24 hours of collection and analyzed for microbial water quality parameters according to the European Directive (EU) 2020/2184. Additionally, all samples were analyzed by MALDI-TOF MS and metabarcoding methods. Sample preparation differed according to the technique and microbial load.

2.3. Heterotrophic bacteria plate count

HPC was performed by mass inoculation according to the ISO Standard 6222:1999. Water samples concentrated by the DEUF method (Rexeed) were eluted as described by Hill et al. (2007) and Gunnarsdottir et al. (2020). The final eluate was used for mass inoculation. For water samples with a higher microbial load that did not require DEUF concentration, direct mass inoculation was performed. Different sample volumes (0.001-1 mL) were inoculated in duplicate in ISO Water Plate Count Agar (Oxoid, UK) and incubated at $22 \pm 2^\circ\text{C}$ for $72\text{h} \pm 3\text{h}$.

2.4. Analysis of bacterial communities by MALDI-TOF MS

The bacterial composition of water samples was analyzed using the Microflex LT MALDI-TOF MS device (Bruker Daltonics, Germany). In HPC dilution plates containing well-isolated colonies (from 10 to 130), a total of 30 colonies per sample (where possible) were randomly selected for identification purposes and diversity studies (Bianchi and Bianchi, 1982). To achieve fresh pure cultures, bacterial isolates were subcultured in Water Plate Count Agar and incubated at $22 \pm 2^\circ\text{C}$ for $72\text{h} \pm 3\text{h}$ or until growth (maximum 7 days). Experiments were carried out in a biosafety level 2 cabinet (BSL-2). Samples for MALDI-TOF MS analysis were prepared using the formic acid extended direct

transfer method recommended by Bruker Daltonics under a fume extraction cabinet, as previously described (Pinar-Méndez et al., 2021).

The resulting spectrum patterns were classified according to log score values and consistency category and identified by matches with the reference mass spectra database. This study was conducted using two databases simultaneously: the Bruker Daltonics (BDAL) library (8,468 reference spectra) and the Drinking Water Library (DWL) (319 reference spectra), the latter being previously developed in our laboratory specifically for the identification of bacteria associated with water for human consumption (Pinar-Méndez et al., 2021). Results were interpreted according to the Bruker score values: highly probable species identification (green; ≥ 2.300 – 3.000), secure genus and probable species identification (also green; 2.000 – 2.299), probable genus identification (yellow; 1.700 – 1.999) and unreliable identification (red; ≤ 1.699). Another approach to taxonomic assignment (Bruker Daltonics) classifies the top ten best matches according to categories of consistency: A (species consistency: top 10 matches > 2.000 are of the same species, or > 1.700 are of the same genus); B (genus consistency: top 10 matches > 1.700 are of the same genus but not of the same species); C (no consistency: top 10 matches < 1.700 , or > 1.700 are not of the same genus).

2.4.1. Unidentified isolates

A fraction of colonies that could not be identified were subcultured, incubated as described above and submitted to further analysis. If the second attempt at identification failed, isolates were reprocessed but using a different sample preparation method for an improved protein extraction. Thus, the acid/acetonitrile extraction method was performed following Bruker's instructions, and 1 μ L of protein suspension was spotted in triplicate on a MALDI 96 target plate, which was air-dried, and samples were covered with 1 μ L of matrix solution for the analysis. If the colonies remained unidentified, they were treated separately for clustering purposes (**Figure S1**). First, for the non-identified isolates, a mini Main Spectrum Profile (mMSP) *ad hoc* library was created using MALDI Biotyper 3.1 software offline (Bruker Daltonics, Germany). Thus, 9 mass spectra per isolate were checked for the following criteria: 3,000 to 10,000 Da and maximum error tolerance of 500 ppm. The mMSP library was used as a database for offline analysis of the unidentified isolates one-by-one and score values ranging from 3.000 to 1.700 were grouped into a cluster, assuming a maximum resolution at genus level.

BioNumerics 7.6 software (Apply Maths, Belgium) was used to process and analyze the large MALDI-TOF spectra datasets. For clustering purposes, similarities between all spectra were analyzed using the peak-based Pearson correlation coefficient and clustered by the UPGMA clustering algorithm to generate a dendrogram. The maximum distance level cut-off for clustering spectra was set at 50.

2.5. Comparison of culturable bacterial communities with 16S rRNA metabarcoding data for the total bacteria

The results obtained by MALDI-TOF MS were compared with previously obtained 16S rRNA (V4) metabarcoding data derived from the same samples (Pinar-Méndez et al., 2022) (**Figure S2**).

2.6. Statistical analyses

Data were analyzed using R version 4.0.5 and RStudio version 1.2.1335. All samples were \log_{10} transformed for statistical purposes. The Shapiro-Wilk Normality test was run to check if data fitted a normal distribution. The Kruskal–Wallis test was used to look for differences in HPC between all the DWTP samples, and then the Mann-Whitney U test was run to find out which samples were different. The alpha diversity metrics of Shannon (diversity index) and Observed (richness) were calculated at the genus level to assess bacterial diversity, including all isolates (identified and non-identified). To determine seasonal differences in HPC and diversity, the samples were grouped according to the ambient temperature during the sampling campaign (low $\leq 18^{\circ}\text{C}$ and high $\geq 22^{\circ}\text{C}$) and the Mann-Whitney test was used to determine possible significant differences between groups.

3. Results

3.1. Heterotrophic bacteria

The studied drinking water fully met all the quality standards of the EU Directive (data not shown). All samples gave positive HPC results except for two: one after the DEC stage and the other after OZ. The total number of culturable heterotrophic bacteria differed significantly between all the treatment stages (Kruskal-Wallis test, $p < 0.01$). Average values ranged from 0.027 \log_{10} CFU/mL to 4.62 \log_{10} CFU/mL (**Figure 2A**).

When comparing the two water sources, HPC values were 4 log₁₀ units higher in RW compared to GW (4.47 log₁₀ CFU/mL and 0.37 log₁₀ CFU/mL, respectively). After DEC, the HPC remained high (4.62 log₁₀ CFU/mL). A decrease of about 1-log was observed after the SF stage (3.63 log₁₀ CFU/mL), and a further decrease of 2 log₁₀ units after the subsequent processes of OZ (2.18 log₁₀ CFU/mL) and CF (2.16 log₁₀ CFU/mL). The HPC after the advanced treatment (1.48 log₁₀ CFU/mL) was 3 log₁₀ units lower compared to RW. MIX samples (2.23 log₁₀ CFU/mL) gave similar HPC values to those of CF, followed by a strong reduction of more than 2 log₁₀ units after the final chlorination (0.027 log₁₀ CFU/mL).

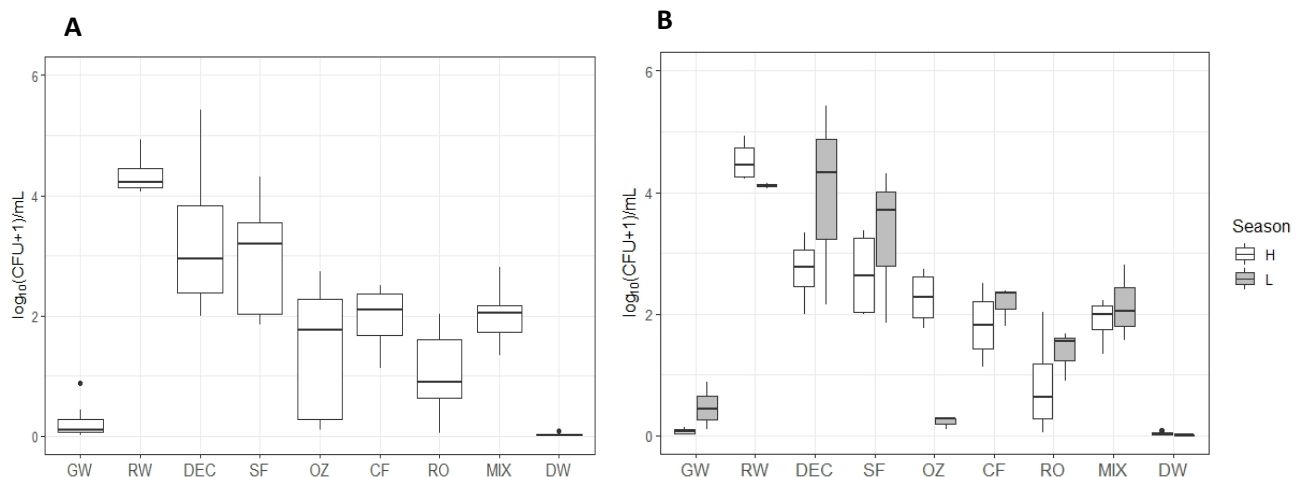


Figure 2. Boxplot charts displaying the distribution of heterotrophic bacteria plate counts (log₁₀ CFU/mL) in water samples from different stages in the DWTP: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Boxplots in (A) show total counts and in (B) the results are grouped by the ambient temperature when sampling was carried out: high (H) or low (L). Summary of data includes minimum score (lower whisker), first, median and third quartiles (box) and maximum scores (upper whisker). Statistically significant outliers are represented by dots.

Seasonal differences in HPC were found for some stages in the DWTP (**Figure 2 B**). Thus, RW, OZ and DW samples presented significantly higher HPC values at high

temperature (HT) than low temperature (LT) (Mann-Whitney U test, $p < 0.05$). Conversely, in GW, DEC, SF, CF and MIX samples, counts were slightly higher at LT, although the differences were not statistically significant. Interestingly, bacterial regrowth was observed after CF at LT, which accounted for a 2-log increase with regard to the previous stage (OZ).

3.2. Analysis of bacterial communities by MALDI-TOF MS analysis

3.2.1. Identification of bacterial isolates using BDAL and DWL databases

A total of 1,807 colonies were isolated from the DWTP stages and further analyzed by MALDI-TOF MS: GW (211), RW (210), DEC (182), SF (203), OZ (180), CF (207), RO (211), MIX (211) and DW (192). The DWL database (Pinar-Méndez et al., 2021) was used to improve identification, as with only the BDAL database, more than half of the isolates (1,010 out of 1,807) remained unidentified (56%) (**Figure 3**). The simultaneous use of both databases gave better results, reducing unreliable identification to 38% (686 out of 1,807), and classification at the genus level increasing to 32% and species level to 30%.

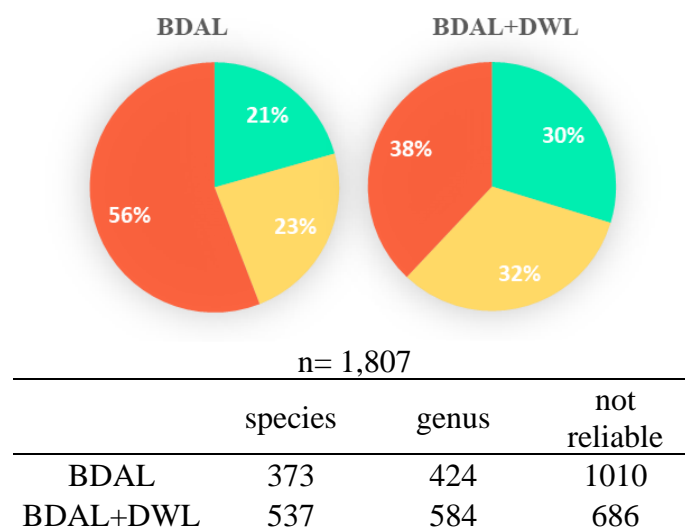


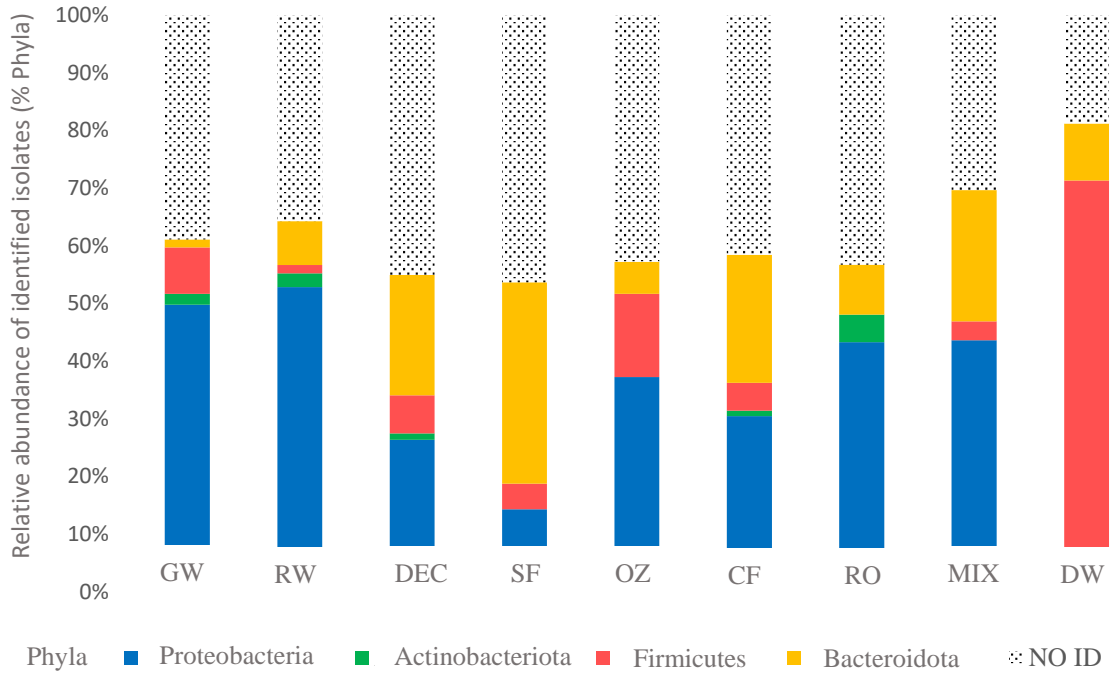
Figure 3. Identification of the 1,807 water bacterial isolates when using the Bruker Daltonics (BDAL) database alone or in conjunction with the Drinking Water Library (DWL). Isolates were classified according to scores based on MALDI-TOF MS analysis: highly probable species (≥ 2.300 – 3.000 ; green), secure genus and probable species (2.000 – 2.29 ; yellow), probable genus (1.700 – 1.999 ; yellow) or unreliable identification (≤ 1.699 ; red).

While using the BDAL database alone, unreliable identification at the different DWTP stages varied from 32 to 71%, compared to 30 to 46% when using the extended database (BDAL+DWL) (**Figure S3**), the rate being even lower in DW (19%). Accordingly, combining both databases led to an improvement in results in all DWTP samples and the classification rate at genus or species level ranged from 54 up to 81%.

3.2.2. Characterization of bacterial isolates

The bacterial community composition in the DWTP was highly diverse and differed at each stage, not only at the genus level but also at the phylum level. Four phyla were detected (**Figure 4A**). Overall, Proteobacteria predominated, except in DEC, SF and CF samples, in which the proportion of Bacteroidota was higher, and in DW, where Firmicutes was clearly dominant. Bacteroidota was well represented throughout the DWTP and was the second most dominant phylum. In contrast with its abundance in chlorinated DW, Firmicutes was scarcely found during the previous treatments, with the exception of OZ samples, in which 14.4% of the isolates were affiliated to this phylum. Actinobacteriota was also detected in the treatment samples, but to a low extent (4.7 % of the isolates).

A



B

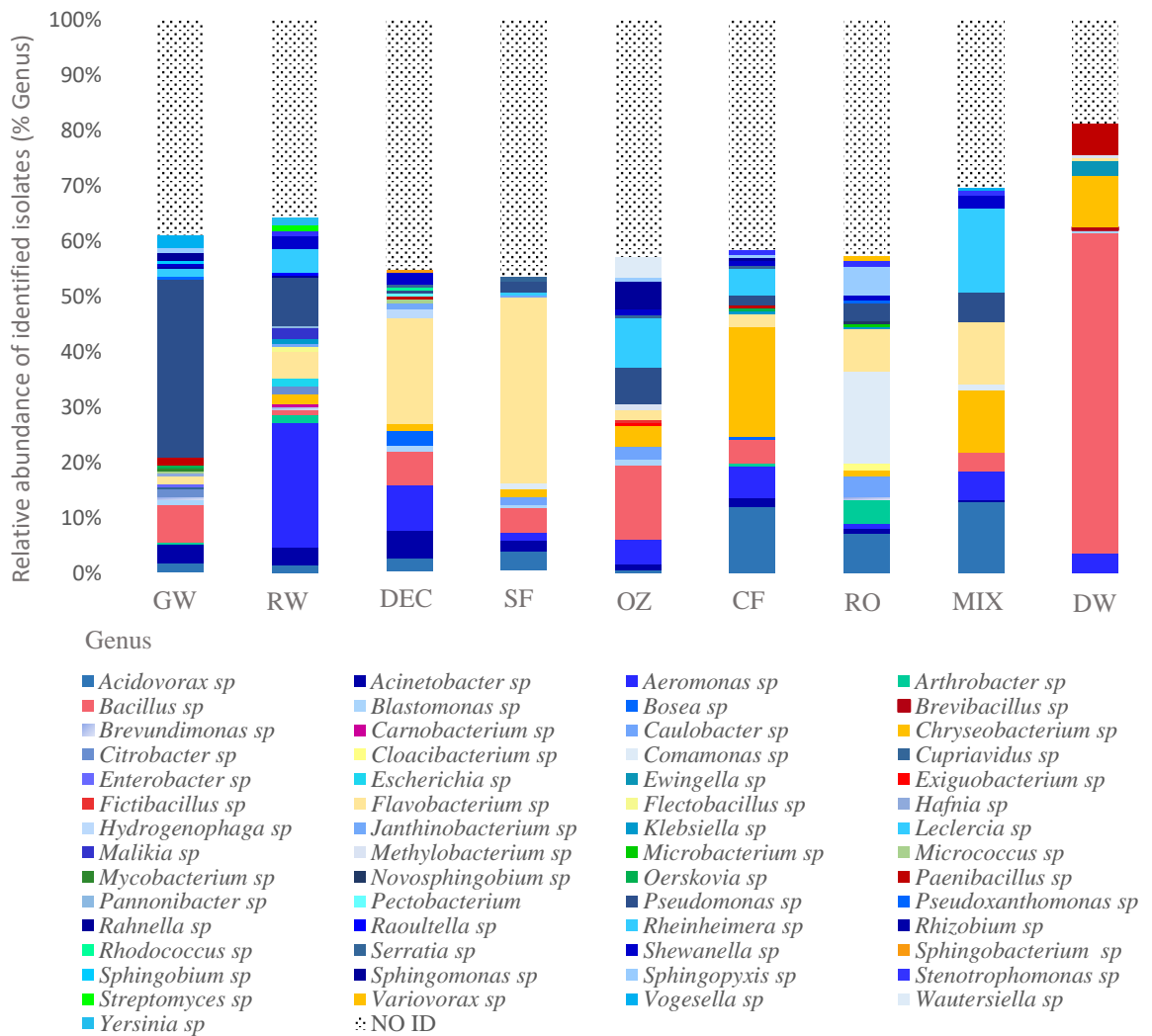


Figure 4. Patterns of relative abundance of isolates identified at the A) phylum level and B) genus level, and their distribution along the treatment stages: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Dots represent unidentified fractions (NO ID).

At the genus level, a total of 57 genera were successfully identified by MALDI-TOF MS (**Figure 4B**). Analysis of GW and RW revealed each contained 24 different genera, *Aeromonas* (22%) and *Pseudomonas* (32%) predominating, respectively. The first change in bacterial composition was observed in the pretreatment stages, as *Flavobacterium*, poorly represented in RW (5%), was the most abundant genus in DEC (19%) and SF (34%). Further shifts were observed during the conventional treatments, with a reduction of *Flavobacterium* and an increase of *Bacillus* (13%) and *Rheinheimera* (9%) after OZ, and a higher relative abundance of *Chryseobacterium* (20%) and *Acidovorax* (12%) after CF. Regarding the advanced treatment, the predominant genus in RO samples was *Comamonas* (17%), not detected upstream except after SF (1%). The genus profile in MIX samples was similar to that of CF (*Rheinheimera* 15%, *Acidovorax* 13%, *Flavobacterium* 11%, *Chryseobacterium* 11%), with little influence from the RO stage. Overall, the most notably shift in composition was observed in DW, in which *Bacillus* was clearly dominant (58%), and *Chryseobacterium* (9%) and *Paenibacillus* (6%) were also detected.

3.2.3. Unidentified isolates

The unidentified isolates (686 colonies) were grouped into a total of 91 clusters (C:565 colonies). 121 colonies had a single protein profile (S:80) or exhibited altered or stopped growth, probably due to stress (S:41) (**Figure 5**). One of the most abundant clusters, C-79 (green), was only detected at HT in the DEC (10%), SF (10%), CF (2%) and RO (3%) stages, and was not found in source water or the final DW. C-47 (light brown) was only detected at LT in OZ (5%) and DW (13%) samples. Other clusters had less seasonal association, such as C-31 (dark brown), which in CF samples was more abundant at LT (26%) than HT (7%), but was also detected in low relative abundance (2-7%) in DEC, SF, OZ, and MIX stages in both seasons.

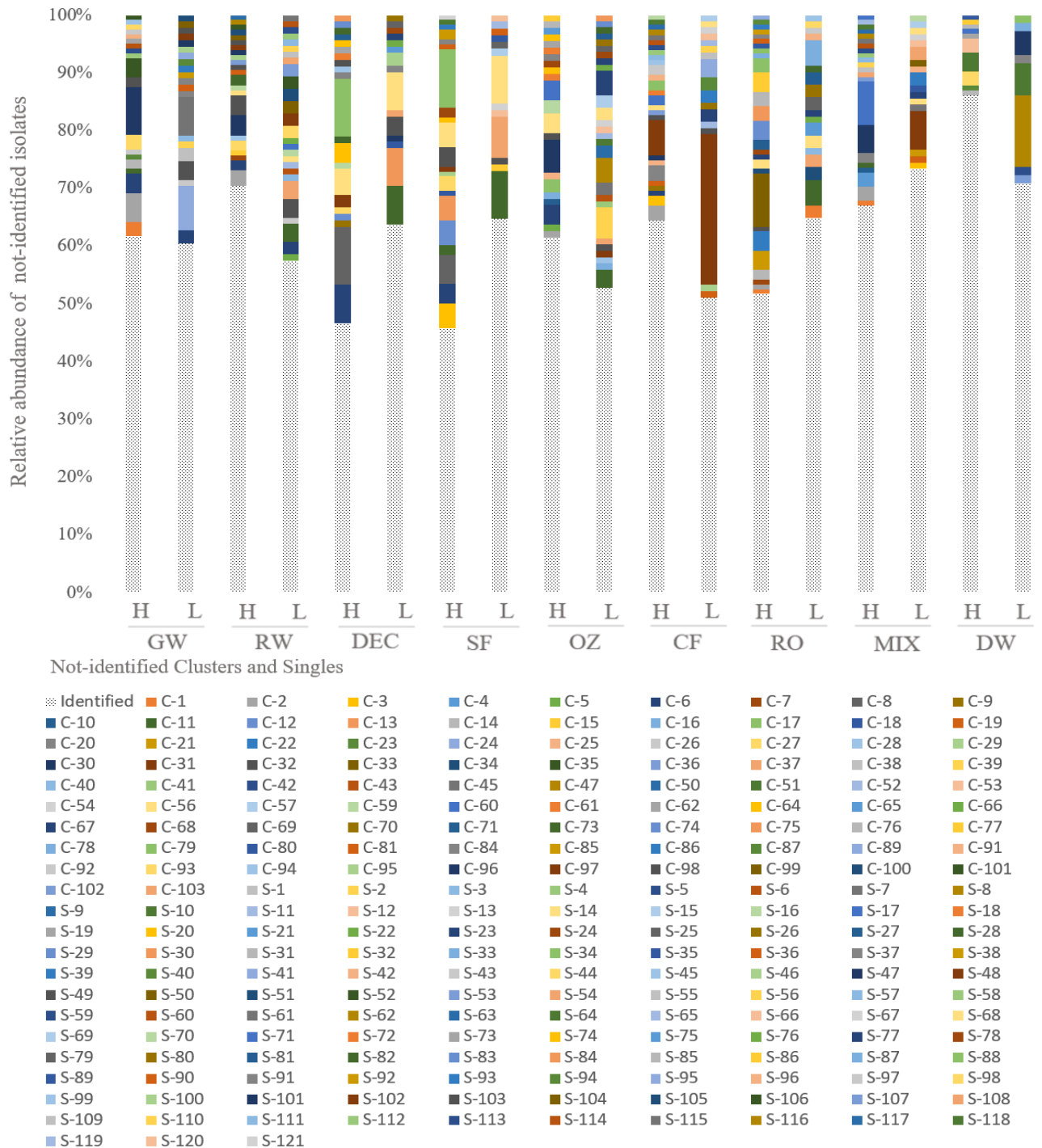


Figure 5. Distribution of unidentified isolates in the DWTP: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Samples are grouped according to the ambient temperature during the sampling campaign: high (H) or low (L). The colors show the relative abundance of all isolates,

which were either single (S) or grouped in clusters (C). Dots represent the identified fraction previously shown in **Figure 4**.

3.2.4. Alpha diversity analysis

Alpha diversity metrics were calculated at the genus level to assess bacterial diversity. For this analysis, the clusters described in the previous section were included, as they represent possible genera (**Figure S4**). Changes in diversity were observed throughout the treatment, a decrease occurring from source waters to the final DW. Statistically significant differences in bacterial communities were found between some treatment stages (Kruskal Wallis, $p < 0.01$). RW presented the highest diversity indices, followed by GW, while the lowest values corresponded to DW, where after the chlorination procedure diversity and richness was significantly lower compared to source water (Tukey HSD test, $p = 0.002$). Small differences in diversity were observed between treatment stages, but they were not statistically significant. Nevertheless, all treatment stages clearly differed from DW.

3.2.5. Seasonal variation

In order to detect possible seasonal variations in bacterial diversity in the DWTP stages, samples were grouped into two categories according to ambient temperature, high (HT; $\geq 22^\circ\text{C}$) and low (LT; $\leq 18^\circ\text{C}$), but only small differences were observed in some samples, without statistical significance (**Figure 6**). Species richness (number of observed species) and the Shannon index followed a similar trend. The Shannon index revealed a small increase in diversity at HT in the DEC, SF, OZ, CF and MIX stages, and at LT in RW, RO and DW, whereas GW samples remained quite stable between seasons. Richness

increased slightly at HT in DEC, SF, CF, and MIX samples and at LT in RW and RO samples, but no differences were observed for GW, OZ and DW.

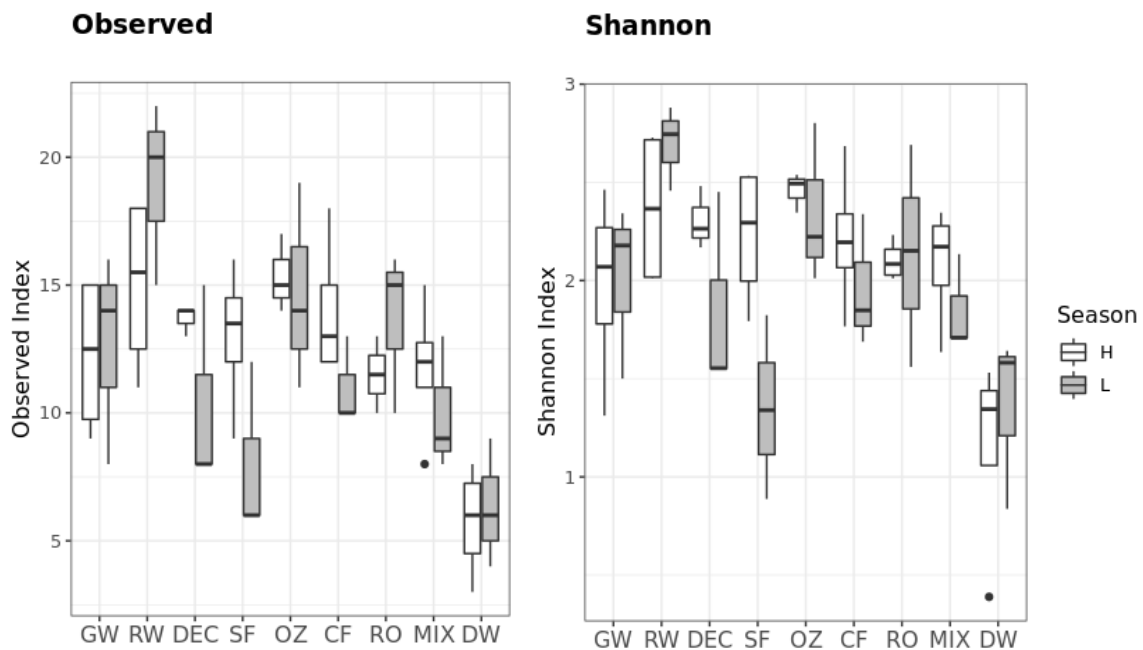
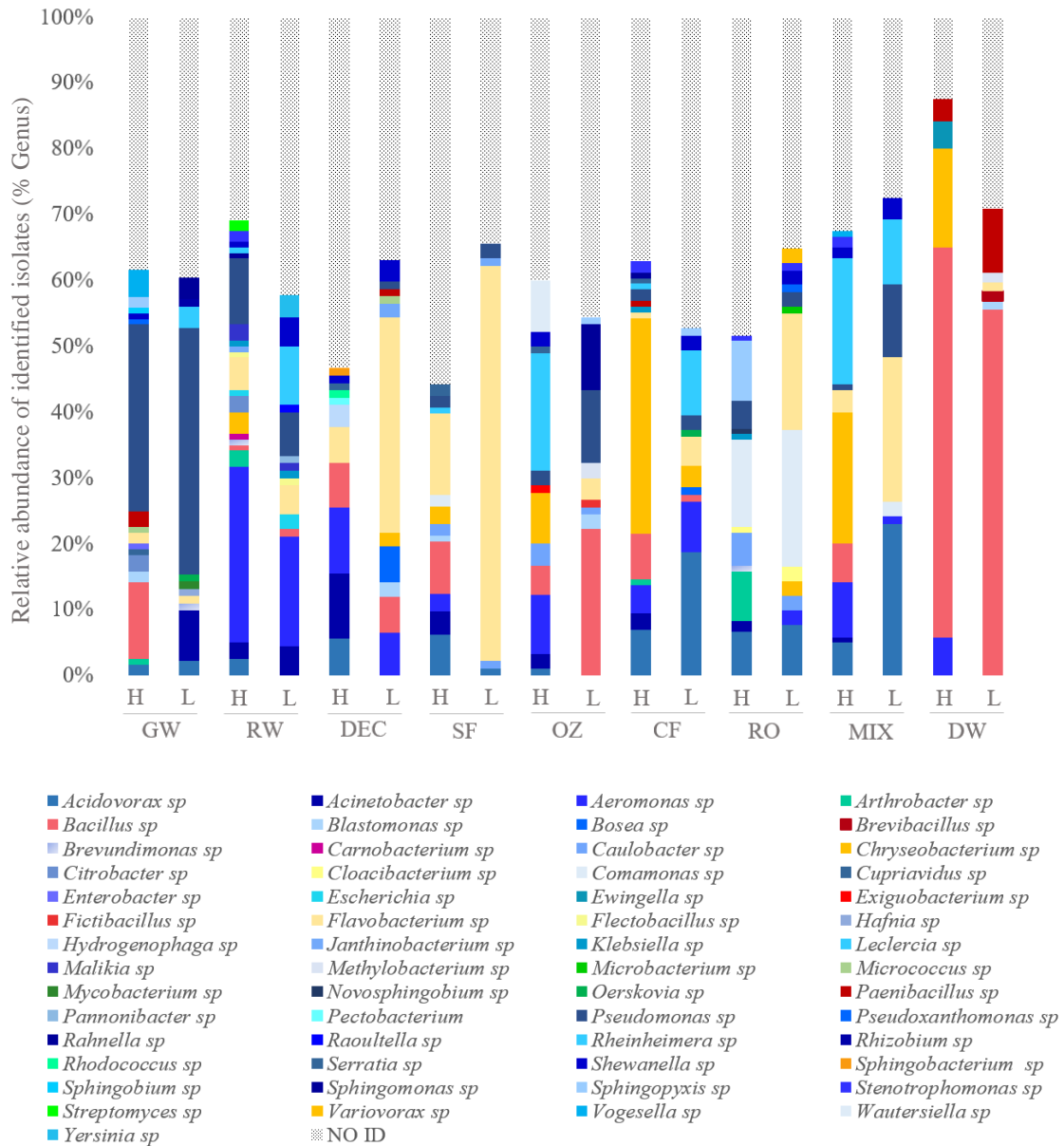


Figure 6. Boxplots representing alpha diversity indices of genus richness (Observed) and diversity (Shannon) for all DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Samples are grouped according to the ambient temperature during the sampling campaign: high (H, white) or low (L, grey).

Small differences in taxonomy at the genus level were also associated with seasons (**Figure 7**). At LT, DEC and SF samples were mostly dominated by *Flavobacterium* (33% and 60%, respectively), whereas at HT, when genus diversity was higher, the relative abundance of *Flavobacterium* was lower (5% and 12%, respectively). In conventional treatment stages, the temperature also affected the taxa. Thus, the dominant genus in OZ samples was *Bacillus* (22%) at HT and *Rheinheimera* (18%) at LT. In CF samples, the high relative abundance of *Chryseobacterium* at HT (32%) was reduced at LT (3%). In the advanced treatment, RO samples showed different profiles according to temperature: at LT, *Comamonas* (19%) and *Flavobacterium* (16%) were predominant, whereas the latter was not detected at HT. Moreover, two genera were only found at HT in RO

samples: *Spingopyxis* (10%) and *Arthrobacter* (9%), both of which were detected in a few stages upstream only in very low abundance (<1%). In MIX samples, *Acidovorax* (21%) and *Flavobacterium* (20%) were well represented at LT, whereas at HT the predominant genera were *Chryseobacterium* (20%) and *Rheinheimera* (20%). Finally,



DW did not show large differences in genus composition, being dominated by *Bacillus* in both seasons (HT: 70%, LT: 57%), although *Aeromonas*, *Chryseobacterium* and *Ewingella* were only detected at HT, and *Blastomonas*, *Brevibacillus*, *Flavobacterium* and *Methylobacterium* only at LT.

Figure 7. Seasonal variation in genus diversity in the 63 DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW). Samples are grouped according to the ambient temperature during the sampling campaign: high (H) or low (L). Dots represent the unidentified fraction (NO ID).

3.3. Comparison of culturable bacterial communities with the total bacteria detected by 16S rRNA metabarcoding

The study of culturable bacterial communities provides very useful information on metabolically active cells. However, culturable bacteria represent only a minor fraction of the entire water microbiome and other approaches are needed to provide complementary and more in-depth information. Thus, the identified heterotrophic bacterial populations were compared with the data obtained in a previous 16S rRNA metabarcoding study (Pinar-Méndez et al., 2022).

3.3.1. Detection of MALDI-TOF MS-identified genera in metabarcoding reads

The 57 genera identified by MALDI-TOF MS in samples from the different water treatments were compared with metabarcoding data (presence/absence and relative abundance reads) (**Table 1**). On 41 occasions (27 different genera), MALDI-TOF MS allowed the detection of a genus that was missed by metabarcoding (boxes edged in black). Conversely, on 152 occasions, 33 genera were only detected by metabarcoding (boxes edged in white).

It is also noteworthy that the most abundant genera in each matrix according to MALDI-TOF MS analysis were all successfully detected by metabarcoding, although with different relative abundance. Thus, in GW, *Pseudomonas* abundance was higher by MALDI-TOF MS (32%) than by metabarcoding (1%); a similar difference was observed in *Aeromonas* in RW (22% versus 0.4%, respectively). In DEC samples, relative abundance of *Flavobacterium* was similar by MALDI-TOF MS (19%) and metabarcoding (14%), and in SF samples, *Flavobacterium* was the prevailing genus by both techniques (33.5% and 15%, respectively). In conventional treatments, at the OZ stage *Bacillus* was predominant according to MALDI-TOF MS analysis (13%), whereas metabarcoding indicated a very low abundance (0.005 %). Similarly, 19.8% of the CF

isolates corresponded to *Chryseobacterium*, whose incidence was much lower according to metabarcoding results (0.08 %). In the advanced treatments, *Comamonas* was identified in RO samples by both MALDI-TOF MS (16.6%) and metabarcoding (2%). In MIX samples, *Rheinheimera* represented 15% of the culturable fraction and was also detected by metabarcoding (3%). Finally, in DW samples, culturable *Bacillus* predominated (57.8%) according to MALDI-TOF MS analysis and was also identified by metabarcoding but with much lower abundance (0.05%).

Regarding the total genera in the final DW, some genera were only detected by MALDI-TOF MS but not by metabarcoding, in abundances ranging from 0.5 to 3.9%: *Arthrobacter*, *Brevibacillus*, *Carnobacterium*, *Enterobacter*, *Leclercia*, *Micrococcus*, *Oerskovia*, *Rahnella*, *Rhodococcus*, *Sphingobacterium*, *Streptomyces* and *Wautersiella*. On the other hand, metabarcoding revealed a greater diversity of bacteria in DW not detected by MALDI-TOF MS, the relative abundance reads ranging from 0.003 to 2%: *Acidovorax*, *Bosea*, *Brevundimonas*, *Cloacibacterium*, *Comamonas*, *Hafnia*, *Hydrogenophaga*, *Janthinobacterium*, *Klebsiella*, *Mycobacterium*, *Novosphingobium*, *Pseudomonas*, *Pseudoxanthomonas*, *Raoultella*, *Rheinheimera*, *Rhizobium*, *Serratia*, *Shewanella*, *Sphingobium*, *Sphingomonas*, *Sphingopyxis*, *Stenotrophomonas*, *Variovorax* and *Yersinia*.

Table 1. Heatmap showing the relative abundance of the 57 bacterial genera identified by MALDI-TOF MS and their presence/absence and abundance according to metabarcoding analysis in all the DWTP stages. Data are illustrated by a color scale spanning from green (not detected) to red (high relative abundance). Boxes edged in black indicate identification only by MALDI-TOF MS, while boxes edged in white correspond to exclusive identification by metabarcoding. Stars indicates differences in taxon classification by metabarcoding: **Escherichia-Shigella*, ***Hafnia-Obesumbacterium*, ****Methylobacterium-Methylorubrum*, *****Microbacteriaceae*, ******Allorhizobium-Neorhizobium-Pararhizobium-Rhizobium*.

Genera	GW		RW		DEC		SF		OZ		CF		RO		MIX		DW	
	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S	MALDI	16S
<i>Acidovorax</i>	1.9	0.007	1.4	0.8	2.7	0.5	3.9	0.6	0.6	1	12.1	0.3	7.1	0.8	12.8	0.2	0	0.04
<i>Acinetobacter</i>	3.3	0.2	3.3	2.0	4.9	0.8	2.0	0.7	1.1	0.2	1.4	0.02	0.9	2.0	0.5	0.004	0	0
<i>Aeromonas</i>	0	0.02	22.4	0.4	8.2	0.2	1.5	0.3	4.4	2	5.8	3	0.9	0.8	5.2	4	3.6	2
<i>Arthrobacter</i>	0.5	0	1.4	0	0	0	0	0	0	0	0.5	0	4.3	0	0	0	0	0
<i>Bacillus</i>	6.6	0	1.0	0	6.0	0.010	4.4	0	13.3	0.005	4.3	0	0	3.3	0	57.8	0.05	0
<i>Blastomonas</i>	0.9	0	0	0	1.1	0.3	0.5	0.3	1.1	0.5	0.0	0.1	0	0.6	0	0.1	0.5	0
<i>Bosea</i>	0	0.006	0	0.03	2.7	0.2	0	0.1	0	0.2	0.5	0.06	0	0.06	0	0.08	0	0.03
<i>Brevibacillus</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0
<i>Brevundimonas</i>	0.5	0.008	0.5	0.2	0	0.03	0	0.2	0	0.2	0	0.2	0.5	4.0	0	0.3	0	0.009
<i>Carnobacterium</i>	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Caulobacter</i>	0	0.07	0.0	0.08	0	0.1	1.5	0.2	2.2	0.06	0	0.0	3.8	7.0	0	0.2	0	0
<i>Chryseobacterium</i>	0	0	1.9	0.01	1.1	0	1.5	0.01	3.9	0.4	19.8	0.08	0.9	0.06	11.4	0.1	9.4	0.9
<i>Citrobacter</i>	1.4	0.008	1.4	0	0	0	0	0.02	0	0	0	0.01	0	0.02	0	0.01	0	0
<i>Cloacibacterium</i>	0	0.008	0	0.2	0	0.08	0	0.2	0	0.06	0	0.02	1.4	0.1	0	0.01	0	0.003
<i>Comamonas</i>	0	0.4	0	0	0	0	1.0	0.02	0	0.2	0	0	16.6	2.0	0.9	0.1	0	0.8
<i>Cupriavidus</i>	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0	0	0
<i>Enterobacter</i>	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Escherichia</i> *	0	0	1.4	0.003	0	0.007	0	0.009	0	0	0	0	0	0	0	0	0	0
<i>Ewingella</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.6	0
<i>Exiguobacterium</i>	0	0	0	0	0	0	0	0	0.6	0.001	0	0	0	0	0	0	0	0
<i>Fictibacillus</i>	0	0	0	0	0	0	0	0	0.6	0.0	0	0	0	0	0	0	0	0
<i>Flavobacterium</i>	1.4	0.05	4.8	11.0	19.2	14.0	33.5	15.0	1.7	9.0	2.4	7.0	7.6	8.0	11.4	7.0	0.5	2.0
<i>Flectobacillus</i>	0	0.002	1	0.008	0	0.02	0	0.008	0	0	0	0.002	0	0.03	0	0.003	0	0
<i>Hafnia</i> **	0.5	0	0	0	0	0	0	0	0	0	0	0.001	0	0	0	0	0	0.1
<i>Hydrogenophaga</i>	0	0.8	0	2.0	1.6	5.0	0	2.0	0	1.0	0	3.0	0	0.6	0	2.0	0	0.2
<i>Janthinobacterium</i>	0	0	0.5	0.009	1.1	0.02	0.5	0.1	0	0.05	0	0.06	0	0.8	0	0.07	0	0.01
<i>Klebsiella</i>	0	0.003	1.0	0.02	0	0.01	0	0.02	0	0.8	0.5	0.004	0.5	0.002	0	0	0	0.2
<i>Leclercia</i>	0	0	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0
<i>Malikia</i>	0	0	1.9	0.9	0	1	0	1	0	0	0	0.2	0	0.3	0	0.2	0	0
<i>Methylobacterium</i> ***	0	0	0	0	0	0	0	0	1.1	0.03	0	0	0	0	0	0	0.5	0
<i>Microbacterium</i> ****	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.09	0	0	0	0
<i>Micrococcus</i>	0.5	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Mycobacterium</i>	0.5	0.002	0	0	0	0.02	0	0.02	0	0.3	0	0.02	0	0.009	0	0.003	0	0.2
<i>Novosphingobium</i>	0	0.03	0	0.5	0	1.0	0	0.3	0	0.2	0	0.2	0.5	0.3	0	0.1	0	0.8
<i>Oerskovia</i>	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Paenibacillus</i>	1.4	0	0	0	0.5	0	0	0	0	0	0.5	0	0	0	0	0	5.7	0.004
<i>Pannonibacter</i>	0	0	0.5	0.003	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Pectobacterium</i>	0	0	0	0.002	0.5	0	0	0.002	0	0	0	0	0	0	0	0	0	0
<i>Pseudomonas</i>	32.2	1.0	8.6	3.0	0.5	2.0	2.0	3.0	6.7	8.0	1.9	2.0	3.3	5.0	5.2	3.0	0	1.0
<i>Pseudoxanthomonas</i>	0.5	0	0	0.1	0	0.007	0	0	0	0.003	0	0	0.5	0.02	0	0	0	0.004
<i>Rahnella</i>	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Raoultella</i>	0	0	0.5	0	0	0	0	0.005	0	0	0	0	0.008	0	0.02	0	0	0.01
<i>Rheinheimera</i>	1.4	0.3	4.3	0.6	0	0.8	0	1	8.9	6.0	4.8	4.0	0	0.2	15.2	3.0	0	0.6
<i>Rhizobium</i> *****	0.5	0	0	0.06	0	0.1	0	0.08	0	0.05	0	0.1	0	0.2	0	0.2	0	0.005
<i>Rhodococcus</i>	0	0	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Serratia</i>	0	0	0	0	0.5	0	1	0	0.6	0	0.5	0	0	0	0	0	0	0.07
<i>Shewanella</i>	0.5	0.01	2.4	0.06	2.2	0.02	0	0.03	1.1	0.9	1.0	0.3	0.9	1.0	2.4	0.2	0	0.003
<i>Sphingobacterium</i>	0	0	0	0	0.5	0	0	0	0	0.001	0	0	0	0	0	0	0	0
<i>Sphingobium</i>	0.5	0.006	0	0.002	0	0	0	0.003	0	0.2	0	0	0	0.05	0	0	0	0.06
<i>Sphingomonas</i>	1.4	0.05	0	0.09	0	0.09	0	0.4	5.0	0.5	0.5	3.0	0	0	0	1.0	0	0.3
<i>Sphingopyxis</i>	0.9	0.02	0	0.01	0	0.6	0	0.006	0.6	0.2	0.5	1.0	5.2	1.0	0	0.5	0	0.04
<i>Stenotrophomonas</i>	0	0	1.0	0.02	0	0.002	0	0.001	0	0.01	1.0	0.01	0.9	0.005	0.9	0.006	0	0.006
<i>Streptomyces</i>	0	0	1.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Variovorax</i>	0	0	0	0.006	0	0.07	0	0.03	0	0	0	0.6	0.9	0.4	0	0.3	0	0.02
<i>Vogesella</i>	2.4	0.04	0	0.02	0	0.003	0	0.004	0	0.06	0	0.05	0	0.5	0.5	0.09	0	0
<i>Wautersiella</i>	0	0	0	0	0	0	0	0	3.9	0	0	0	0	0	0	0	0	0
<i>Yersinia</i>	0	0	1.4	0.007	0	0.005	0	0.02	0	0.02	0	0.02	0	0.006	0	0.01	0	0.2

3.3.2. Fluctuation of the six most abundant genera along DWTP stages

According to MALDI-TOF MS analysis, the most abundant genera recovered from all DWTP samples were *Bacillus* (206 out of 1,807 isolates), *Flavobacterium* (166), *Aeromonas* (124), *Pseudomonas* (123), *Chryseobacterium* (106) and *Acidovorax* (87).

With the aim of tracking their fluctuation along the different treatment stages, and to evaluate the capacity of MALDI-TOF and metabarcoding techniques to detect their presence, these genera and their abundance were compared in a schematic diagram of the DWTP (**Figure 8**).

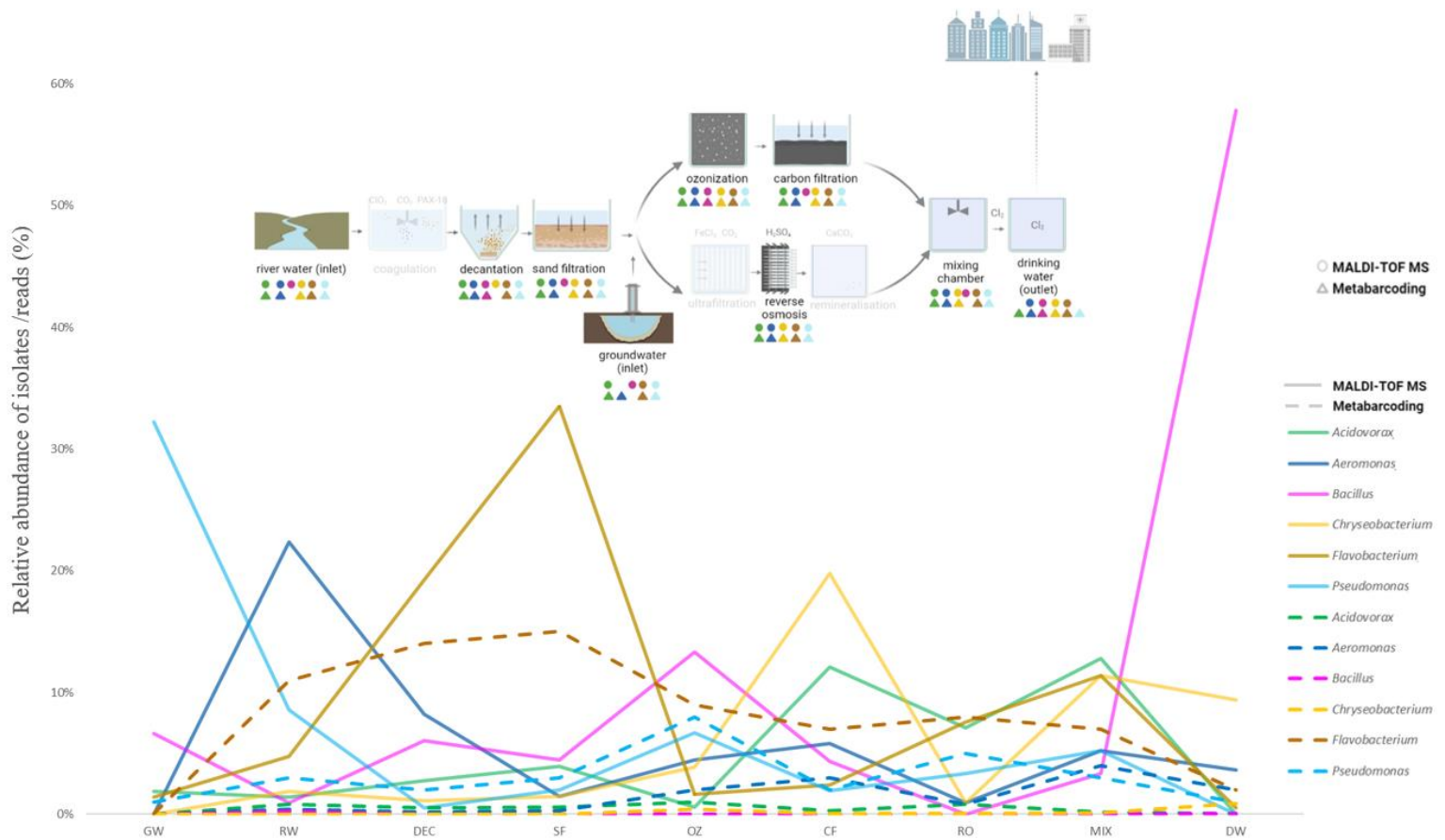


Figure 8. Fluctuation graph of the relative abundance of the six most abundant genera in the DWTP stages according to MALDI-TOF MS results (solid line) and relative abundance reads by metabarcoding (dotted line). In the schematic diagram of the DWTP the genera are indicated by a circle for MALDI-TOF results or a triangle for metabarcoding. Samples corresponded to groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination drinking water (DW).

Flavobacterium was the only genus identified by both approaches at all treatment stages, from source to tap. The genera *Acidovorax* and *Pseudomonas* were detected by both methods at all DWTP stages but in DW metabarcoding revealed a low abundance (0.04% and 1%, respectively). *Aeromonas* was identified by both techniques at all the stages except GW, where it was detected only by metabarcoding. *Bacillus* was tracked from inlet water to outlet DW, with a massive increase after chlorination, although its detection and abundance differed according to the technique. Analysis by MALDI-TOF MS provided positive results for *Bacillus* in GW, RW, SF, CF, MIX and DW, but by metabarcoding only in DEC, OZ and DW samples. Finally, *Chryseobacterium* was detected in RW (but not in GW), SF, OZ, CF, RO, MIX and DW by both approaches, and in DEC only by MALDI-TOF MS.

Likewise, to gain insight into the DWTP locations of the most abundant taxa, the six most abundant genera (>5%) according to MALDI-TOF MS (same genera as in Fig. 8) and metabarcoding (Pinar-Méndez et al., 2022) were compared. Their hotspots are depicted in **Figure S5**.

Flavobacterium and *Pseudomonas* were the most abundant genera in the DWTP according to both techniques. *Flavobacterium*, with a relative abundance >5%, was located mainly in RW, DEC, SF, OZ, CF, RO and MIX, although in OZ and CF samples it was only detected by metabarcoding. *Pseudomonas* (>5%) was identified in RW, GW, OZ, MIX only by MALDI-TOF MS, in OZ by both techniques, and additionally in RO by metabarcoding.

Regarding the top six genera according to the MALDI-TOF MS approach (all with a relative abundance >5%), *Aeromonas* and *Bacillus* were identified in source water, treatment stages and the final DW, whereas *Acidovorax* and *Chryseobacterium* were mainly detected in the late stages and DW (Fig S5).

Regarding the six most prevalent genera according to metabarcoding (with relative abundance reads >5%), *Limnohabitans*, unknown Comamonadaceae and unknown Chloroplast were detected in source water and treatment stages, while unknown Obscuribacteraceae was mainly found in DW. Therefore, hotspots for these genera consisted of early treatment processes for *Limnohabitans* (RW, DEC and SF); RW, OZ and DW for unknown Chloroplast; a range of treatments for unknown Comamonadaceae (RW, DEC, DS, OZ and RO), and finally, unknown Obscuribacteraceae presented a high abundance only in post-chlorination DW.

4. Discussion

The bacterial communities in a full-scale DWTP in Barcelona (Catalonia, North-East Spain) were compared via two identification methods, MALDI-TOF MS (culture-dependent) and high-throughput amplicon sequencing of the 16S rRNA gene (culture-independent), using metabarcoding data obtained in a previous study (Pinar-Méndez et al., 2022).

The studied water sources (river and groundwater) and the final drinking water fully met the quality standard requirements of the European Drinking Water Directive. HPC values decreased progressively along the successive DWTP treatments from source to tap (from 4.62 log₁₀ to 0.027 log₁₀ CFU/mL), with a significant reduction after chlorination. During the warmer seasons, the abundance of heterotrophic bacteria in RW, OZ and DW samples increased significantly. However, at the OZ stage, it should be noted that besides temperature, the level of residual ozone may be another influential seasonal factor, as the DWTP operational procedure stipulates a lower ozone dosage in summer (mean: 0.05 ppm) than in winter (mean: 0.22 ppm). This is to avoid the formation of bromate, an ozonation by-product and potential human carcinogen produced at higher temperatures (Von Gunten, 2003). Therefore, lower residual disinfectant may contribute to bacterial regrowth in the warmer seasons (Li et al., 2018). Additionally, although CF samples did not show seasonal variation, at low temperatures they had higher HPC values than samples from the preceding OZ stage. This may indicate the establishment of a microbial biofilm community in the carbon filters, as the HPC values of the CF effluent were 2 log higher than in the CF influent. Although drinking water regulations do not stipulate upper limits for the HPC, they require that there should be no abnormal changes (Anonymous, 2020). In this context, it is of interest that the two identification methods gave different results in terms of bacterial abundance and diversity, even in water samples with a normal HPC, as observed by other researchers using multiparametric approaches (Lautenschlager et al., 2013; Prest et al., 2014).

MALDI-TOF MS is described as a rapid identification tool and a reliable alternative to amplicon sequencing (Kraková et al., 2017), although its resolution capacity depends directly on the available reference spectra. In the present work, to analyze the diversity of culturable bacteria, 1,807 isolates from HPC plates were analyzed by MALDI-TOF MS using the Bruker BDAL database extended by our previously developed DWL (Pinar-Méndez et al., 2021). The rate of identification at the different treatment stages varied

from 54% to 81%. Such differences can occur when predominant taxa are not represented in the libraries, which increases the percentage of misidentification. Thus, to improve identification power, it is advisable to customize databases with taxa of interest, as previously reported (De Carolis et al., 2014; Kim et al., 2016; Seuylemezian et al., 2018). Accordingly, DW samples contributed the lowest percentage of unidentified isolates (19%), probably because in the DWL, 92 out of 319 main spectrum profiles (MSPs) corresponded to DW samples, whereas process water was represented by only 28 MSPs. Notably, in SF samples the rate of identification increased by 24% when both databases were used, although they still accounted for 46% of unidentified isolates. According to the alpha diversity indices, SF samples had a low diversity and therefore the unidentified fraction may correspond to numerous isolates of only a few taxa not represented in the database.

The bacterial diversity was influenced by the treatment applied, especially chlorination, which caused a strong reduction in the final drinking water. Nevertheless, the results need careful interpretation, as the identified and unidentified isolates (clusters or singles) were evaluated using alpha diversity indices with a cut-off at the genus level.

Trends in diversity differed according to the identification method used. Metabarcoding indicated that diversity was highest in GW followed by RW, with a gradual reduction during the successive treatments until CF; after a temporary increase, it was drastically reduced in the DW. According to MALDI-TOF MS results, diversity was highest in RW followed by GW, and after a similar decrease in subsequent stages, underwent an increase in OZ samples; only small differences in diversity were observed between the other stages, except the post-chlorination reduction.

Regarding taxonomic assignments, MALDI-TOF MS was able to identify 30% of isolates at the species level, while in some cases metabarcoding could not provide classification beyond the genus level due to the small size of the sequenced DNA fragment. A total of 36% of reads were unsuccessfully assigned to genera, and 6% of reads corresponded to *Candidatus* (sequence-based potential new taxa as-yet uncultured).

At the phylum level, MALDI-TOF MS identified Proteobacteria as predominant in source water (51% on average), its relative abundance decreasing during the pretreatment stages, when Bacteroidota increased, and prevailing again in conventional and advanced treatments. Both phyla predominated throughout the treatment until chlorination, when Firmicutes clearly increased its relative abundance (64.1%), becoming the dominant phylum in the final DW. The proportion of Firmicutes was less than 8% in source water

and treatment samples, except at the OZ stage, when it reached 14%. Similar trends have been described by other researchers (Atnafu et al., 2021; Li et al., 2017; Sala-Comorera et al., 2017). The domination of Firmicutes in drinking water may be attributed to the greater resistance of Gram-positive bacteria to chlorine disinfection in water systems compared to Gram-negative bacteria, which are more susceptible to disinfectants (Mir et al., 1997).

Succession was also observed at the genus level: *Aeromonas* and *Pseudomonas* predominated in source water, *Flavobacterium* in pretreatment stages, *Bacillus* in OZ, *Chryseobacterium* in CF, *Comamonas* in RO, *Rheinheimera* in MIX and *Bacillus* in DW. All these genera have been previously reported in drinking water environments (Atnafu et al., 2021; Fish and Boxall, 2018; Sala-Comorera et al., 2020, 2017).

In the studied system, spore-forming bacteria belonging to the *Bacillus* genus resisted all treatments and became dominant in the final drinking water, increasing after the disinfection procedures of ozonization and chlorination. Different species of *Bacillus* were identified in the DW: *B. cereus*, *B. cibi*, *B. horneckiae*, *B. indicus*, *B. idriensis*, *B. licheniformis*, *B. megaterium*, *B. mojavensis*, *B. mycoides*, *B. simplex*, *B. pumilus*, *B. sonoriensis*, *B. thuringiensis* and *B. weihenstemphanensis*. However, reflecting the difficulties in differentiating between *Bacillus* species, and perhaps influenced by the degree of sporulation, the MALDI-TOF identification scores ranged widely from 1.707 to 2.459 (Shu and Yang, 2017). Moreover, as some species within this genus are closely related, their identification with the 16S rRNA sequencing reference method was also challenging, and most isolates were not classified beyond the genus level. *B. cereus* is included in the Risk Group Database (ABSA, 2020) as a human and animal pathogen, but in this study it was not possible to safely distinguish it from other similar species such as *B. thuringiensis*, *B. mycoides* and *B. weihenstemphanensis*, none of which are human pathogens. On the other hand, some surfactin-producing *Bacillus* species have been reported as control agents against *Legionella pneumophila* due to the antagonistic activity of the biosurfactant (Loiseau et al., 2015) and some species are used in bioremediation of aquaculture water (Hlordzi et al., 2020).

MALDI-TOF MS analysis was able to detect taxonomic changes related to seasonality in some of the treatment stages. Temperature has been described as a modulator of microbial dynamics, which may disrupt the community composition in water (Degerman et al., 2013). For instance, both DEC and SF samples presented a clear predominance of *Flavobacterium* at low temperatures, whereas at higher temperatures, there was a greater

diversity of genera. Conversely, *Chryseobacterium* was more prevalent in CF at high than at low temperatures.

At the phylum level, the two techniques provided similar results, showing DWTP domination by *Proteobacteria* except in the final water, in which chlorination exerted a strong selective pressure, resulting in lower diversity and a taxonomic shift to Cyanobacteria (metabarcoding) or Firmicutes (MALDI-TOF MS). At lower taxonomic levels, the results differed to a greater extent, although the most abundant genera identified by MALDI-TOF MS were all detected by metabarcoding. In DEC, SF, RO and MIX samples, the most abundant genera were the same according to both methods: *Flavobacterium* in the pretreatment stages and MIX, and *Comamonas* (or uncultured Comamonadaceae by metabarcoding) in RO. The taxa identified in the other treatments varied (Pinar-Méndez et al., 2022). For instance, in GW the dominant genus was *Candidatus Omnitrophus* (25%) according to metabarcoding and *Aeromonas* (22%) by MALDI-TOF MS, while the final DW was dominated by uncultured members of Obscuribacteraceae (31%) or by *Bacillus* (58%), respectively.

Additionally, some genera were identified only by one of the two techniques. For instance, metabarcoding detected a greater variety of Proteobacteria at the genus level in DW. Despite the use of disinfectant residuals, certain chlorine-resistant taxa remained in the final treated water. According to other studies, the class Gammaproteobacteria is more resistant to free chlorine than Alphaproteobacteria (Mathieu et al., 2009). Moreover, some of the genera reported in the present study are associated with biofilm formation, such as *Bosea*, *Comamonas*, *Methylobacterium*, *Pseudomonas*, *Sphingobium* and *Sphingomonas* (Fish and Boxall, 2018). It is important to detect which taxa are seeding from biofilms into the distribution network to prevent possible regrowth on water pipes, which is also favored by high chlorine concentrations or rechlorination (Fish and Boxall, 2018; Mathieu et al., 2009). Water quality can be subsequently affected by bacterial mobilization into the water column. However, only a few isolates of these taxa were detected by MALDI, perhaps hindered by the dominance of *Bacillus* at this stage.

The most abundant genera detected in DW, *Bacillus* (Firmicutes) according to MALDI-TOF and unknown Obscuribacteraceae (Cyanobacteria) according to metabarcoding, have high resistance to disinfection procedures. However, the Cyanobacteria group includes a large proportion of uncultured strains. In the NCBI database, about 49% of cyanobacterial genomes correspond to uncultured metagenome-assembled strains (Dextro et al., 2021) and may therefore be undetectable by culture-based techniques such

as MALDI-TOF MS. Conversely, *Bacillus* was identified by metabarcoding, although in very low relative abundance.

Both technologies have benefits and shortcomings. Metabarcoding analysis provided a more detailed view of the bacterial communities in the DWTP water samples than MALDI-TOF MS, revealing a higher diversity and additional taxa. However, amplicon sequencing is currently unable to differentiate between live or dead cells, so the resulting data cannot provide a reliable assessment of hazards associated with specific taxa. Moreover, the results may be strongly influenced by nucleic acid extraction procedures, be biased against rare taxa, and samples low in DNA, such as chlorinated water, may be contaminated by the DNA reagent (Salter et al., 2014). To overcome the analytical problem of low microbial biomass in the DWTP, ultrafiltration by Rexeed was used, which concentrates large volumes of water and thus allows a more representative characterization of bacterial communities. This method allowed sufficient microbial biomass to be recovered for both MALDI-TOF and metabarcoding analyses.

As a cultured-based method, MALDI-TOF MS can focus on the viable microorganisms of a microbial community, although the characterization results can be skewed by other factors. Thus, sample processing (direct plating, filtration/concentration procedures), media nutrient composition, growth conditions (aerobic or anaerobic, incubation time and temperature), the protein extraction method as well as the representativeness of the mass spectra database are all factors that can impede identification. However, with a suitable database, proteomics studies by MALDI-TOF MS represent a rapid and reliable microbial identification strategy that provides data directly from isolated bacteria in a few minutes. Easy to use and low cost, it can be applied in routine bacterial monitoring in DWTPs and distribution systems, and the data interpretation does not require advanced bioinformatics skills. Nevertheless, culturable bacteria represent a minor fraction of the overall microbiome, so the evaluation of bacteria only by culture-based methods may underestimate the microbial diversity in a water system.

Both MALDI-TOF MS and metabarcoding detected shifts in the bacterial community composition in the DWTP that could be explained by bacterial attachment to or detachment from biofilms, the displacement or removal of certain taxa by the treatments, or selective microbial growth. Although these changes may influence water quality, they were not reflected in the quantitative data of HPC analysis required by the water safety regulations. Therefore, the information generated by the combined application of these proteomic and metagenomic tools in addition to the monitoring of regulated microbial

indicators should increase our understanding of how DWTPs function and improve water quality management. However, a more complete understanding of the complexity of microbial dynamics in a DWTP would require more knowledge about the metabolic activity and functions of the resident microbiota.

5. Conclusions

The HPC in analyzed samples decreased along the successive water treatments, increasing with ambient temperature in some processes. Both MALDI-TOF MS and 16S rRNA metabarcoding revealed a high bacterial diversity in the DWTP, with the highest values obtained in source water and the lowest in the final drinking water. Bacterial monitoring by the two techniques provided insights into the taxa present in the different DWTP stages, with low-abundance species being detected and some genera identified by only one approach. Spore-forming bacteria of the *Bacillus* genus (Firmicutes) resisted all treatments and was the dominant genus in drinking water according to MALDI-TOF MS, whereas metabarcoding indicated that members of the unculturable Obscuribacteraceae (Cyanobacteria) were more abundant. Monitoring microbial fluctuations between treatment stages and seasons is of interest for drinking water management and can help assess the impact of specific events on treatment effectivity and water quality. Overall, the combined use of multiple approaches constitutes a powerful monitoring tool of bacterial community dynamics in DWTPs.

Supplementary material

Supplementary material associated with this article: Additional file 1_ Figures S1-S5.

CrediT authorship contribution statement

A.P.-M.: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Visualization, Writing – original draft, Funding acquisition. **B.G.:** Conceptualization, Resources, Project administration, Supervision, Writing – review & editing. **A.R.B.:** Conceptualization, Methodology, Resources, Project administration, Supervision, Writing – review & editing. **C.G.A.:** Conceptualization, Methodology, Resources, Formal analysis, Supervision, Project administration, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by Pla de Doctorats Industrials of the Catalan Government (2016 DI 083) and the Water Research Institute. Graphical abstract was created with BioRender.com. We thank Jordi Martin, Meritxell Minoves and Sergio Montes for their technical support expertise in DWTP procedures.

References

- ABSA, 2020. Risk Group Database. Am. Biol. Saf. Assoc. Risk Gr. Database. <https://my.absa.org/Riskgroups>.
- Angelakis, E., Million, M., Henry, M., Raoult, D., 2011. Rapid and accurate bacterial identification in probiotics and yoghurts by MALDI-TOF mass spectrometry. *J. Food Sci.* 76. <https://doi.org/10.1111/j.1750-3841.2011.02369.x>
- Anonymous, 2020. Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption, Official Journal of the European Union.
- Atnafu, B., Desta, A., Assefa, F., 2021. Microbial community structure and diversity in drinking water supply, distribution systems as well as household point of use sites in Addis Ababa City, Ethiopia. *Microb. Ecol.* <https://doi.org/10.1007/S00248-021-01819-3>
- Bartram, J., Cotruvo, J., Exner, M., Fricker, C., Glasmacher, A., 2003. Heterotrophic Plate Counts and Drinking-water Safety The Significance of HPCs for Water Quality and Human Health. *IWA Publ.* <https://doi.org/10.2166/9781780405940>
- Benner, J., Helbling, D.E., Kohler, H.-P.E., Wittebol, J., Kaiser, E., Prasse, C., Ternes, T.A., Albers, C.N., Aamand, J., Horemans, B., Springael, D., Walravens, E., Boon, N., 2013. Is biological treatment a viable alternative for micropollutant removal in drinking water treatment processes? *Water Res.* 15, 5955–76. <https://doi.org/10.1016/j.watres.2013.07.015>
- Bianchi, M.A.G., Bianchi, A.J.M., 1982. Statistical sampling of bacterial strains and its use in bacterial diversity measurement. *Microb. Ecol.* <https://doi.org/10.1007/BF02011462>
- Blanch, A.R., Galofré, B., Lucena, F., Terradillos, A., Vilanova, X., Ribas, F., 2007.

- Characterization of bacterial coliform occurrences in different zones of a drinking water distribution system. *J. Appl. Microbiol.* 102, 711–721. <https://doi.org/10.1111/J.1365-2672.2006.03141.X>
- Boers, S.A., Jansen, R., Hays, J.P., 2019. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur. J. Clin. Microbiol. Infect. Dis.* 38, 1059–1070. <https://doi.org/10.1007/s10096-019-03520-3>
- De Carolis, E., Vella, A., Vaccaro, L., Torelli, R., Posteraro, P., Ricciardi, W., Sanguinetti, M., Posteraro, B., 2014. Development and validation of an in-house database for matrix-assisted laser desorption ionization-time of flight mass spectrometry-based yeast identification using a fast protein extraction procedure. *J. Clin. Microbiol.* 52, 1453–1458. <https://doi.org/10.1128/JCM.03355-13>
- Degerman, R., Dinasquet, J., Riemann, L., De Luna, S.S., Andersson, A., 2013. Effect of resource availability on bacterial community responses to increased temperature. *Aquat. Microb. Ecol.* 68, 131–142. <https://doi.org/10.3354/AME01609>
- Dextro, R.B., Delbaje, E., Cotta, S.R., Zehr, J.P., Fiore, M.F., 2021. Trends in free-access genomic data accelerate advances in Cyanobacteria taxonomy. *J. Phycol.* 57, 1392–1402. <https://doi.org/10.1111/JPY.13200>
- Douterelo, I., Boxall, J.B., Deines, P., Sekar, R., Fish, K.E., Biggs, C.A., 2014. Methodological approaches for studying the microbial ecology of drinking water distribution systems. *Water Res.* 65, 134–156. <https://doi.org/10.1016/j.watres.2014.07.008>
- Fish, K.E., Boxall, J.B., 2018. Biofilm microbiome (re)growth dynamics in drinking water distribution systems are impacted by chlorine concentration. *Front. Microbiol.* 9, 2519. <https://doi.org/10.3389/FMICB.2018.02519/BIBTEX>
- Gitis, V., Hankins, N., 2018. Water treatment chemicals: Trends and challenges. *J. Water Process Eng.* 25, 34–38. <https://doi.org/10.1016/j.jwpe.2018.06.003>
- Gunnarsdottir, M.J., Gardarsson, S.M., Figueras, M.J., Puigdomènech, C., Juárez, R., Saucedo, G., Arnedo, M.J., Santos, R., Monteiro, S., Avery, L., Pagaling, E., Allan, R., Abel, C., Eglitis, J., Hamsch, B., Hügler, M., Rajkovic, A., Smigic, N., Udovicki, B., Albrechtsen, H.J., López-Avilés, A., Hunter, P., 2020. Water safety plan enhancements with improved drinking water quality detection techniques. *Sci. Total Environ.* 698, 134185. <https://doi.org/10.1016/j.scitotenv.2019.134185>
- Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D., Cromeans, T.L., 2007.

- Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-Liter tap water samples. *Appl. Environ. Microbiol.* 73, 4218–4225. <https://doi.org/10.1128/AEM.02713-06>
- Hlordzi, V., Kuebutornye, F.K.A., Afriyie, G., Abarike, E.D., Lu, Y., Chi, S., Anokyewaa, M.A., 2020. The use of *Bacillus* species in maintenance of water quality in aquaculture: A review. *Aquac. Reports* 18, 2352–5134. <https://doi.org/10.1016/J.AQREP.2020.100503>
- Hou, L., Zhou, Q., Wu, Q., Gu, Q., Sun, M., Zhang, J., 2018. Spatiotemporal changes in bacterial community and microbial activity in a full-scale drinking water treatment plant. *Sci. Total Environ.* 625, 449–459. <https://doi.org/10.1016/j.scitotenv.2017.12.301>
- Kim, E., Cho, E.-J., Yang, S.-M., Kim, M.-J., Kim, H.-Y., 2021. Novel approaches for the identification of microbial communities in kimchi: MALDI-TOF MS analysis and high-throughput sequencing. *Food Microbiol.* 94, 103641. <https://doi.org/10.1016/j.fm.2020.103641>
- Kim, E., Cho, Y., Lee, Y., Han, S.-K., Kim, C.-G., Choo, D.-W., Kim, Y.-R., Kim, H.-Y., 2016. A proteomic approach for rapid identification of *Weissella* species isolated from Korean fermented foods on MALDI-TOF MS supplemented with an in-house database. <https://doi.org/10.1016/j.ijfoodmicro.2016.11.027>
- Kraková, L., Soltys, K., Otlewska, A., Pietrzak, K., Purkrtová, S., Savická, D., Purskářová, A., Bucková, M., Szemes, T., Budis, J., Demnerová, K., Gutarowska, B., Pangallo, D., 2017. Comparison of methods for identification of microbial communities in book collections: Culture-dependent (sequencing and MALDI-TOF MS) and culture-independent (Illumina MiSeq). *Int. Biodeterior. Biodegradation* 131, 51–59. <https://doi.org/10.1016/j.ibiod.2017.02.015>
- Lautenschlager, K., Hwang, C., Liu, W.T., Boon, N., Köster, O., Vrouwenvelder, H., Egli, T., Hammes, F., 2013. A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Res.* 47, 3015–3025. <https://doi.org/10.1016/j.watres.2013.03.002>
- Li, Q., Yu, S., Li, L., Liu, G., Gu, Z., Liu, M., Liu, Z., Ye, Y., Xia, Q., Ren, L., 2017. Microbial communities shaped by treatment processes in a drinking water treatment plant and their contribution and threat to drinking water safety. *Front. Microbiol.* 8, 2465. <https://doi.org/10.3389/FMICB.2017.02465>
- Li, W., Zhang, J., Wang, F., Qian, L., Zhou, Y., Qi, W., Chen, J., 2018. Effect of

- disinfectant residual on the interaction between bacterial growth and assimilable organic carbon in a drinking water distribution system. *Chemosphere* 202, 586–597. <https://doi.org/10.1016/j.chemosphere.2018.03.056>
- Liu, G., Tao, Y., Zhang, Y., Lut, M., Knibbe, W.J., van der Wielen, P., Liu, W., Medema, G., van der Meer, W., 2017. Hotspots for selected metal elements and microbes accumulation and the corresponding water quality deterioration potential in an unchlorinated drinking water distribution system. *Water Res.* 124, 435–445. <https://doi.org/10.1016/j.watres.2017.08.002>
- Liu, S., Gunawan, C., Barraud, N., Rice, S.A., Harry, E.J., Amal, R., 2016. Understanding, monitoring, and controlling biofilm growth in drinking water distribution systems. *Environ. Sci. Technol.* 50, 8954–8976. <https://doi.org/10.1021/ACS.EST.6B00835>
- Loiseau, C., Schlusshuber, M., Bigot, R., Bertaux, J., Berjeaud, J.M., Verdon, J., 2015. Surfactin from *Bacillus subtilis* displays an unexpected anti-*Legionella* activity. *Appl. Microbiol. Biotechnol.* 99, 5083–5093. <https://doi.org/10.1007/S00253-014-6317-Z>
- Mathieu, L., Bouteleux, C., Fass, S., Angel, E., Block, J.C., 2009. Reversible shift in the alpha, beta and gamma proteobacteria populations of drinking water biofilms during discontinuous chlorination. *Water Res.* 43, 3375–3386. <https://doi.org/10.1016/j.watres.2009.05.005>
- Mir, J., Morató, J., Ribas, F., 1997. Resistance to chlorine of freshwater bacterial strains. *J. Appl. Microbiol.* 82, 7–18. <https://doi.org/10.1111/J.1365-2672.1997.TB03292.X>
- Pinar-Méndez, A., Fernández, S., Baquero, D., Vilaró, C., Galofré, B., González, S., Rodrigo-Torres, L., Arahal, D.R., Macián, M.C., Ruvira, M.A., Aznar, R., Caudet-Segarra, L., Sala-Comorera, L., Lucena, F., Blanch, A.R., Garcia-Aljaro, C., 2021. Rapid and improved identification of drinking water bacteria using the Drinking Water Library, a dedicated MALDI-TOF MS database. *Water Res.* 203, 117543. <https://doi.org/10.1016/J.WATRES.2021.117543>
- Pinar-Méndez, A., Wangensteen, O.S., Præbel, K., Galofré, B., Méndez, J., Blanch, A.R., García-Aljaro, C., 2022. Monitoring bacterial community dynamics in a drinking water treatment plant: An integrative approach using metabarcoding and microbial indicators in large water volumes. *Water* 14, 1435. <https://doi.org/10.3390/W14091435>
- Prest, E.I., El-Chakhtoura, J., Hammes, F., Saikaly, P.E., van Loosdrecht, M.C.M.,

- Vrouwenvelder, J.S., 2014. Combining flow cytometry and 16S rRNA gene pyrosequencing: A promising approach for drinking water monitoring and characterization. *Water Res.* 63, 179–189. <https://doi.org/10.1016/j.watres.2014.06.020>
- Proctor, C.R., Hammes, F., 2015. Drinking water microbiology-from measurement to management. *Curr. Opin. Biotechnol.* 33, 87–94. <https://doi.org/10.1016/j.copbio.2014.12.014>
- Rhodes, E.R., Hamilton, D.W., See, M.J., Wymer, L., 2011. Evaluation of hollow-fiber ultrafiltration primary concentration of pathogens and secondary concentration of viruses from water. *J. Virol. Methods* 176, 38–45. <https://doi.org/10.1016/j.jviromet.2011.05.031>
- Rossel, S., Khodami, S., Martínez Arbizu, P., 2019. Comparison of rapid biodiversity assessment of meiobenthos using MALDI-TOF MS and Metabarcoding. *Front. Mar. Sci.* 6, 659. <https://doi.org/10.3389/FMARS.2019.00659/BIBTEX>
- Sala-Comorera, L., Blanch, A.R., Vilaró, C., Galofré, B., García-Aljaro, C., 2017. Heterotrophic monitoring at a drinking water treatment plant by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry after different drinking water treatments. *J. Water Health* 15, 885–897. <https://doi.org/10.2166/wh.2017.090>
- Sala-Comorera, L., Blanch, A.R., Vilaró, C., Galofré, B., García-Aljaro, C., 2016a. Pseudomonas-related populations associated with reverse osmosis in drinking water treatment. *J. Environ. Manage.* 182, 335–341. <https://doi.org/10.1016/J.JENVMAN.2016.07.089>
- Sala-Comorera, L., Caudet-Segarra, L., Galofré, B., Lucena, F., Blanch, A.R., García-Aljaro, C., 2020. Unravelling the composition of tap and mineral water microbiota: Divergences between next-generation sequencing techniques and culture-based methods. *Int. J. Food Microbiol.* 334, 108850. <https://doi.org/10.1016/j.ijfoodmicro.2020.108850>
- Sala-Comorera, L., Vilaró, C., Galofré, B., Blanch, A.R., García-Aljaro, C., 2016b. Use of matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry for bacterial monitoring in routine analysis at a drinking water treatment plant. *Int. J. Hyg. Environ. Health* 219, 577–584. <https://doi.org/10.1016/J.IJHEH.2016.01.001>
- Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner,

- P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12, 1–12. <https://doi.org/10.1186/s12915-014-0087-z>
- Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W., Vaishampayan, P., 2018. Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. *Front. Microbiol.* 9, 1–8. <https://doi.org/10.3389/fmicb.2018.00780>
- Shu, L.J., Yang, Y.L., 2017. *Bacillus* classification based on matrix-assisted laser desorption ionization time-of-flight mass spectrometry—Effects of culture conditions. *Sci. Rep.* 7, 15546. <https://doi.org/10.1038/S41598-017-15808-5>
- Skjevrak, I., Lund, V., Ormerod, K., Due, A., Herikstad, H., 2004. Biofilm in water pipelines; a potential source for off-flavours in the drinking water. *Water Sci. Technol.* 49, 211–217. <https://doi.org/https://doi.org/10.2166/WST.2004.0573>
- United Nations, 2020. Monitoring water and sanitation in the 2030. Agenda for Sustainable Development Integrated Monitoring Initiative for SDG 6. <https://www.unwater.org/publications/monitoring-water-and-sanitation-in-the-2030-agenda-for-sustainable-development/>.
- Von Gunten, U., 2003. Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res.* 37, 1469–1487. [https://doi.org/10.1016/S0043-1354\(02\)00458-X](https://doi.org/10.1016/S0043-1354(02)00458-X)
- W.Q. Betancourt and J.B. Rose, 2004. Drinking water treatment processes for removal of *Cryptosporidium* and *Giardia*. *Vet. Parasitol.* 219–234. <https://doi.org/10.1016/j.vetpar.2004.09.002>
- Wingender, J., Flemming, H.C., 2011. Biofilms in drinking water and their role as reservoir for pathogens. *Int. J. Hyg. Environ. Health* 214, 417–423. <https://doi.org/10.1016/j.ijheh.2011.05.009>

Comparison of bacterial diversity assessment by MALDI-TOF MS and 16S rRNA metabarcoding in a full-scale drinking water treatment plant

Anna Pinar-Méndez^{a,b,c}, Belén Galofré^a, Anicet R. Blanch^{b,c}, Cristina García-Aljaro^{b,c}

^a Aigües de Barcelona, Empresa Metropolitana de Gestió del Cicle Integral de l'Aigua, Barcelona, Spain

^b Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona (UB), Barcelona, Spain

^c Institut de Recerca de l'Aigua (IdRA), Universitat de Barcelona (UB), Barcelona, Spain

***Corresponding authors:**

Dr. Cristina García Aljaro

Mailing address: crgarcia@ub.edu

Avinguda Diagonal, 643, 08028 Barcelona, Spain

Phone: (+34) 934 02 14 84

Ms. Anna Pinar Méndez

Mailing address: apinar@aiguesdebarcelona.cat

General Batet 1-7, 08028 Barcelona, Spain

Phone: (+34) 93 247 98 12

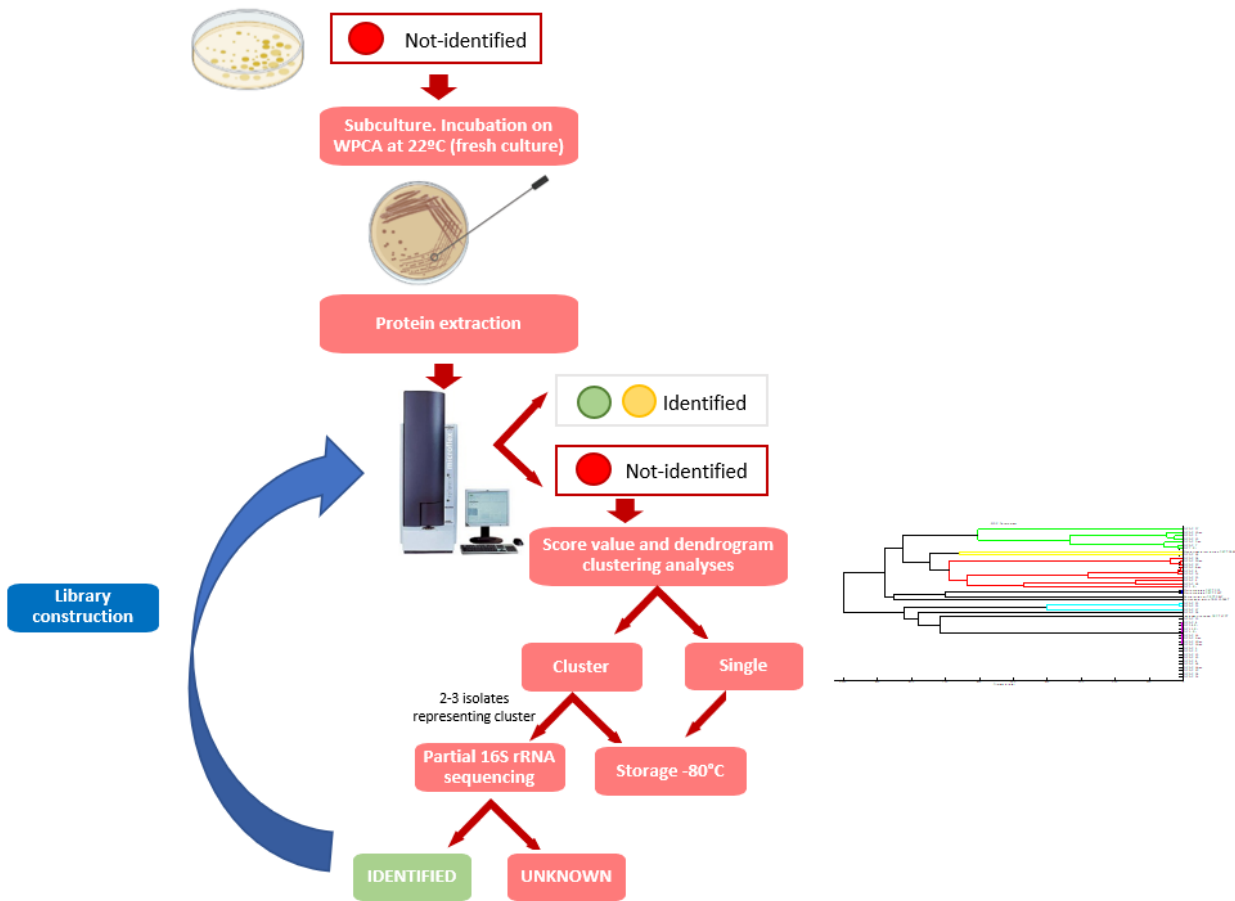


Figure S1. Workflow of non-identified isolates by MALDI-TOF MS. Isolates were reanalyzed and followed acid/acetonitrile protein extraction protocol to improve identification results. If identification results were successful, they contributed to bacterial unravel bacterial diversity at DWTP. Isolates with unreliable identification were processed by BioNumerics and clustered when possible. Some isolates were stored for sequencing purposes further on and updating DWL library.

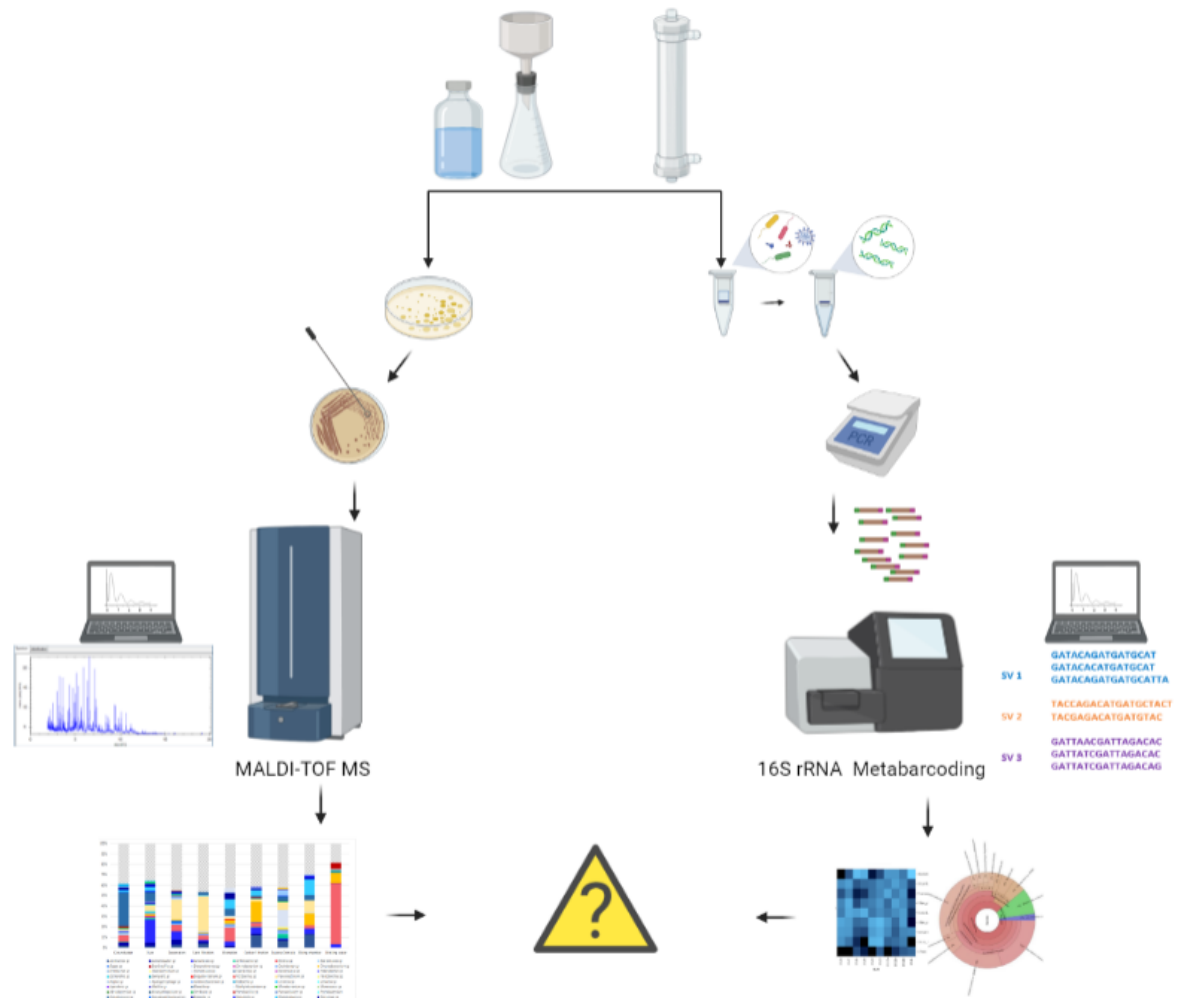


Figure S2. Workflow of bacterial community identifications using MALDI-TOF MS and metabarcoding of DWTP samples. Heterotrophic bacteria isolated from water samples was used for MALDI-TOF MS identification at genus or species level using two databases: BDAL and DWL. DNA extracted from water samples was used to analyze relative abundant reads of bacterial communities by 16S rRNA metabarcoding using SILVA database. Both approaches were used to describe culturable and unculturable bacteria at DWTP stages.

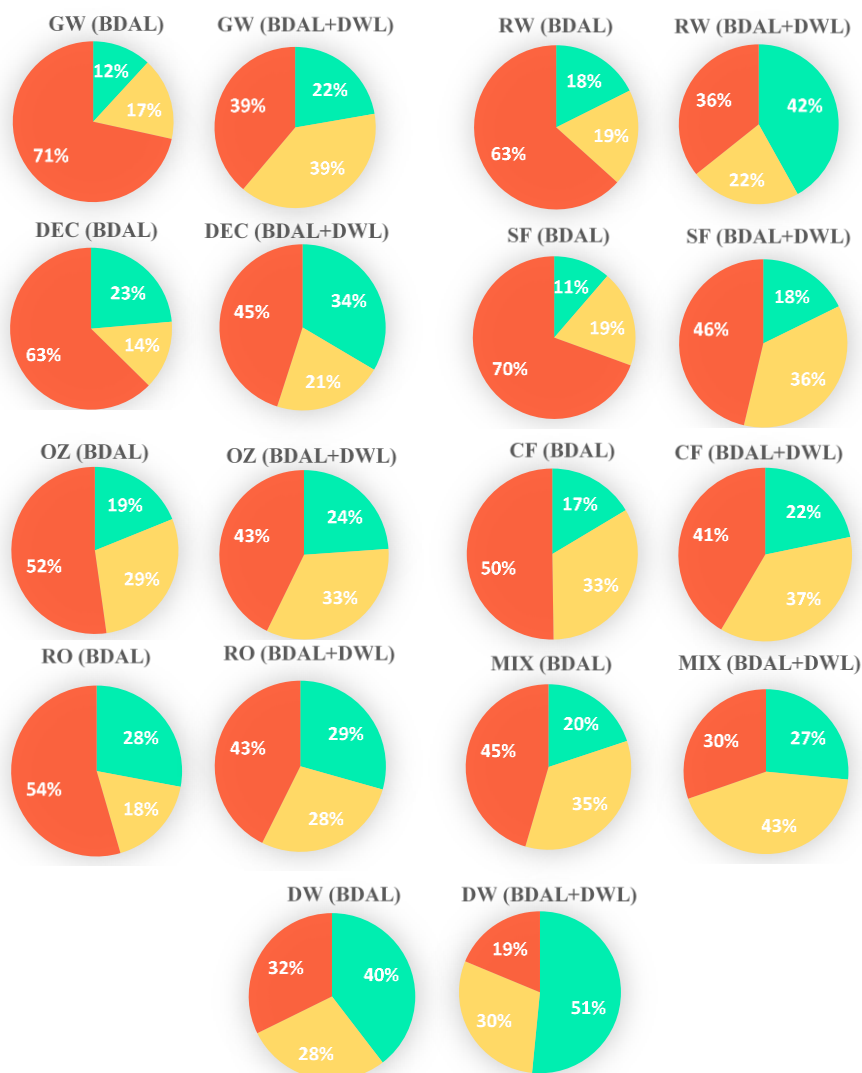


Figure S3. Comparison of the identification percentage for all DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW), using the Bruker Daltonics (BDAL) database and the extended database Bruker Daltonics (BDAL) with the Drinking Water Library (DWL). Results were classified according to following scores: highly probable species (≥ 2.300 – 3.000 ; green), secure genus and probable species (2.000 – 2.29 ; yellow); probable genus (1.700 – 1.999 ; yellow) or unreliable results (≤ 1.699 ; red).

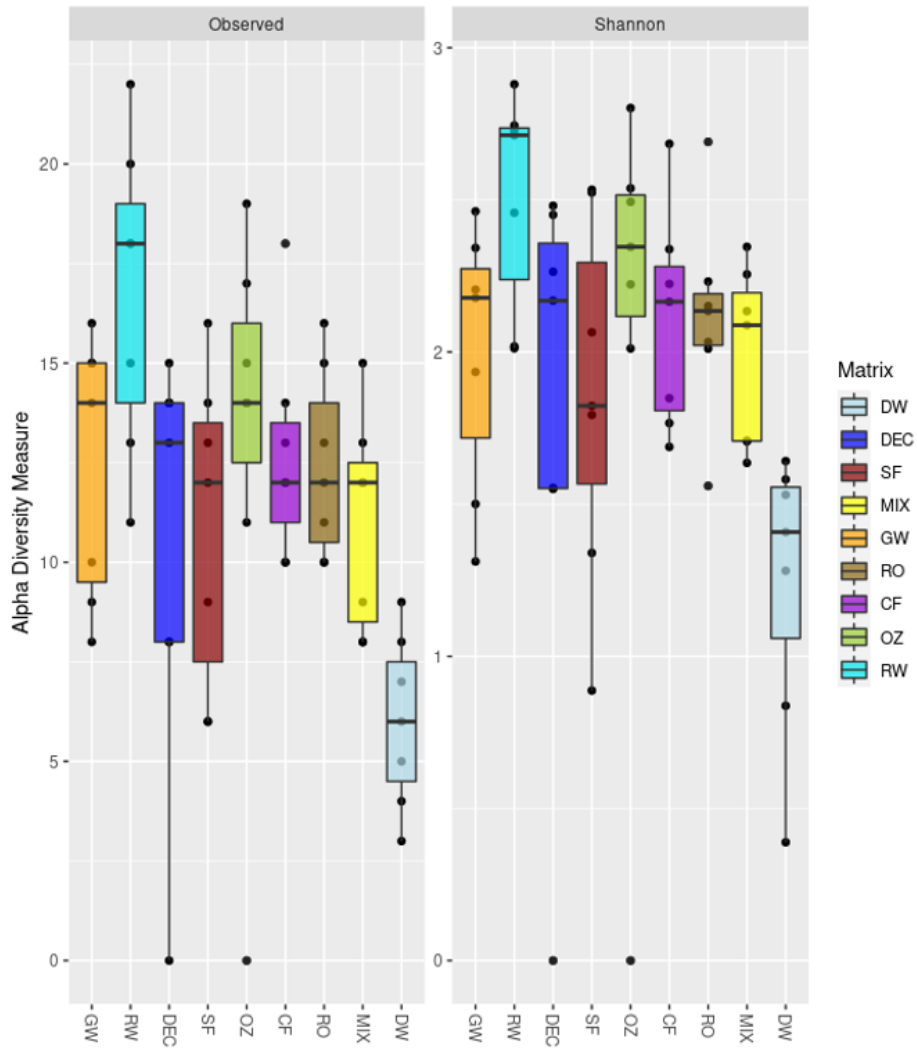


Figure S4. Boxplots representing alpha diversity indices of genus richness (Observed) and diversity (Shannon) for all DWTP samples: groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing chamber (MIX) and post-chlorination water (DW). After chlorination procedure diversity was significantly reduced in DW.

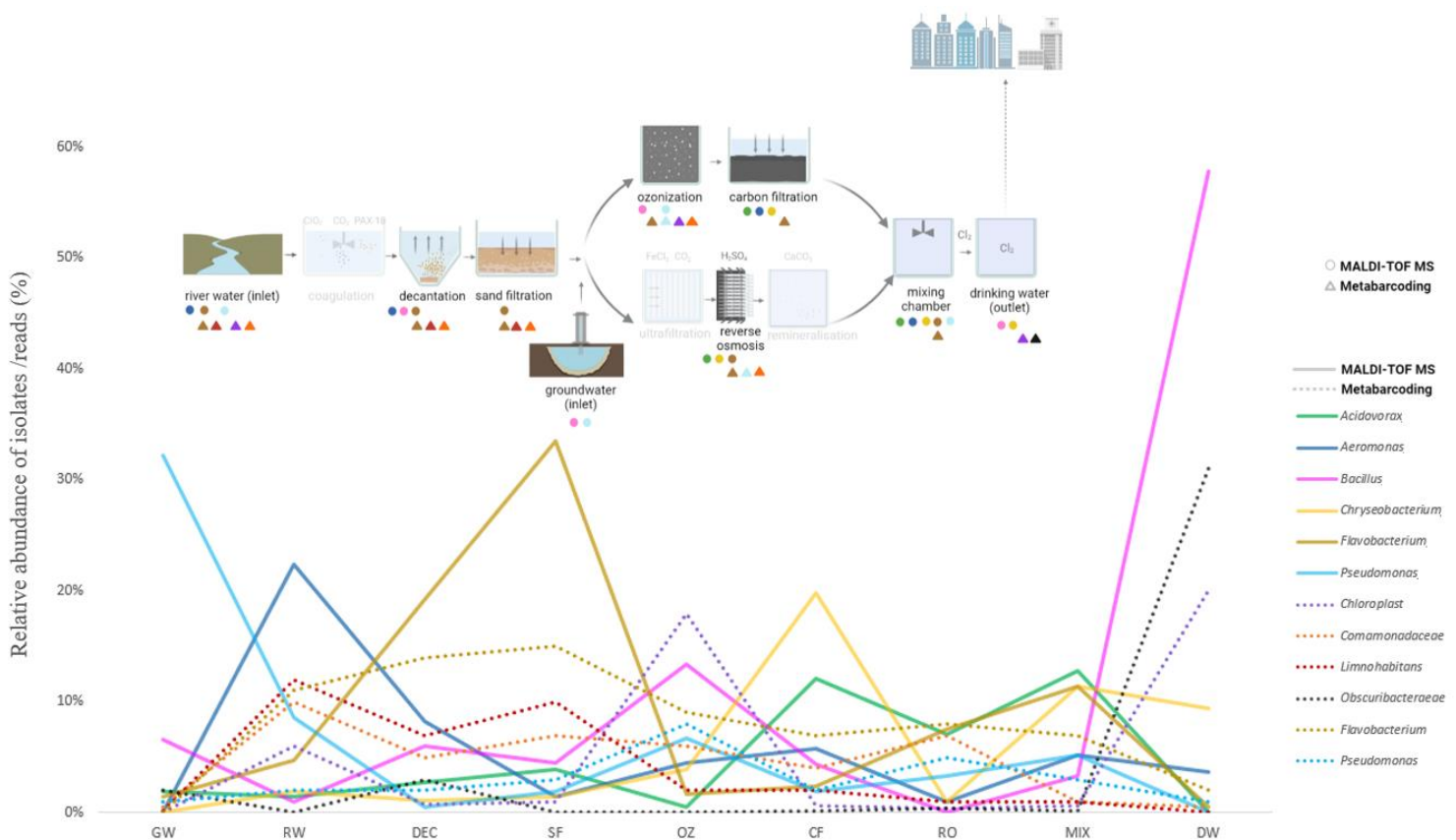


Figure S5. Fluctuation graph along the DWTP stages of the 6 most abundant genera according to MALDI-TOF MS results (solid line) and the 6 most abundant genera according to metabarcoding results (dotted line). Hotspots (abundance >5%) are pointed into the DWTP schematic diagram with a circle for MALDI-TOF results or a triangle for metabarcoding. Samples corresponded to groundwater (GW), river water (RW), decantation (DEC), sand filtration (SF), ozonization (OZ), carbon filtration (CF), reverse osmosis (RO), mixing.

Flavobacterium and *Pseudomonas* were the most abundant genera in the DWTP according to both techniques. *Flavobacterium*, with a relative abundance >5%, was located mainly in RW, DEC, SF, OZ, CF, RO and MIX, although in OZ and CF samples it was only detected by metabarcoding. *Pseudomonas* (>5%) was identified in RW, GW, OZ, MIX only by MALDI-TOF MS, in OZ by both techniques, and additionally in RO by metabarcoding.

Regarding the top six genera according to the MALDI-TOF MS approach (all with a relative abundance >5%), *Aeromonas* and *Bacillus* were identified in source water, treatment stages and the final DW, whereas *Acidovorax* and *Chryseobacterium* were mainly detected in the late stages and DW (Fig S5).

Regarding the six most prevalent genera according to the metabarcoding (with relative abundance reads >5%), *Limnohabitans*, unknown Comamonadaceae and unknown Chloroplast were detected in source water and treatment stages, while unknown Obscuribacteraceae was mainly found in DW. Therefore, hotspots for these genera consisted of early treatment processes for *Limnohabitans* (RW, DEC and SF); RW, OZ and DW for unknown Chloroplast; a range of

treatments for unknown Comamonadaceae (RW, DEC, DS, OZ and RO), and finally, unknown Obscuribacteraceae presented a high abundance only in post-chlorination DW.

DISCUSSIÓ GENERAL

A les plantes de tractament d'aigua potable s'utilitzen diversos processos de tractament per a eliminar els contaminants i microorganismes de l'aigua crua. Aquests processos poden constar de diverses etapes i diferents tecnologies de tractament, que conjuntament constitueixen barreres efectives que garanteixen la qualitat de l'aigua potable. L'ETAP de Sant Joan Despí, operada per Aigües de Barcelona i objecte de la present tesi doctoral, és particularment complexa degut, entre d'altres, al context d'ubicació de la planta de tractament i la qualitat dels recursos hídrics que utilitza. Aquesta complexitat implica que el tractament global de l'ETAP consti de 3 fases de tractament: pretractament, tractament convencional amb ozonització i filtració per carbó, tractament avançat per ultrafiltració o osmosi inversa, i cloració final.

El control de l'aigua mitjançant l'ús d'indicadors bacteriològics de la contaminació fecal i el nombre de bacteris heteròtrofs, han estat clau per l'assegurament de la qualitat de l'aigua des de la seva implantació a les successives directives d'aigües de consum, evitant exposar a la població a riscos sanitaris. La rellevància de la transmissió de patògens vírics a les aigües per ruta fecal-oral, ha reforçat la necessitat d'implantar també l'ús d'indicadors vírics, ja recollits a l'actual Directiva EU 2020/2184, per complimentar la informació de contaminació fecal que en ocasions no era possible detectar únicament amb els indicadors bacteriològics.

Malgrat l'eficiència dels processos de potabilització, que asseguren la qualitat sanitària de l'aigua final, l'eliminació de la totalitat dels bacteris a l'aigua potable no és possible (ni tampoc l'objectiu), i per tant és imprescindible acceptar la seva gran ubiqüitat. Poder identificar aquesta microbiota i la seva abundància relativa, permet entendre el seu comportament, i aplicar aquest coneixement per una millor gestió de la qualitat de l'aigua. Actualment la normativa vigent no contempla l'estudi de la totalitat de les comunitats bacterianes, i per tant únicament amb l'anàlisi d'indicadors, s'obté una informació limitada des del punt de vista dels processos bioquímics i biològics que tenen lloc a la planta en el tractament de l'aigua.

L'objectiu principal d'aquesta tesi ha estat aprofundir en el coneixement de les comunitats microbianes implicades en els processos de tractament de l'aigua a l'ETAP de Sant Joan Despí. El coneixement d'aquestes comunitats s'ha abordat des de dues perspectives diferents, l'anàlisi de les comunitats bacterianes cultivables utilitzant un mètode basat en cultiu (MALDI-TOF MS) i l'anàlisi de les comunitats totals utilitzant un mètode independent de cultiu (*metabarcoding* 16S ARNr).

Aquesta tesi ha comportat l'aplicació d'un conjunt d'avenços metodològics pel que fa l'anàlisi de les comunitats en un procés de potabilització en funcionament i operativitat real. L'adaptació d'aquestes metodologies ha requerit ajustar-les per a analitzar mostres amb baixa concentració

microbiana i ha comportat la necessitat de disposar d'una llibreria de perfils de MALDI-TOF MS adaptada a mostres d'aigua.

Per assolir el segon objectiu de la tesi, s'han utilitzat i optimitzat dues les metodologies de concentració diferents: la filtració membrana i la ultrafiltració DEUF. Per superar el problema analític de la baixa concentració de microorganismes a les etapes més avançades de l'ETAP, la ultrafiltració DEUF es va utilitzar a les etapes on més fàcilment es perdia la senyal dels indicadors de contaminació fecal (pous, i de tractament convencional en endavant). Aquest mètode va permetre concentrar grans volums d'aigua (fins a 1.100 L) augmentant la senyal dels indicadors i ajudant a una caracterització més representativa de les comunitats bacterianes. D'aquesta manera es van poder detectar i quantificar indicadors microbians en etapes més avançades del tractament quan la concentració esperable d'aquests és molt baixa i habitualment no detectada pels mètodes de filtració convencionals legiscats (basats en l'anàlisi de 100 mL). També va permetre observar una gran diversitat per MALDI-TOF i *metabarcoding* sobretot en mostres amb baixa concentració de microorganismes (Salter et al., 2014) com els processos de desinfecció.

Totes les mostres d'aigua potable analitzades complien amb els estàndards de qualitat microbiològics. El tercer objectiu de la tesi va poder determinar la cinètica d'eliminació dels indicadors al llarg dels tractaments. A excepció del riu, pretractament i tractament convencional, on si es va observar la presència d'indicadors de contaminació fecal amb una reducció progressiva al llarg dels tractaments d'entre 4 a 6 log respecte el riu, la resta de matrius van presentar resultats negatius (<1 UFC/100 mL) per a aquests paràmetres. No obstant, gràcies a l'ús de DEUF es van detectar algunes mostres positives d'indicadors també al tractament avançat i a la cambra de mescla en concentracions baixes, però no a l'aigua potable. Tanmateix els valors de HPC es van reduir progressivament amb un descens significatiu després de la cloració (de 4,62 log₁₀ a 0,027 log₁₀ CFU/mL).

Es va observar un factor estacional durant les estacions més càlides, on l'abundància de bacteris heteròtrofs a les mostres de riu, ozonització i aigua potable va augmentar significativament. Cal remarcar però que la consigna de dosificació d'ozó a l'estiu és més baixa que a l'hivern per a evitar la formació de bromat, un subproducte de l'ozonització i potencial carcinogen humà produït a temperatures més altes (Von Gunten, 2003). Per tant, un desinfectant residual més baix pot contribuir al creixement bacterià en les estacions més càlides (Li et al., 2018). A més, tot i que les mostres de filtració per carbó no van mostrar variacions estacionals, a baixes temperatures es van observar 2 log de HPC superiors a les mostres de l'etapa anterior (ozonització), que podria indicar un possible recreixement als filtres. En aquest context, és vital

per reforçar la seguretat de l'aigua emprant altres eines per l'estudi de les comunitats bacterianes. La seva importància rau en entendre la seva dinàmica espacial i temporal als sistemes d'aigua potable (C. Li et al., 2017), conèixer en profunditat quins paràmetres o tractaments poden modificar l'estructura bacteriana i com aquesta pot influir en la qualitat de l'aigua, detectar l'ocurrència d'esdeveniments anormals abans d'una degradació significativa de la qualitat i la seguretat de l'aigua, així com establir un punt de partida cercant indicadors que permetin detectar alteracions en el bon funcionament de la planta. Per exemple, els processos de desinfecció, com la cloració durant la potabilització, fan disminuir significativament la concentració bacteriana a l'aigua, però es poden produir recreixements que alterin la composició microbiològica que es trobava a l'aigua potable final de l'ETAP abans de la seva impulsió a la xarxa. Aquests canvis en la microbiota també es poden produir al llarg dels tractaments de potabilització per un efecte selectiu del tractament envers la microbiota resistent influenciant la composició de l'aigua final. Malgrat que les concentracions bacterianes a les aigües potables solen ser baixes, diversos estudis posen de manifest una gran diversitat de gèneres presents a les aigües de consum (Li et al., 2018; Luo et al., 2021; Pinto et al., 2012).

Davant de la necessitat de disposar de mètodes ràpids per a la caracterització de la microbiota bacteriana de l'ETAP, s'ha optimitzat el sistema d'espectrometria de masses MALDI-TOF que es basa en la detecció de proteïnes ribosomals majoritàriament, per a l'anàlisi de bacteris de l'aigua. En microbiologia, els mètodes d'identificació es basen principalment en procediments de cultiu, com l'ús de medis cromogènics o fenotipat bioquímic (API, Phene Plate System, Biolog, etc.), o mètodes moleculars, com la seqüenciació massiva d'alt rendiment d'amplicons del gen 16S ARNr. Dels diferents mètodes basats en cultiu per a la identificació de bacteris, el fenotipat bioquímic és altament discriminatori per a la identificació en mostres d'aigua, però és laboriós i requereix molt de temps, sobretot en les avaluacions de la biodiversitat, ja que cada identificació requereix entre 24 hores i fins a 5 dies d'incubació (Blanch et al., 2007; Hou et al., 2018; Sala-Comorera et al., 2016a, 2016b; Tsuchida et al., 2020). La tècnica de MALDI-TOF MS és una eina d'identificació alternativa als mètodes de fenotipat bioquímic i molecular amb un poder resolutiu comparable o superior, però amb una destacable rapidesa i facilitat d'anàlisi que permet obtenir resultats en pocs minuts i una caracterització de diverses soques en un sol dia (Sala-Comorera et al., 2016b; Sárvári et al., 2018; Timperio et al., 2017). Tot i que la inversió inicial de l'equip és elevada, altres estudis han mostrat un estalvi net anual del 87,8% en costos de reactius en comparació amb els mètodes tradicionals (Tan et al., 2012). El seu gran poder resolutiu s'ha comparat fins i tot a la seqüenciació massiva d'amplicons (Kraková et al., 2017), encara que la seva capacitat d'identificació depèn directament dels espectres de referència

disponibles a la base de dades de l'equip. La gran majoria de llibreries de perfils d'espectres de masses estan orientats a àmbits clínics, per aquest motiu diversos autors han destacat la necessitat de creació de bases de dades internes per aconseguir una bona resolució taxonòmica, especialment per als estudis d'ecologia microbiana (De Carolis et al., 2014; Kopcakova et al., 2014; Rahi et al., 2016; Seuylemezian et al., 2018). Amb el que, el primer objectiu de la tesi va ser posar a punt la metodologia per tal de generar una base de dades de perfils proteics (que s'anomenen llibreries) de bacteris aïllats de l'aigua de la xarxa de distribució, d'aigües minerals naturals, de fonts i de processos de potabilització, així com de soques de referència d'origen ambiental relacionades amb l'aigua. Aquest objectiu s'engloba dins del marc del projecte RETOS Drinking Water Library (RTC-2015-4496-2) del Ministerio de Ciencia e Innovación en que el que el laboratori d'Aigües de Barcelona va participar (i en el que vaig participar com a doctoranda), conjuntament amb la Universitat de Barcelona i la Colección Espanyola de Cultivos Tipo.

Així, la llibreria resultant anomenada *Drinking water Library* (DWL) va permetre disposar en total de 319 perfils d'espectres principals (MSP) incorporant un total de 96 gèneres diferents, dels quals 44 gèneres i 189 espècies (67 potencialment nous tàxons), no es trobaven a la llibreria original de l'equip (BDAL). Aquesta nova configuració va permetre una millora significativa en la identificació de les soques que restaven sense identificar i que es van re-identificar mitjançant l'ús conjunt de BDAL i DWL, reduint d'un 55% a un 8% les soques sense identificar, i augmentant d'un 29% a un 76% la identificació a nivell d'espècie.

La tecnologia MALDI-TOF MS, gràcies a la creació prèvia de la llibreria DWL, ens ha permès assolir el quart objectiu de la tesi que ha consistit en monitoritzar i caracteritzar la planta de tractament d'aigua potable (ETAP) de Sant Joan Despí durant gairebé dos anys. Així, rere la creació de la llibreria, es van processar 1.807 aïllaments de colònies recuperades de les plaques de HPC de 9 etapes diferents de tractament: captació (riu i pous), pretractament (decantació i filtració per sorra), tractament convencional (ozonització i filtració per carbó), tractament avançat (osmosi inversa) i aigua final (cambra de mescla i aigua potable clorada). L'objectiu d'aquest estudi es va centrar en la identificació dels bacteris heterotròfics cultivables obtinguts prèviament de l'anàlisi de qualitat microbiològica analitzada segons la Directiva d'aigües, i per tant aquests bacteris estaven sotmesos a unes condicions controlades de creixement, amb un medi concret i una temperatura i un temps d'incubació determinats, que són els que s'utilitzen en l'anàlisi requerit per la legislació.

La seva identificació es va fer mitjançant l'ús conjunt de les llibreries BDAL i DWL, i la taxa d'identificació en les diferents etapes del tractament va variar del 54% al 81%, essent la major

resolució per a les mostres d'aigua potable probablement per la major representativitat de MSP (perfil d'espectre principal) a la llibreria destinada a aigües de consum. L'ús de la DWL va permetre una clara millora en la identificació, ja que únicament amb l'ús de la llibreria BDAL el percentatge d'identificació es situava entre un 29% i un 68%. Malgrat l'ús combinat de les dues llibreries, una part de les soques no es van poder identificar (686), però el seu agrupament en funció dels seus espectres va permetre realitzar un estudi de diversitat d'aquests bacteris no identificats. Així, es van agrupar per les seves similituds en 91 clústers, mentre que 121 soques van presentar un perfil proteic únic. Amb l'objectiu de continuar actualitzant la biblioteca de perfils proteics per millorar la caracterització dels bacteris no identificats, es van seleccionar i seqüenciar diversos representants d'aquests clústers per tal de continuar properament amb l'actualització continua de la llibreria, tal com expressen molts autors d'aquesta necessitat per l'èxit d'identificació de soques d'interès (De Carolis et al., 2014; Kim et al., 2016; Kopcakova et al., 2014; Rahi et al., 2016; Santos et al., 2016; Seuylemezian et al., 2018).

Respecte la caracterització i diversitat per MALDI-TOF MS es va observar una gran diversitat de gèneres bacterians (57 diferents identificats) i canvis successius en la diversitat al llarg de les etapes del tractament, principalment dominades per Proteobacteria i Bacteroidota, amb comunitats molt diverses al riu i aigües subterrànies, fins a una disminució significativa de la diversitat després de l'etapa de cloració, on es va produir un emplaçament dominat per *Bacillus* (Firmicutes) a l'aigua potable.

D'altra banda, donades les limitacions que suposa caracteritzar únicament la fracció bacteriana cultivable que representa una fracció molt petita (<1%) de la biodiversitat real del món microbià de l'aigua (Amann et al., 1995), es va plantejar el cinquè objectiu de la tesi per a la caracterització de la microbiota bacteriana de l'ETAP amb mètodes moleculars independents de cultiu. Per aquest nou objectiu, es va seleccionar la tècnica de *metabarcoding* (seqüenciació massiva) on el gen diana és un fragment d'uns 250 pb del gen del 16S rRNA del fragment de la regió variable 4. El gen de l'ARN ribosòmic 16S és el gen marcador més utilitzat en ecologia microbiana. Aquest gen està molt conservat per al qual s'han desenvolupat bases de dades i està acceptat per la comunitat científica per realitzar estudis filogenètics (Baker et al., 2003; Bukin et al., 2019; Morgan and Huttenhower, 2012). No obstant, la utilització d'aquest gen també presenta una sèrie de limitacions, com biaixos derivats de l'elecció del primer o pel fet que les estratègies de fragments curts, en la major part dels casos, no és possible realitzar una assignació taxonòmica de manera fiable a nivell de gènere o espècie (Janda and Abbott, 2007). D'entre la resta de regions variables del gen 16S ARNr emprades per a la identificació taxonòmica, es va seleccionar

la regió variable 4 per estar àmpliament utilitzada en diversos estudis (Caporaso et al., 2012; Liu et al., 2020) amb l'objectiu de poder comparar els resultats obtinguts amb altres estudis.

Un dels desavantatges de treballar amb el gen 16S és que es tracta d'un gen multicopia, i les espècies bacterianes contenen més d'una còpia d'operons d'ARN ribosòmic en els seus genomes, fins i tot algunes en contenen fins a 15 (Espejo and Plaza, 2018; Louca et al., 2018). Donada aquesta diferència entre el número de còpies del gen 16S, el nombre de variants de seqüència d'amplicons (ASV) estan esbiaixats cap als taxons que tenen un número superior de còpies, fet que pot induir un biaix en els resultats. Tot i que hi ha eines bioinformàtiques que intenten corregir aquests biaixos mitjançant l'estimació de còpies per taxó, com PICRUST (Langille et al., 2013) o CopyRighter (Angly et al., 2014), la seva precisió està força qüestionada, amb el que molts estudis recomanen no corregir, ja que una predicció inexacta del número de còpies de 16S per taxó pot introduir un soroll important als perfils de la comunitat bacteriana, que pot empitjorar els biaixos originals assumint un número de còpies desigual, a banda que dificulta la comparació entre estudis, en particular quan els mètodes de predicció són diferents. Per aquest motiu, i altres factors que poden produir biaixos, com l'extracció d'ADN o l'elecció dels primers per la seqüenciació dirigida per amplicons, es pot parlar d'abundància relativa o de proporcions, però actualment la seqüenciació massiva no és una eina quantitativa suficientment robusta. No obstant, es podria considerar com semi-quantitativa, perquè permet comparar dins de cada mostra i entre mostres la diferència en el nombre de seqüències, l'abundància relativa i com es diferencien, permetent conèixer la diferència entre proporcions dels tàxons dins i entre les mostres.

Tanmateix, atesa la seva limitació en quant a la quantificació, es poden dissenyar estratègies que permetin tenir una aproximació més precisa, com per exemple dissenyar els primers (encebadors) per a amplificar les espècies més dominants i quantificar-les mitjançant PCR quantitativa.

En la present tesi doctoral, l'estudi de les comunitats bacterianes mitjançant *metabarcoding* va revelar una gran diversitat de tàxons així com canvis de les poblacions en les diferents etapes. La successió de tàxons va marcar un clar predomini del filum Proteobacteria, Bacteroidota i Bdellovibrionota durant tot el tractament, amb un gran canvi a l'aigua final amb un clar predomini de Cyanobacteria. Les etapes amb major diversitat van correspondre a pous, cambra de mescla i filtració per carbó, amb una observada pèrdua de diversitat rere les etapes de desinfecció (decantació (per l'addició de diòxid de clor) , ozonització i cloració) i osmosi inversa, i per contra, un augment diversitat en els filtres (sorra i carbó). La major diversitat observada als

filtres podria indicar un microbioma establert associat a la filtració, tal com s'ha descrit en altres estudis (Pinto et al., 2012). En canvi, la diversitat de l'aigua final va estar marcada per l'efecte del clor que és un fort bactericida que regula la presència dels pocs grups resisteixen el procés de desinfecció. En aquest cas amb clar predomini del filum Cyanobacteria però amb detecció de generes resistents a la desinfecció, i potencials formadors de biofilms, com *Aeromonas*, *Clostridium* o *Pseudomonas* (Luo et al., 2021).

La detecció de certs ASV sota les condicions operatives normals de l'ETAP és d'interès, ja que poden proporcionar noves eines de seguiment per detectar tractaments deteriorats, un potencial aplicació que aquesta tesi ha permès posar de manifest, i per la que encara és necessita continuar treballant en l'àmbit de la recerca, però amb un peu ja posat a la operativitat de l'ETAP.

També va permetre observar una eventualitat a la planta derivada del trencament del col·lector de salmorres riu amunt que va produir una aturada de la captació superficial i en conseqüència, canvi a captació d'aigua de pous. Possiblement atribuïble a un canvi sobtat de temperatura de l'aigua de la captació (11°C de diferencia entre ambdós orígens), es podria haver modificat temporalment les propietats d'adsorció del filtre de carbó en el que sembla va derivar en un augment d'algunes poblacions bacterianes que podrien estar desadsorbint-se del filtre, com s'ha descrit aquest fenomen en altres estudis (Moona et al., 2019). Atès a que només es va produir un sol cop, les dades cal analitzar-les amb precaució i no es poden extreure conclusions suficientment recolzades amb els presents resultats.

La tendència observada en la caracterització bacteriana per *metabarcoding* en comparació amb altres estudis a l'ETAP o de l'aigua potable de la xarxa de distribució, coincideix amb un clar predomini de Proteobacteria als tractaments, però es diferencia en que la majoria d'estudis reporten Proteobacteria també com a filum dominant a l'aigua potable (Atnafu et al., 2021; Chao et al., 2013; Hou et al., 2018; Q. Li et al., 2017; Pinto et al., 2012; Poitelon et al., 2010; Sala-Comorera et al., 2019; Wang et al., 2018), a diferència del que vàrem observar al nostre sistema amb un clar predomini de Cyanobacteria, com també s'ha observat en alguns estudis (Jalili et al., 2021; Liu et al., 2016; Pei et al., 2017). Aquests resultats variables podrien estar marcats per diversos factors com les diferents ubicacions geogràfiques de les plantes, el tipus de recurs hídic emprat per la captació, així com la seva qualitat, l'ús (o no) de desinfectant, etc. que, tot i que comparteixen similituds, configuren un hàbitat únic a cada ETAP o xarxa de distribució.

Un dels factors que va influenciar en la diversitat i la taxonomia va ser la temperatura amb una correlació significativa d'augment de diversitat en temperatures més elevades, especialment a

la decantació i filtració per carbó, també observada a altres estudis co-relacionant temperatura elevada i els nivells de nutrients elevats, amb un augment de la taxa de creixement de les comunitats bacterianes (Degerman et al., 2013; Douterelo et al., 2016; Hou et al., 2018; Pinto et al., 2012; Romero et al., 2021). Aquests resultats poden ser d'interès per als gestors d'aigua potable ja que donen suport a la preocupació que el canvi climàtic podria promoure el creixement de poblacions bacterianes heterotròfiques als sistemes aquàtics davant d'un escenari d'escalfament global on els nivells de nutrients s'espera que augmentaran considerablement. En aquests entorns, la composició de la comunitat bacteriana podria canviar i les seves taxes de creixement podrien augmentar (Calero Preciado et al., 2021). La implicació de les altes temperatures també podrien afavorir la presència de certs patògens bacterians (Calero Preciado et al., 2021; van der Wielen and van der Kooij, 2013).

En aquest estudi, al llarg dels tractaments es va poder detectar seqüències de gèneres pels quals s'han descrit espècies patògenes de transmissió hídrica com *Aeromonas*, *Arcobacter*, *Legionella*, *Mycobacterium*, *Escherichia-Shigella* i *Pseudomonas*, o l'ordre Campylobacterales, eliminades a l'aigua potable, amb excepció de *Pseudomonas* i *Aeromonas*, demostrant l'eficàcia dels tractaments de potabilització com han descrit en altres estudis (Collado et al., 2010).

El sisè i últim objectiu de la tesi va consistir en comparar les semblances i similituds entre les diferents aproximacions utilitzades.

Respecte a l'aigua potable, la detecció d'indicadors cultivables (segons la Directiva) en comparació amb la detecció de seqüències corresponents als indicadors (segons *metabarcoding*), van mostrar discrepàncies en quant a la seva detecció, indicant que l'ús de les eines moleculars per si soles podrien malinterpretar la qualitat de l'aigua (Acharya et al., 2020). Pel paràmetre HPC, mostres que compleixen la normativa i no presenten canvis anòmals, s'han registrat tant per MALDI-TOF MS com per *metabarcoding* resultats diferents de diversitat bacteriana, tal com han observat altres investigadors que han comparat recomptes d'heteròtrofs amb altres tècniques moleculars (Lautenschlager et al., 2013; Prest et al., 2014).

Tant amb el mètode de MALDI-TOF com el de *metabarcoding*, la diversitat bacteriana va estar influenciada pel tractament aplicat, especialment la cloració, que va provocar una forta reducció de l'aigua potable final. Les tendències en la diversitat, però, variaven segons el mètode d'identificació utilitzat. Mentre que amb *metabarcoding* la diversitat va ser més alta a pous, cambra de mescla i filtració per carbó, amb pèrdua de diversitat a les desinfeccions i un augment diversitat en els filtres acabant amb una reducció dràstica a l'aigua potable, amb MALDI-TOF MS, la diversitat va ser més alta a rius i pous, amb un descens al pretractament, seguit d'un augment

amb la ozonització i la resta d'etapes es van observar petites diferències no significatives, excepte la reducció significativa posterior a la cloració a l'aigua final. No obstant, s'ha de remarcar que en el cas del MALDI-TOF MS la diversitat es va analitzar a un rang taxonòmic superior, a nivell de gènere, que en el cas del *metabarcoding* que es va realitzar a nivell d'ASV.

Pel que fa a les assignacions taxonòmiques, a nivell de fílum, en termes generals les dues tècniques van donar resultats similars, mostrant el domini de Proteobacteria al llarg dels tractaments excepte a l'aigua final, en què la cloració va exercir una forta pressió selectiva, donant lloc a una diversitat més baixa i un desplaçament taxonòmic cap a Cyanobacteria (*metabarcoding*) o Firmicutes (MALDI-TOF MS).

Ambdues tècniques van coincidir identificant pocs però els mateixos gèneres que es trobaven en major proporcions a algunes etapes (ex. *Flavobacterium* a decantació, filtració per sorra, osmosi inversa i cambra de mescla), però majoritàriament diferien a nivells baixos d'identificació sent el predominant la família Obscuribacteraceae segons *metabarcoding*, i en canvi *Bacillus* segons MALDI-TOF MS. Alguns gèneres només es van identificar mitjançant una de les dues tècniques, per exemple, per *metabarcoding* es va detectar una major varietat de gèneres Alfaproteobacteria i Gammaproteobacteria a l'aigua potable no detectables per MALDI-TOF com *Bosea*, *Comamonas*, *Methylobacterium*, *Pseudomonas*. Aquests gèneres han estat associats a la formació de biofilms i afavorits per altes concentracions de clor (Fish and Boxall, 2018; Mathieu et al., 2009). Per tant és de vital importància aprofundir en la seva detecció, en particular la seva viabilitat, per a evitar un possible recreixement a la xarxa de distribució que pugui derivar en canvis organolèptics o donar lloc a l'aparició de patògens oportunistes.

Ambdues tecnologies tenen avantatges i inconvenients. L'anàlisi per *metabarcoding* proporciona una visió més detallada de la totalitat de comunitats bacterianes a les mostres d'aigua que pel MALDI-TOF MS, revelant una major diversitat i tàxons addicionals. No obstant, no permet diferenciar entre cèl·lules vives o mortes, de manera que les dades resultants per si soles dificulten l'avaluació dels perills associats a la detecció de tàxons específics. D'altra banda, la identificació per MALDI-TOF MS pot veure's influenciada per factors com el medi de cultiu o condicions de creixement, mètode d'extracció de proteïnes, la base de dades d'espectres de masses, però pot identificar de manera ràpida i fiable, i a baix cost els microorganismes viables, que permet un ús en rutina, però, els bacteris cultivables representen una fracció petita del microbioma global i per si sola podria subestimar la diversitat a l'aigua.

Per tant, acoblar als indicadors microbiològics de la contaminació fecal la informació integrada de les eines proteòmiques i metagenòmiques, ens ha permès observar variacions de les

comunitats bacterianes en diferents condicions ambientals així com conèixer les condicions normals de funcionament del sistema i detectar anomalies no captades pels mètodes requerits per legislació. Així doncs aquestes tècniques ajuden a millorar la nostra comprensió sobre el funcionament i l'eficiència de les diferents etapes del tractament de l'aigua. Per exemple, gràcies a la seqüenciació massiva o *metabarcoding* es permet la detecció d'indicadors, patògens, i altra microbiota poc explorada cultivable i no cultivable en un mateix anàlisi. També, MALDI-TOF i *metabarcoding* permeten la observació de canvis en l'estructura de les comunitats bacterianes a temperatures elevades, o la reducció de Proteobacteria al llarg dels tractaments i la presència de gèneres resistents a l'aigua potable (Cyanobacteria o Firmicutes) que *a priori* no haurien de ser perjudicials per a la salut ni afectar a les propietats de l'aigua, des de que la informació dels indicadors ens ha aportat una seguretat en quant al consum. De fet, la legislació vigent també contempla l'anàlisi de microcistines (toxines de Cyanobacteria) a l'aigua tractada de la sortida de l'ETAP. Així doncs aquesta microbiota podria tenir un paper important en mantenir la bioestabilitat a l'aigua i per tant, conèixer-la mitjançant una aproximació holística és important per a una millor gestió integral de la qualitat de l'aigua davant d'un escenari de canvi climàtic i ús integral dels diferents recursos hídrics (OneWater) per mitigar els riscos que podrien comprometre la qualitat de l'aigua.

Perspectives futures en quant a la utilitat d'aquestes eines

La gran versatilitat d'aquestes eines pot derivar en nous estudis en que podrien resultar de gran utilitat per a la traçabilitat de les poblacions bacterianes per les empreses implicades en la gestió de l'aigua i en particular per AGBAR podrien aplicar-se en els següents àmbits:

- Anàlisi de les comunitats bacterianes a la xarxa de distribució de Barcelona que rep diàriament una composició d'aigua potable d'origens diferents (aquífers Besòs i Llobregat, superficials Llobregat i Ter, i dessalinitzadora).
- Avaluació de l'impacte en les comunitats bacterianes a les etapes de decantació i filtració per sorra per l'imminent canvi d'operativa planta a l'ETAP de SANT Joan Despí mitjançant la dosificació de permanganat potàssic a l'entrada de la decantació i reducció de dosificació de diòxid de clor que es preveu tingui un impacte positiu en la qualitat de l'aigua.
- Estudi de les comunitats bacterianes i les seves funcions metabòliques als pous de l'aqüífer del Llobregat amb recàrrega d'aigua filtrada pel filtre de sorra (Projecte Queen)
- Estudi de la viabilitat de les comunitats bacterianes per tècniques de metagenòmica mitjançant l'ús de marcadors de viabilitat com el propidi monoazida (PMA) que s'uneix i inactiva l'ADN lliure i cèl·lules tenen les parets danyades impeding la seva ampliació i seqüenciació.

- Creació d'una llibreria MALDI-TOF MS amb espectres de Cyanobacteria.
- Estudis de metabolòmica de les principals poblacions dominats a l'aigua potable com Cyanobacteria (ex. per confirmar que no són productores de toxines) i *Bacillus* (ex. per determinar si poden presentar factors de patogenicitat).
- Estudi dels biofilms a les etapes de tractament i a la xarxa de distribució
- Determinar la concentració de taxons d'interès, com els majoritaris segons *metabarcoding* a l'aigua potable, quantificant per PCR quantitativa per tal de disposar d'indicadors de procés quantificables.
- Generar dades de *metabarcoding* i integrar-les per desenvolupar models basats en intel·ligència artificial (IA) , emprant tècniques d'aprenentatge automàtic que interpretin les dades de seqüenciació, per exemple identificant un grup de marcadors genètics que expliquin la variació de les comunitats en el microbioma de l'aigua potable, i que en combinació amb dades fisicoquímiques dels processos, permeti desenvolupar models predictius de distribució espacial i la dinàmica temporal de les comunitats bacterianes d'aigua potable per facilitar vigilància de la qualitat de l'aigua.

CONCLUSIONS

Les conclusions generals dels diferents estudis realitzats a la present tesi doctoral són:

- S'ha optimitzat una metodologia per tal de crear una **llibreria de perfils proteics d'espectrometria de masses MALDI-TOF de bacteris relacionats amb l'aigua** que complementa la base de dades (llibreria) original, donada la seva limitació d'espectres de soques d'origen ambiental. L'ús conjunt ha millorat substancialment els resultats d'identificació.
- La tecnologia **MALDI-TOF MS** ha permès la **identificació de soques bacterianes aïllades de mostres d'aigües** de diferents orígens i destaca per la seva notable especificitat, rapidesa i facilitat d'ús que permet fer-ne ús en rutina als laboratoris de microbiologia d'aigües i empreses de subministrament d'aigua. Cal tenir en compte però que per a fer estudis de biodiversitat amb un nombre elevat de soques, el procediment i el resultat s'alenteix.
- S'han ajustat els processos de **concentració de biomassa** per les diferents etapes de tractament de l'ETAP per l'estudi de la seva microbiota amb l'ús de filtració per membrana i per ultrafiltració DEUF de grans volums de mostra. El volum i la metodologia de concentració de les mostres d'aigua és essencial per a una correcta representativitat de la matriu real.
- El **mètode DEUF** ha permès **concentrar grans volums d'aigua** de diferents matrius de l'ETAP mostrant una millor sensibilitat en la detecció d'indicadors de contaminació fecal, i conseqüentment permetent calcular una reducció més gran al llarg de les etapes. Tanmateix ha permès obtenir la biomassa adequada (també en les mostres clorades) requerida per la seqüenciació.
- S'ha definit un procés d'**extracció d'ADN** per a totes les matrius de l'ETAP i que permet ser aplicat en mostres d'aigua clorada i amb poca quantitat de biomassa per estudis de seqüenciació massiva. L'**extracció d'ADN** ha de permetre obtenir una quantitat i qualitat suficient dels àcids nucleics de la microbiota present a les mostres.
- L'**aigua potable** ha complert la normativa d'aigües de consum pel que fa als indicadors de contaminació fecal i recompte d'heteròtrofs. Malgrat això les poblacions d'heteròtrofs han mostrat canvis a nivell taxonòmic no detectats amb els paràmetres microbiològics de la normativa actual.
- La **caracterització per MALDI-TOF MS** de les poblacions bacterianes **cultivables** de l'ETAP de Sant Joan Despí al llarg del procés de potabilització d'aigua han mostrat el següent patró:

Proteobacteria i Bacteroidota dominant totes les etapes amb excepció de l'aigua potable on s'ha observat un domini de Firmicutes, concretament de *Bacillus*, que ha resistit tots els tractaments.

- La **caracterització de la totalitat** de les **poblacions bacterianes** de l'ETAP per **metabarcoding** (seqüenciació massiva) de la regió variable del fragment V4 del gen 16S ARNr ha permès observar: Proteobacteria, Bacteroidota i Bdellovibionota dominant totes les etapes amb excepció de l'aigua potable on s'ha observat Cyanobacteria en un clar domini del sistema, concretament d'Obscuribacteraceae.
- La utilització de les dues tècniques (**MALDI-TOF MS i metabarcoding**) ha permès una caracterització en profunditat de la microbiota a l'ETAP, ja que ambdues tècniques s'han mostrat **complementàries**.
- Totes les etapes de tractament tenen una **gran diversitat** observada tant per MALDI-TOF MS com per metabarcoding. De manera global s'ha observat una **disminució de la biodiversitat** als processos de **desinfecció** com la dioxiclорació de la decantació, l'ozonització i la cloració final a l'aigua potable i un **augment de diversitat** en els **filtres** de sorra i de carbó.
- S'ha observat una **influència de la temperatura** en quant a l'abundància de bacteris heteròtrofs a les mostres de riu, ozonització i aigua potable que va augmentar significativament, així com en augment de la diversitat i canvis en la seva composició, especialment a la decantació i filtració per carbó.
- Els **processos de tractament han influït** en la composició de la microbiota de l'aigua final, sobretot els processos de filtració i la cloració induint una forta pressió selectiva a favor dels grups resistents.
- Per **metabarcoding** s'han detectat **seqüències de gèneres patògens bacterians** de transmissió hídrica al llarg dels tractaments no detectables per MALDI-TOF MS ni pels mètodes regulats per la legislació. Aquestes han estat eliminades a l'aigua potable, amb excepció de baixes proporcions d'*Aeromonas* i *Pseudomonas*, demostrant l'eficàcia dels tractaments de potabilització.
- El coneixement de les comunitats bacterianes mitjançant l'**ús conjunt de diverses aproximacions addicionals als paràmetres legiscats**, permet disposar d'una imatge global, conèixer possibles canvis deguts, entre altres, a la temperatura i a l'efecte dels tractaments, i en conseqüència, permet avançar cap a una millor gestió integral dels sistemes d'aigua

que es preveu de gran utilitat davant d'un escenari de canvi climàtic i ús integral dels diferents recursos hídrics (OneWater) que podrien comprometre la qualitat de l'aigua.

Afectacions de la tesi per la pandèmia del SARS-CoV-2

A l'iniciar la pandèmia de la COVID-19, la comunitat científica reporta l'existència del virus en femta a elevades concentracions que aleshores, i tot i que en principi és un virus que no correspondria que fos infectiu per transmissió hídrica, encara no s'havia corroborat, i per tant no se sabia si podia tractar-se també d'un virus patògen de transmissió fecal-oral. El riu Llobregat a la vall baixa, com s'ha esmentat a la introducció, rep una aportació important d'aigua residual tractada. Per tal d'assegurar la innocuïtat de l'aigua potable, va ser necessari comprovar quina afectació podia tenir el virus del SARS-CoV-2 a l'ETAP mitjançant la seva detecció tant a la captació com a l'aigua potable. Pel que s'inicia ràpidament la posada a punt de tècniques de detecció del virus per RT-qPCR en col·laboració amb el grup de virus entèrics de la Universitat de Barcelona, així com una sèrie de mesures de protecció i avaluació de riscos biològics mitjançant equips de protecció individual o *EPIs* (tot i la limitació de disponibilitat que es tenia aleshores) pels treballadors de les plantes depuradores així com pel personal de recollida de mostres i tècniques d'anàlisi. Un cop es verifica que el virus no està present a les aigües de consum, i també mitjançant l'anàlisi de colífags somàtics, es posa en marxa a l'abril el projecte *REVEAL* (Suez) en col·laboració amb diferents institucions (acadèmia i empreses) autonòmiques, estatals i internacionals per tal d'avaluar l'impacte de la COVID-19 a les aigües residuals per assegurar la seva eliminació per l'alliberació segura al medi ambient, com es fa habitualment, així com el reaprofitament de fangs de les depuradores per finalitats agrícoles, construcció, etc. Durant aquest període, fruit de la col·laboració universitat-empresa mitjançant el doctorat industrial, vaig aturar la tesi doctoral temporalment durant 1,5 anys per tal de poder donar resposta a la situació pandèmica que requeria dels esforços de tots i que, amb els coneixements adquirits al llarg de la tesi, em va permetre participar activament en aquest projecte, que arrel que s'ha demostrat la correlació entre la concentració del virus en aigües residuals i la prevalença de la malaltia, actualment ha derivat en una vigilància epidemiològica periòdica mitjançant la detecció de la circulació del virus a les aigües residuals de les entrades de les depuradores principals de l'àmbit metropolità (Prat, Besòs, Montcada, Gavà i Sant Feliu), així com alguns col·lectors residuals municipals amb notificació de dades al departament de Salut i plataforma de seguiment de l'evolució del virus. Els resultats d'aquest projecte a l'inici de la pandèmia es troben publicats a l'apartat "altres publicacions" a la tesi doctoral.

BIBLIOGRAFIA

- ACA, 2019. SEQUERA Les sequeres a Catalunya Principals episodis [WWW Document]. URL https://sequera.gencat.cat/web/.content/02_LA_SEQUERA/02_QUE_LA_SEQUERA/Doc_LesSequeresCatalunya.pdf (accessed 5.8.22).
- Acharya, K., Halla, F.F., Massawa, S.M., Mgana, S.M., Komar, T., Davenport, R.J., Werner, D., 2020. Chlorination effects on DNA based characterization of water microbiomes and implications for the interpretation of data from disinfected systems. *J. Environ. Manage.* 276, 111319. <https://doi.org/10.1016/j.jenvman.2020.111319>
- Aigües de Barcelona, 2020. Memòria d'explotació 2020 [WWW Document]. URL https://www.aiguesdebarcelona.cat/documents/42802/619803/MEMORIA_EXPLLOTACIO_2020-11.pdf/b23e1738-fe54-5a00-c077-5a742bd5b0e3?t=1623339956221
- Aizenberg-Gershtein, Y., Vaizel-Ohayon, D., Halpern, M., 2012. Structure of bacterial communities in diverse freshwater habitats. *Can. J. Microbiol.* 58, 326–335. https://doi.org/10.1139/W11-138/SUPPL_FILE/W11-138SUPPLB.DOC
- Amann, R.L., Ludwig, W., Schleifer, K.-H., 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143. <https://doi.org/10.1128/MR.59.1.143-169.1995>
- Angelakis, E., Million, M., Henry, M., Raoult, D., 2011. Rapid and Accurate Bacterial Identification in Probiotics and Yoghurts by MALDI-TOF Mass Spectrometry. *J. Food Sci.* 76, M568–M572. <https://doi.org/10.1111/J.1750-3841.2011.02369.X>
- Angly, F.E., Dennis, P.G., Skarshewski, A., Vanwonterghem, I., Hugenholtz, P., Tyson, G.W., 2014. CopyRighter: A rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* 2, 1–13. <https://doi.org/10.1186/2049-2618-2-11/FIGURES/6>
- Anhalt, J.P., Fenselau, C., 1975. Identification of Bacteria using Mass Spectrometry. *Anal. Chem.* 47, 219–225. <https://doi.org/10.1021/AC60352A007>
- Anonymous, 2020. Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption [WWW Document]. URL <https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32020L2184> (accessed 2.9.22).
- Armengol, J., Dolz, J., 2009. L'abastament d'aigua a Catalunya i la seva garantia. "Nota d'economia: revista d'economia catalana i de sector públic." *General. Catalunya. Dir. Gen. d'Anàlisi i Política Econòmica* 127–129.
- Asquith, E., Evans, C., Dunstan, R.H., Geary, P., Cole, B., 2018. Distribution, abundance and activity of geosmin and 2-methylisoborneol-producing *Streptomyces* in drinking water reservoirs. *Water Res.* 145, 30–38. <https://doi.org/10.1016/J.WATRES.2018.08.014>
- ATL, n.d. ETAP del Ter | Ens d'Abastament d'Aigua Ter-Llobregat [WWW Document]. n.d. URL https://www.atl.cat/es/etap-del-ter_3170 (accessed 5.17.22a).
- ATL, n.d. ETAP del Llobregat | Ens d'Abastament d'Aigua Ter-Llobregat [WWW Document]. n.d. URL https://www.atl.cat/es/etap-del-llobregat_2974 (accessed 5.17.22b).
- ATL, n.d. Desalinizadora del Llobregat | Ens d'Abastament d'Aigua Ter-Llobregat [WWW Document]. URL https://www.atl.cat/es/desalinizadora-del-llobregat_2458 (accessed 5.13.22c).

- Atnafu, B., Desta, A., Assefa, F., 2021. Microbial Community Structure and Diversity in Drinking Water Supply, Distribution Systems as well as Household Point of Use Sites in Addis Ababa City, Ethiopia. *Microb. Ecol.* <https://doi.org/10.1007/S00248-021-01819-3>
- Baker, G.C., Smith, J.J., Cowan, D.A., 2003. Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* 55, 541–555. <https://doi.org/10.1016/j.mimet.2003.08.009>
- Barreiro, J.R., Ferreira, C.R., Sanvido, G.B., Kostrzewa, M., Maier, T., Wegemann, B., Böttcher, V., Eberlin, M.N., dos Santos, M. V., 2010. Short communication: Identification of subclinical cow mastitis pathogens in milk by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J. Dairy Sci.* 93, 5661–5667. <https://doi.org/10.3168/JDS.2010-3614>
- Bartram, J., Cotruvo, J., Exner, M., Fricker, C., Glasmacher, A., 2003. Heterotrophic Plate Counts and Drinking-water Safety The Significance of HPCs for Water Quality and Human Health. IWA Publ. <https://doi.org/10.2166/9781780405940>
- Bastin, B., Bird, P., Joseph Benzinger, M., Crowley, E., Agin, J., Goins, D., Sohler, D., Timke, M., Shi, G., Kostrzewa, M., 2019. Confirmation and Identification of Salmonella spp., Cronobacter spp., and Other Gram-Negative Organisms by the Bruker MALDI Biotyper Method: Collaborative Study Method Extension to Include Campylobacter Species, Revised First Action 2017.09. *J. AOAC Int.* 102, 1595–1616. <https://doi.org/10.5740/JAOACINT.18-0329>
- Berry, D., Xi, C., Raskin, L., 2006. Microbial ecology of drinking water distribution systems. *Curr. Opin. Biotechnol.* 17, 297–302. <https://doi.org/10.1016/j.copbio.2006.05.007>
- Blanch, A.R., Galofré, B., Lucena, F., Terradillos, A., Vilanova, X., Ribas, F., 2007. Characterization of bacterial coliform occurrences in different zones of a drinking water distribution system. *J. Appl. Microbiol.* 102, 711–721. <https://doi.org/10.1111/J.1365-2672.2006.03141.X>
- Blázquez-Pallí, N., Rosell, M., Varias, J., Bosch, M., Soler, A., Vicent, T., Marco-Urrea, E., 2019. Integrative isotopic and molecular approach for the diagnosis and implementation of an efficient in-situ enhanced biological reductive dechlorination of chlorinated ethenes. *Water Res.* 167. <https://doi.org/10.1016/j.watres.2019.115106>
- Bouwer, E.J., Crowe, P.B., 1988. Biological Processes in Drinking Water Treatment. *J. Am. Water Works Assoc.* 80, 82–93. <https://doi.org/10.1002/J.1551-8833.1988.TB03103.X>
- Brandt, J., Albertsen, M., 2018. Investigation of detection limits and the influence of DNA extraction and primer choice on the observed microbial communities in drinking water samples using 16S rRNA gene amplicon sequencing. *Front. Microbiol.* 9, 1–11. <https://doi.org/10.3389/fmicb.2018.02140>
- Bridle, H., 2014. Waterborne pathogens : detection methods and applications.
- Bukin, Y.S., Galachyants, Y.P., Morozov, I. V., Bukin, S. V., Zakharenko, A.S., Zemskaia, T.I., 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Sci. Data* 2019 6:1–14. <https://doi.org/10.1038/sdata.2019.7>
- Burak, S., Margat, J., 2016. Water Management in the Mediterranean Region: Concepts and Policies. *Water Resour. Manag.* 30, 5779–5797. <https://doi.org/10.1007/S11269-016-1389-4/TABLES/5>
- Calero Preciado, C., Boxall, J., Soria-Carrasco, V., Martínez, S., Doutereiro, I., 2021. Implications of Climate Change: How Does Increased Water Temperature Influence Biofilm and Water Quality of Chlorinated Drinking Water Distribution Systems? *Front. Microbiol.* 12. <https://doi.org/10.3389/fmicb.2021.658927>

- Calgua, B., Mengewein, A., Grunert, A., Bofill-Mas, S., Clemente-Casares, P., Hundesa, A., Wyn-Jones, A.P., López-Pila, J.M., Girones, R., 2008. Development and application of a one-step low cost procedure to concentrate viruses from seawater samples. *J. Virol. Methods* 153, 79–83. <https://doi.org/10.1016/J.JVIROMET.2008.08.003>
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J.A., Smith, G., Knight, R., 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* 6, 1621–1624. <https://doi.org/10.1038/ismej.2012.8>
- CETAQUA, 2018. D35.1 Evaluation of the results and impacts on ESS of a flexible ASR system in Barcelona (ES) demo site. Guidelines and recommendations for transfer this innovative solution. [WWW Document]. URL <https://dessin-project.eu/?wpdmp=pro=d35-1-evaluation-of-the-results-and-impacts-on-ess-of-a-flexible-asr-system-in-barcelona-es-demo-site-guidelines-and-recommendations-for-transfer-this-innovative-solution>
- Chao, Y., Ma, L., Yang, Y., Ju, F., Zhang, X.X., Wu, W.M., Zhang, T., 2013. Metagenomic analysis reveals significant changes of microbial compositions and protective functions during drinking water treatment. *Sci. Rep.* 3, 1–9. <https://doi.org/10.1038/srep03550>
- Chao, Y., Mao, Y., Wang, Z., Zhang, T., 2015. Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. *Sci. Reports* 2015 51 5, 1–13. <https://doi.org/10.1038/srep10044>
- Christner, M., Trusch, M., Rohde, H., Kwiatkowski, M., Schlüter, H., Wolters, M., Aepfelbacher, M., Hentschke, M., 2014. Rapid MALDI-TOF Mass Spectrometry Strain Typing during a Large Outbreak of Shiga-Toxigenic *Escherichia coli*. *PLoS One* 9, e101924. <https://doi.org/10.1371/JOURNAL.PONE.0101924>
- Collado, L., Kasimir, G., Perez, U., Bosch, A., Pinto, R., Saucedo, G., Huguet, J.M., Jose, M., 2010. Occurrence and diversity of *Arcobacter* spp . along the Llobregat River catchment , at sewage effluents and in a drinking water treatment plant. *Water Res.* 44, 3696–3702. <https://doi.org/10.1016/j.watres.2010.04.002>
- Cuevas-Ferrando, E., Randazzo, W., Pérez-Cataluña, A., Sánchez, G., 2020. HEV Occurrence in Waste and Drinking Water Treatment Plants. *Front. Microbiol.* 10, 2937. <https://doi.org/10.3389/FMICB.2019.02937/BIBTEX>
- D’Amore, R., Ijaz, U.Z., Schirmer, M., Kenny, J.G., Gregory, R., Darby, A.C., 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, 1–20. <https://doi.org/10.1186/S12864-015-2194-9/TABLES/9>
- De Carolis, E., Vella, A., Vaccaro, L., Torelli, R., Posteraro, P., Ricciardi, W., Sanguinetti, M., Posteraro, B., 2014. Development and validation of an in-house database for matrix-assisted laser desorption ionization-time of flight mass spectrometry-based yeast identification using a fast protein extraction procedure. *J. Clin. Microbiol.* 52, 1453–1458. <https://doi.org/10.1128/JCM.03355-13>
- Degerman, R., Dinasquet, J., Riemann, L., De Luna, S.S., Andersson, A., 2013. Effect of resource availability on bacterial community responses to increased temperature. *Aquat. Microb. Ecol.* 68, 131–142. <https://doi.org/10.3354/AME01609>
- Deulofeu, M., García-Cuesta, E., Peña-Méndez, E.M., Conde, J.E., Jiménez-Romero, O., Verdú, E., Serrando, M.T., Salvadó, V., Boadas-Vaello, P., 2021. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining MALDI-TOF MS and Artificial Intelligence.

- Front. Med. 8, 398. <https://doi.org/10.3389/FMED.2021.661358/BIBTEX>
- Douterelo, I., Husband, S., Loza, V., Boxall, J., 2016. Dynamics of biofilm regrowth in drinking water distribution systems. *Appl. Environ. Microbiol.* 82, 4155–4168. <https://doi.org/10.1128/AEM.00109-16>
- Egli, T., 2010. How to live at very low substrate concentration. *Water Res.* 44, 4826–4837. <https://doi.org/10.1016/J.WATRES.2010.07.023>
- Eichler, S., Christen, R., Höltje, C., Westphal, P., Bötzel, J., Brettar, I., Mehling, A., Höfle, M.G., 2006. Composition and dynamics of bacterial communities of a drinking water supply system as assessed by RNA- and DNA-based 16S rRNA gene fingerprinting. *Appl. Environ. Microbiol.* 72, 1858–1872. <https://doi.org/10.1128/AEM.72.3.1858-1872.2006>
- Eljarrat, E., Barceló, D., 2012. Occurrence and Behavior of Brominated Flame Retardants in the Llobregat River Basin. *Handb. Environ. Chem.* 21, 135–150. https://doi.org/10.1007/698_2011_139
- Emami, K., Askari, V., Ullrich, M., Mohinudeen, K., Anil, A.C., Khandeparker, L., Burgess, J.G., Mesbahi, E., 2012. Characterization of bacteria in ballast water using MALDI-TOF mass spectrometry. *PLoS One* 7. <https://doi.org/10.1371/JOURNAL.PONE.0038515>
- Espejo, R.T., Plaza, N., 2018. Multiple Ribosomal RNA operons in bacteria; Their concerted evolution and potential consequences on the rate of evolution of their 16S rRNA. *Front. Microbiol.* 9, 1–6. <https://doi.org/10.3389/fmicb.2018.01232>
- FAO and UN Water, 2021. Progress on Level of Water Stress. Global status and acceleration needs for SDG Indicator 6.4.2, 2021. <https://doi.org/10.4060/cb6241en>
- Favere, J., Barbosa, R.G., Sleutels, T., Verstraete, W., De Gussem, B., Boon, N., 2021. Safeguarding the microbial water quality from source to tap. *npj Clean Water* 4, 1–6. <https://doi.org/10.1038/s41545-021-00118-1>
- Ferguson, G., McIntosh, J.C., Warr, O., Sherwood Lollar, B., Ballentine, C.J., Famiglietti, J.S., Kim, J.H., Michalski, J.R., Mustard, J.F., Tarnas, J., McDonnell, J.J., 2021. Crustal Groundwater Volumes Greater Than Previously Thought. *Geophys. Res. Lett.* 48, e2021GL093549. <https://doi.org/10.1029/2021GL093549>
- Fernández-Turiel, J.L., Gimeno, D., Rodríguez, J.J., Carnicero, M., Valero, F., 2003. Spatial and seasonal variations of water quality in a Mediterranean catchment: The Llobregat River (NE Spain). *Environ. Geochem. Health* 25, 453–474. <https://doi.org/10.1023/B:EGAH.0000004566.75757.98>
- Ferrari, S., Frosth, S., Svensson, L., Fernström, L.L., Skarin, H., Hansson, I., 2019. Detection of *Campylobacter* spp. in water by dead-end ultrafiltration and application at farm level. *J. Appl. Microbiol.* 127, 1270–1279. <https://doi.org/10.1111/JAM.14379>
- Fish, K.E., Boxall, J.B., 2018. Biofilm microbiome (re)growth dynamics in drinking water distribution systems are impacted by chlorine concentration. *Front. Microbiol.* 9, 2519. <https://doi.org/10.3389/FMICB.2018.02519/BIBTEX>
- Florio, W., Baldeschi, L., Rizzato, C., Tavanti, A., Ghelardi, E., Lupetti, A., 2020. Detection of Antibiotic-Resistance by MALDI-TOF Mass Spectrometry: An Expanding Area. *Front. Cell. Infect. Microbiol.* 10, 572909. <https://doi.org/10.3389/FCIMB.2020.572909>
- Francy, D.S., Stelzer, E.A., Brady, A.M.G., Huitger, C., Bushon, R.N., Ip, H.S., Ware, M.W., Villegas, E.N., Gallardo, V., Lindquist, H.D.A., 2013. Comparison of Filters for Concentrating

- Microbial Indicators and Pathogens in Lake Water Samples. *Appl. Environ. Microbiol.* 79, 1342. <https://doi.org/10.1128/AEM.03117-12>
- Fung, D.Y.C., 2002. Rapid Methods and Automation in Microbiology. *Compr. Rev. Food Sci. Food Saf.* 1, 3–22. <https://doi.org/10.1111/J.1541-4337.2002.TB00003.X>
- Giorgi, F., Lionello, P., 2007. Climate change projections for the Mediterranean region. *Glob Planet Chang.* 63, 90–104. <https://doi.org/10.1016/j.gloplacha.2007.09.005>
- Gleick, P.H., 2018. Transitions to freshwater sustainability. *Proc. Natl. Acad. Sci. U. S. A.* 115, 8863–8871. <https://doi.org/10.1073/PNAS.1808893115/-/DCSUPPLEMENTAL>
- Gorostiza Langa, S., 2014. Potash Extraction and Historical Environmental Conflict in the Bages Region (Spain). *Investig. Geográficas* 5. <https://doi.org/10.14198/ingeo2014.61.01>
- Gunnarsdottir, M.J., Gardarsson, S.M., Figueras, M.J., Puigdomènech, C., Juárez, R., Saucedo, G., Arnedo, M.J., Santos, R., Monteiro, S., Avery, L., Pagaling, E., Allan, R., Abel, C., Eglitis, J., Hamsch, B., Hügler, M., Rajkovic, A., Smigic, N., Udovicki, B., Albrechtsen, H.J., López-Avilés, A., Hunter, P., 2020. Water safety plan enhancements with improved drinking water quality detection techniques. *Sci. Total Environ.* 698, 134185. <https://doi.org/10.1016/j.scitotenv.2019.134185>
- Heather, J.M., Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics* 107, 1–8. <https://doi.org/10.1016/j.ygeno.2015.11.003>
- Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. B Biol. Sci.* 270, 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hernandez, M.M., Banu, R., Shrestha, P., Patel, A., Chen, F., Cao, L., Fabre, S., Tan, J., Lopez, H., Chiu, N., Shifrin, B., Zapolskaya, I., Flores, V., Lee, P.Y., Castañeda, S., Ramírez, J.D., Jhang, J., Osorio, G., Gitman, M.R., Nowak, M.D., Reich, D.L., Cordon-Cardo, C., Sordillo, E.M., Paniz-Mondolfi, A.E., 2021. RT-PCR/MALDI-TOF mass spectrometry-based detection of SARS-CoV-2 in saliva specimens. *J. Med. Virol.* 93, 5481–5486. <https://doi.org/10.1002/JMV.27069>
- Hill, V.R., Kahler, A.M., Jothikumar, N., Johnson, T.B., Hahn, D., Cromeans, T.L., 2007. Multistate evaluation of an ultrafiltration-based procedure for simultaneous recovery of enteric microbes in 100-Liter tap water samples. *Appl. Environ. Microbiol.* 73, 4218–4225. <https://doi.org/10.1128/AEM.02713-06>
- Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenka, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLoS One* 12, 1–17. <https://doi.org/10.1371/journal.pone.0170199>
- Holland, R.D., Wilkes, J.G., Ralli, F., Sutherland, J.B., Persons, C.C., Voorhees, K.J., Lay, J. O', 1996. Rapid Identification of Intact Whole Bacteria Based on Spectral Patterns using Matrix-assisted Laser Desorption/Ionization with Time-of-flight Mass Spectrometry. *RAPID Commun. MASS Spectrom.* 10, 1227–1232. [https://doi.org/10.1002/\(SICI\)1097-0231\(19960731\)10:10](https://doi.org/10.1002/(SICI)1097-0231(19960731)10:10)
- Hou, L., Zhou, Q., Wu, Q., Gu, Q., Sun, M., Zhang, J., 2018. Spatiotemporal changes in bacterial community and microbial activity in a full-scale drinking water treatment plant. *Sci. Total Environ.* 625, 449–459. <https://doi.org/10.1016/j.scitotenv.2017.12.301>

- Izaguirre, G., Hwang, C.J., Krasner, S.W., McGuire, M.J., 1982. Geosmin and 2-Methylisoborneol from Cyanobacteria in Three Water Supply Systems. *Appl. Environ. Microbiol.* 43, 708. <https://doi.org/10.1128/AEM.43.3.708-714.1982>
- Jalili, F., Trigui, H., Guerra Maldonado, J.F., Dorner, S., Zamyadi, A., Shapiro, B.J., Terrat, Y., Fortin, N., Sauvé, S., Prévost, M., 2021. Can Cyanobacterial Diversity in the Source Predict the Diversity in Sludge and the Risk of Toxin Release in a Drinking Water Treatment Plant? *Toxins (Basel)*. 13. <https://doi.org/10.3390/toxins13010025>
- Janda, J.M., Abbott, S.L., 2007. 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: Pluses, perils, and pitfalls. *J. Clin. Microbiol.* 45, 2761–2764. <https://doi.org/10.1128/JCM.01228-07>
- Kim, E., Cho, Y., Lee, Y., Han, S.-K., Kim, C.-G., Choo, D.-W., Kim, Y.-R., Kim, H.-Y., 2016. A proteomic approach for rapid identification of *Weissella* species isolated from Korean fermented foods on MALDI-TOF MS supplemented with an in-house database. *Int J Food Microbiol.* 243, 9–15. <https://doi.org/10.1016/j.ijfoodmicro.2016.11.027>
- Kopcakova, A., Stramova, Z., Kvasnova, S., Godany, A., Perhacova, Z., Pristas, P., 2014. Need for database extension for reliable identification of bacteria from extreme environments using MALDI TOF mass spectrometry. *Chem. Pap.* 68, 1435–1442. <https://doi.org/10.2478/s11696-014-0612-0>
- Kraková, L., Soltys, K., Otlewska, A., Pietrzak, K., Purkrtová, S., Savická, D., Purskářová, A., Bucková, M., Szemes, T., Budis, J., Demnerová, K., Gutarowska, B., Pangallo, D., 2017. Comparison of methods for identification of microbial communities in book collections: Culture-dependent (sequencing and MALDI-TOF MS) and culture-independent (Illumina MiSeq). *Int. Biodeterior. Biodegradation* 131, 51–59. <https://doi.org/10.1016/j.ibiod.2017.02.015>
- Langille, M.G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepille, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G., Huttenhower, C., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* 31, 814–821. <https://doi.org/10.1038/NBT.2676>
- Lautenschlager, K., Hwang, C., Liu, W.T., Boon, N., Köster, O., Vrouwenvelder, H., Egli, T., Hammes, F., 2013. A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Res.* 47, 3015–3025. <https://doi.org/10.1016/j.watres.2013.03.002>
- Lewis, G.D., Metcalf, T.G., 1988. Polyethylene glycol precipitation for recovery of pathogenic viruses, including hepatitis A virus and human rotavirus, from oyster, water, and sediment samples. *Appl. Environ. Microbiol.* 54, 1983–1988. <https://doi.org/10.1128/AEM.54.8.1983-1988.1988>
- Li, C., Ling, F., Zhang, M., Liu, W.T., Li, Y., Liu, W., 2017. Characterization of bacterial community dynamics in a full-scale drinking water treatment plant. *J. Environ. Sci. (China)* 51, 21–30. <https://doi.org/10.1016/J.JES.2016.05.042>
- Li, D., Li, Z., Yu, J., Cao, N., Liu, R., Yang, M., 2010. Characterization of bacterial community structure in a drinking water distribution system during an occurrence of red water. *Appl. Environ. Microbiol.* 76, 7171–7180. <https://doi.org/10.1128/AEM.00832-10>
- Li, Q., Yu, S., Li, L., Liu, G., Gu, Z., Liu, M., Liu, Z., Ye, Y., Xia, Q., Ren, L., 2017. Microbial communities shaped by treatment processes in a drinking water treatment plant and their contribution and threat to drinking water safety. *Front. Microbiol.* 8, 2465.

- <https://doi.org/10.3389/FMICB.2017.02465>
- Li, W., Zhang, J., Wang, F., Qian, L., Zhou, Y., Qi, W., Chen, J., 2018. Effect of disinfectant residual on the interaction between bacterial growth and assimilable organic carbon in a drinking water distribution system. *Chemosphere* 202, 586–597. <https://doi.org/10.1016/j.chemosphere.2018.03.056>
- Liu, P.-Y., Wu, W.-K., Chen, C.-C., Panyod, S., Sheen, L.-Y., Wu, M.-S., 2020. Evaluation of Compatibility of 16S rRNA V3V4 and V4 Amplicon Libraries for Clinical Microbiome Profiling. *bioRxiv* 2020.08.18.256818. <https://doi.org/10.1101/2020.08.18.256818>
- Liu, T., Kong, W., Chen, N., Zhu, J., Wang, J., He, X., Jin, Y., 2016. Bacterial characterization of Beijing drinking water by flow cytometry and MiSeq sequencing of the 16S rRNA gene. *Ecol. Evol.* 6, 923–934. <https://doi.org/10.1002/ece3.1955>
- Louca, S., Doebeli, M., Parfrey, L.W., 2018. Correcting for 16S rRNA gene copy numbers in microbiome surveys remains an unsolved problem. *Microbiome* 6, 1–12. <https://doi.org/10.1186/s40168-018-0420-9>
- Lucena, F., Jofre, J., 2014. Potential Use of Bacteriophages as Indicators of Water Quality and Wastewater Treatment Processes. *Bacteriophages Control Food- Waterborne Pathog.* 103–118. <https://doi.org/10.1128/9781555816629.CH6>
- Luo, L.W., Wu, Y.H., Yu, T., Wang, Y.H., Chen, G.Q., Tong, X., Bai, Y., Xu, C., Wang, H. Bin, Ikuno, N., Hu, H.Y., 2021. Evaluating method and potential risks of chlorine-resistant bacteria (CRB): A review. *Water Res.* 188, 116474. <https://doi.org/10.1016/j.watres.2020.116474>
- Marcé, R., Honey-Rosés, J., Manzano, A., Moragas, L., Catllar, B., Sabater, S., 2012. The Llobregat River Basin: A Paradigm of Impaired Rivers Under Climate Change Threats. *Handb. Environ. Chem.* 21, 1–26. https://doi.org/10.1007/698_2012_152
- Marizzoni, M., Gurry, T., Provasi, S., Greub, G., Lopizzo, N., Ribaldi, F., Festari, C., Mazzelli, M., Mombelli, E., Salvatore, M., Mirabelli, P., Franzese, M., Soricelli, A., Frisoni, G.B., Cattaneo, A., 2020. Comparison of Bioinformatics Pipelines and Operating Systems for the Analyses of 16S rRNA Gene Amplicon Sequences in Human Fecal Samples. *Front. Microbiol.* 11, 1262. <https://doi.org/10.3389/FMICB.2020.01262/BIBTEX>
- Martin-Alonso, J., 1994. Barcelona's water supply improvement: the brine collector of the Llobregat river. *Water Sci. Technol.* 30, 221–227. <https://doi.org/10.2166/WST.1994.0531>
- Mathieu, L., Bouteleux, C., Fass, S., Angel, E., Block, J.C., 2009. Reversible shift in the alpha, beta and gamma proteobacteria populations of drinking water biofilms during discontinuous chlorination. *Water Res.* 43, 3375–3386. <https://doi.org/10.1016/j.watres.2009.05.005>
- McDaniel, E.A., Wahl, S.A., Ishii, S., Pinto, A., Ziels, R., Nielsen, P.H., McMahan, K.D., Williams, R.B.H., 2021. Prospects for multi-omics in the microbial ecology of water engineering. *Water Res.* 205. <https://doi.org/10.1016/J.WATRES.2021.117608>
- MedECC, 2020. Climate and Environmental Change in the Mediterranean Basin – Current Situation and Risks for the Future. First Mediterranean Assessment Report [Cramer, W., Guiot, J., Marini, K. (eds.)] Union for the Mediterranean, Plan Bleu, UNEP/MAP, Marseille, France, Marseille, France. <https://doi.org/10.5281/zenodo.4768833>
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., Brouwer, A., 2020. Presence of SARS-Coronavirus-2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in the Netherlands. *Environ. Sci. Technol. Lett.* 7, 511–516. https://doi.org/10.1021/ACS.ESTLETT.0C00357/SUPPL_FILE/EZ0C00357_SI_001.PDF

- Moona, N., Wünsch, U.J., Bondelind, M., Bergstedt, O., Sapmaz, T., Pettersson, T.J.R., Murphy, K.R., 2019. Temperature-dependent mechanisms of DOM removal by biological activated carbon filters. *Environ. Sci. Water Res. Technol.* 5, 2232–2241. <https://doi.org/10.1039/c9ew00620f>
- Morgan, X.C., Huttenhower, C., 2012. Chapter 12: Human Microbiome Analysis. *PLoS Comput. Biol.* 8. <https://doi.org/10.1371/journal.pcbi.1002808>
- Owen, D.M., 2015. Changing the paradigm: Managing water throughout the cycle for total water solutions. *J. Am. Water Works Assoc.* 107, 54–59. <https://doi.org/10.5942/JAWWA.2015.107.0092>
- Pérez-Sancho, M., Cerdá, I., Fernández-Bravo, A., Domínguez, L., Figueras, M.J., Fernández-Garayzábal, J.F., Vela, A.I., 2018. Limited performance of MALDI-TOF for identification of fish aeromonas isolates at species level. *J. Fish Dis.* 41, 1485–1493. <https://doi.org/10.1111/JFD.12837>
- Pascual-Benito, M., Emiliano, P., Casas-Mangas, R., Dacal-Rodríguez, C., Gracenea, M., Araujo, R., Valero, F., García-Aljaro, C., Lucena, F., 2020. Assessment of dead-end ultrafiltration for the detection and quantification of microbial indicators and pathogens in the drinking water treatment processes. *Int. J. Hyg. Environ. Health* 230, 113628. <https://doi.org/10.1016/j.ijheh.2020.113628>
- Pei, H., Xu, H., Wang, J., Jin, Y., Xiao, H., Ma, C., Sun, J., Li, H., 2017. 16S rRNA Gene Amplicon Sequencing Reveals Significant Changes in Microbial Compositions during Cyanobacteria-Laden Drinking Water Sludge Storage. *Environ. Sci. Technol.* 51, 12774–12783. <https://doi.org/10.1021/acs.est.7b03085>
- Pérez, S., Rubiano, M.E., Ginebreda, A., Postigo, C., López-Serna, R., Blanco, J., Osorio, V., de Alda, M.L., Petrović, M., Pastor, J.J., Matia, L., Martín-Alonso, J., Munné, A., Jofre, J., Lucena, F., Agulló-Barceló, M., Prat, N., Barceló, D., 2012. Wastewater Reuse in the Llobregat: The Experience at the Prat de Llobregat Treatment Plant. *Handb. Environ. Chem.* 21, 327–346. https://doi.org/10.1007/698_2012_151
- Petrovic, M., Ginebreda, A., Rivers, L., Alda, M.L. De, 2010. Water supply for the Barcelona metropolitan area. *Géosciences* 10, 32–37.
- Pinto, A.J., Xi, C., Raskin, L., 2012. Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Env. Sci Technol* 46, 8851–9. <https://doi.org/10.1021/es302042t>
- Poitelon, J.B., Joyeux, M., Welté, B., Duguet, J.P., Prestel, E., Dubow, M.S., 2010. Variations of bacterial 16S rDNA phylotypes prior to and after chlorination for drinking water production from two surface water treatment plants. *J. Ind. Microbiol. Biotechnol.* 37, 117–128. <https://doi.org/10.1007/s10295-009-0653-5>
- Pollock, J., Glendinning, L., Wisedchanwet, T., Watson, M., 2018. The madness of microbiome: Attempting to find consensus “best practice” for 16S microbiome studies. *Appl. Environ. Microbiol.* 84. https://doi.org/10.1128/AEM.02627-17/E02627_17_PDF.PDF
- Prest, E.I., El-Chakhtoura, J., Hammes, F., Saikaly, P.E., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2014. Combining flow cytometry and 16S rRNA gene pyrosequencing: A promising approach for drinking water monitoring and characterization. *Water Res.* 63, 179–189. <https://doi.org/10.1016/j.watres.2014.06.020>
- Prest, E.I., Hammes, F., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., 2016. Biological stability

- of drinking water: Controlling factors, methods, and challenges. *Front. Microbiol.* 7, 45. <https://doi.org/10.3389/FMICB.2016.00045/BIBTEX>
- Pruden, A., Ashbolt, N., Miller, J., 2019. Overview of issues for water bacterial pathogens. *Water Sanit. 21st Century Heal. Microbiol. Asp. Excreta Wastewater Manag. (Global Water Pathog. Proj.* <https://doi.org/10.14321/waterpathogens.20>
- Pupo, G.M., Lan, R., Reeves, P.R., 2000. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proc. Natl. Acad. Sci. U. S. A.* 97, 10567–10572. https://doi.org/10.1073/PNAS.180094797/SUPPL_FILE/FIG4.PDF
- Rahi, P., Prakash, O., Shouche, Y.S., 2016. Matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry (MALDI-TOF MS) based microbial identifications: Challenges and scopes for microbial ecologists. *Front. Microbiol.* 7, 1–12. <https://doi.org/10.3389/fmicb.2016.01359>
- Randazzo, W., Cuevas-Ferrando, E., Sanjuán, R., Domingo-Calap, P., Sánchez, G., 2020. Metropolitan wastewater analysis for COVID-19 epidemiological surveillance. *Int. J. Hyg. Environ. Health* 230, 1438–4639. <https://doi.org/10.1016/J.IJHEH.2020.113621>
- Real Decreto 1620/2007, 2007. Real Decreto 1620/2007, de 7 de diciembre, por el que se establece el régimen jurídico de la reutilización de las aguas depuradas. [WWW Document]. URL <https://www.boe.es/buscar/doc.php?id=BOE-A-2007-21092> (accessed 5.17.22).
- Regan, J.M., Harrington, G.W., Baribeau, H., De Leon, R., Noguera, D.R., 2003. Diversity of nitrifying bacteria in full-scale chloraminated distribution systems. *Water Res.* 37, 197–205.
- Regan, J.M., Harrington, G.W., Noguera, D.R., 2002. Ammonia- and Nitrite-Oxidizing Bacterial Communities in a Pilot-Scale Chloraminated Drinking Water Distribution System. *Appl. Environ. Microbiol.* 68, 73. <https://doi.org/10.1128/AEM.68.1.73-81.2002>
- Regulation (EU) 2020/741, 2020. Regulation (EU) 2020/741 of the European Parliament and of the Council of 25 May 2020 on minimum requirements for water reuse (Text with EEA relevance) [WWW Document]. URL <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32020R0741> (accessed 5.17.22).
- Revetta, R.P., Pemberton, A., Lamendella, R., Iker, B., Santo Domingo, J.W., 2010. Identification of bacterial populations in drinking water using 16S rRNA-based sequence analyses. *Water Res.* 44, 1353–1360. <https://doi.org/10.1016/j.watres.2009.11.008>
- Rhodes, E.R., Hamilton, D.W., See, M.J., Wymer, L., 2011. Evaluation of hollow-fiber ultrafiltration primary concentration of pathogens and secondary concentration of viruses from water. *J. Virol. Methods* 176, 38–45. <https://doi.org/10.1016/j.jviromet.2011.05.031>
- Riffiani, R., Sulistinah, N., Sunarko, B., 2015. Comparison of three DNA isolation and purification methods of bacterial DNA. *KnE Life Sci.* 2, 491. <https://doi.org/10.18502/KLS.V2I1.199>
- Roeselers, G., Coolen, J., van der Wielen, P.W.J.J., Jaspers, M.C., Atsma, A., de Graaf, B., Schuren, F., 2015. Microbial biogeography of drinking water: Patterns in phylogenetic diversity across space and time. *Environ. Microbiol.* 17, 2505–2514. <https://doi.org/10.1111/1462-2920.12739/SUPPINFO>
- Romero, P.E., Calla-Quispe, E., Castillo-Vilcahuaman, C., Yokoo, M., Fuentes-Rivera, H.L., Ramirez, J.L., Ampuero, A., Ibáñez, A.J., Wong, P., 2021. From the Andes to the desert: 16S

- rRNA metabarcoding characterization of aquatic bacterial communities in the Rimac river, the main source of water for Lima, Peru. *PLoS One* 16, 1–18. <https://doi.org/10.1371/journal.pone.0250401>
- Rubiano, M.E., Agulló-Barceló, M., Casas-Mangas, R., Jofre, J., Lucena, F., 2012. Assessing the effects of tertiary treated wastewater reuse on a Mediterranean river (Llobregat, NE Spain), part III: Pathogens and indicators. *Environ. Sci. Pollut. Res.* 19, 1026–1032. <https://doi.org/10.1007/s11356-011-0562-9>
- Rusiñol, M., Martínez-Puchol, S., Forés, E., Itarte, M., Girones, R., Bofill-Mas, S., 2020. Concentration methods for the quantification of coronavirus and other potentially pandemic enveloped virus from wastewater. *Curr. Opin. Environ. Sci. Heal.* 17, 21. <https://doi.org/10.1016/J.COESH.2020.08.002>
- Sahlin, K., Sipos, B., James, P.L., Medvedev, P., 2021. Error correction enables use of Oxford Nanopore technology for reference-free transcriptome analysis. *Nat. Commun.* 2021 121 12, 1–13. <https://doi.org/10.1038/s41467-020-20340-8>
- Sala-Comorera, L., Blanch, A.R., Casanovas-Massana, A., Monleón-Getino, A., García-Aljaro, C., 2019. Traceability of different brands of bottled mineral water during shelf life, using PCR-DGGE and next generation sequencing techniques. *Food Microbiol.* 82, 1–10. <https://doi.org/10.1016/j.fm.2019.01.006>
- Sala-Comorera, L., Blanch, A.R., Vilaró, C., Galofré, B., García-Aljaro, C., 2016a. Pseudomonas-related populations associated with reverse osmosis in drinking water treatment. *J. Environ. Manage.* 182, 335–341. <https://doi.org/10.1016/J.JENVMAN.2016.07.089>
- Sala-Comorera, L., Vilaró, C., Galofré, B., Blanch, A.R., García-Aljaro, C., 2016b. Use of matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) mass spectrometry for bacterial monitoring in routine analysis at a drinking water treatment plant. *Int. J. Hyg. Environ. Health* 219, 577–584. <https://doi.org/10.1016/J.IJHEH.2016.01.001>
- Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12, 1–12. <https://doi.org/10.1186/s12915-014-0087-z>
- Santos, I., Martin, M., Carlton, D., Amorim, C., Castro, P., Hildenbrand, Z., Schug, K., 2017. MALDI-TOF MS for the Identification of Cultivable Organic-Degrading Bacteria in Contaminated Groundwater near Unconventional Natural Gas Extraction Sites. *Microorganisms* 5, 47. <https://doi.org/10.3390/microorganisms5030047>
- Santos, I.C., Hildenbrand, Z.L., Schug, K.A., 2016. Applications of MALDI-TOF MS in environmental microbiology. *Analyst* 141, 2827–2837. <https://doi.org/10.1039/c6an00131a>
- Sárvári, K.P., Soki, J., Iván, M., Miszti, C., Latkoczy, K., Melegh, S., Urbán, E., 2018. Maldi-TOF MS versus 16S rRNA sequencing: Minor discrepancy between tools in identification of bacteroides isolates. *Acta Microbiol. Immunol. Hung.* 65, 173–181. <https://doi.org/10.1556/030.64.2017.025>
- Scott, B.A., Pepper, I.L., 2010. Water distribution systems as living ecosystems: impact on taste and odor. *J. Environ. Sci. Health. A. Tox. Hazard. Subst. Environ. Eng.* 45, 890–900. <https://doi.org/10.1080/10934521003709115>
- Šedo, O., Sedláček, I., Zdráhal, Z., 2011. Sample preparation methods for MALDI-MS profiling of

- bacteria. *Mass Spectrom. Rev.* 30, 417–434. <https://doi.org/10.1002/MAS.20287>
- Seuylemezian, A., Aronson, H.S., Tan, J., Lin, M., Schubert, W., Vaishampayan, P., 2018. Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. *Front. Microbiol.* 9, 1–8. <https://doi.org/10.3389/fmicb.2018.00780>
- Shiklomanov, I.A. and Rodda, J.C., 2004. World water resources at the beginning of the twenty-first century, UNESCO, International hydrology series.
- Singhal, N., Kumar, M., Kanaujia, P.K., Viridi, J.S., 2015. MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front. Microbiol.* 6. <https://doi.org/10.3389/FMICB.2015.00791>
- Stockwell, C., Geiges, A., Ramalope, D., Gidden, M., Hare, B., de Villafranca Casa, M.J., Moisiu, M., Hans, F., Fekete, H., 2021. Glasgow’s 2030 credibility gap: net zero’s lip service to climate action. *Clim. Action Tracker* 1–30.
- Sun, H., Shi, B., Lytle, D.A., Bai, Y., Wang, D., 2014. Formation and release behavior of iron corrosion products under the influence of bacterial communities in a simulated water distribution system. *Environ. Sci. Process. Impacts* 16, 576–585. <https://doi.org/10.1039/C3EM00544E>
- Tan, G., Opitz, L., Schlapbach, R., Rehrauer, H., 2019. Long fragments achieve lower base quality in Illumina paired-end sequencing. *Sci. Reports* 2019 9, 1–7. <https://doi.org/10.1038/s41598-019-39076-7>
- Tan, K.E., Ellis, B.C., Lee, R., Stamper, P.D., Zhang, S.X., Carroll, K.C., 2012. Prospective evaluation of a matrix-assisted laser desorption ionization-time of flight mass spectrometry system in a hospital clinical microbiology laboratory for identification of bacteria and yeasts: A bench-by-bench study for assessing the impact on ti. *J. Clin. Microbiol.* 50, 3301–3308. <https://doi.org/10.1128/JCM.01405-12>
- Theel, E.S., Schmitt, B.H., Hall, L., Cunningham, S.A., Walchak, R.C., Patel, R., Wengenack, N.L., 2012. Formic acid-based direct, on-plate testing of yeast and *Corynebacterium* species by Bruker Biotyper matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J. Clin. Microbiol.* 50, 3093–3095. <https://doi.org/10.1128/JCM.01045-12>
- Thom, C., Smith, C.J., Moore, G., Weir, P., Ijaz, U.Z., 2022. Microbiomes in drinking water treatment and distribution: a meta-analysis from source to tap. *Water Res.* 212, 118106. <https://doi.org/10.1016/j.watres.2022.118106>
- Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., Prill, R.J., Tripathi, A., Gibbons, S.M., Ackermann, G., Navas-Molina, J.A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J.T., Mirarab, S., Xu, Z.Z., Jiang, L., Haroon, M.F., Kanbar, J., Zhu, Q., Song, S.J., Kosciulek, T., Bokulich, N.A., Lefler, J., Brislawn, C.J., Humphrey, G., Owens, S.M., Hampton-Marcell, J., Berg-Lyons, D., McKenzie, V., Fierer, N., Fuhrman, J.A., Clauset, A., Stevens, R.L., Shade, A., Pollard, K.S., Goodwin, K.D., Jansson, J.K., Gilbert, J.A., Knight, R., Agosto Rivera, J.L., Al-Moosawi, L., Alverdy, J., Amato, K.R., Andras, J., Angenent, L.T., Antonopoulos, D.A., Apprill, A., Armitage, D., Ballantine, K., Bárta, J., Baum, J.K., Berry, A., Bhatnagar, A., Bhatnagar, M., Biddle, J.F., Bittner, L., Boldgiv, B., Bottos, E., Boyer, D.M., Braun, J., Brazelton, W., Brearley, F.Q., Campbell, A.H., Caporaso, J.G., Cardona, C., Carroll, J.L., Cary, S.C., Casper, B.B., Charles, T.C., Chu, H., Claar, D.C., Clark, R.G., Clayton, J.B., Clemente, J.C., Cochran, A., Coleman, M.L., Collins, G., Colwell, R.R., Contreras, M., Crary, B.B., Creer, S., Cristol, D.A., Crump, B.C., Cui, D., Daly, S.E., Davalos, L., Dawson, R.D., Defazio, J., Delsuc, F., Dionisi, H.M., Dominguez-Bello, M.G., Dowell, R., Dubinsky, E.A.,

- Dunn, P.O., Ercolini, D., Espinoza, R.E., Ezenwa, V., Fenner, N., Findlay, H.S., Fleming, I.D., Fogliano, V., Forsman, A., Freeman, C., Friedman, E.S., Galindo, G., Garcia, L., Garcia-Amado, M.A., Garshelis, D., Gasser, R.B., Gerds, G., Gibson, M.K., Gifford, I., Gill, R.T., Giray, T., Gittel, A., Golyshin, P., Gong, D., Grossart, H.P., Guyton, K., Haig, S.J., Hale, V., Hall, R.S., Hallam, S.J., Handley, K.M., Hasan, N.A., Haydon, S.R., Hickman, J.E., Hidalgo, G., Hofmockel, K.S., Hooker, J., Hulth, S., Hultman, J., Hyde, E., Ibáñez-Álamo, J.D., Jastrow, J.D., Jex, A.R., Johnson, L.S., Johnston, E.R., Joseph, S., Jurburg, S.D., Jurelevicius, D., Karlsson, A., Karlsson, R., Kauppinen, S., Kellogg, C.T.E., Kennedy, S.J., Kerkhof, L.J., King, G.M., Kling, G.W., Koehler, A. V., Krezalek, M., Kueneman, J., Lamendella, R., Landon, E.M., Lanede Graaf, K., LaRoche, J., Larsen, P., Laverock, B., Lax, S., Lentino, M., Levin, I.I., Liancourt, P., Liang, W., Linz, A.M., Lipson, D.A., Liu, Y., Lladser, M.E., Lozada, M., Spirito, C.M., MacCormack, W.P., MacRae-Crerar, A., Magris, M., Martín-Platero, A.M., Martín-Vivaldi, M., Martínez, L.M., Martínez-Bueno, M., Marzinelli, E.M., Mason, O.U., Mayer, G.D., McDevitt-Irwin, J.M., McDonald, J.E., McGuire, K.L., McMahan, K.D., McMinds, R., Medina, M., Mendelson, J.R., Metcalf, J.L., Meyer, F., Michelangeli, F., Miller, K., Mills, D.A., Minich, J., Mocali, S., Moitinho-Silva, L., Moore, A., Morgan-Kiss, R.M., Munroe, P., Myrold, D., Neufeld, J.D., Ni, Y., Nicol, G.W., Nielsen, S., Nissimov, J.I., Niu, K., Nolan, M.J., Noyce, K., O'Brien, S.L., Okamoto, N., Orlando, L., Castellano, Y.O., Osuolale, O., Oswald, W., Parnell, J., Peralta-Sánchez, J.M., Petraitis, P., Pfister, C., Pilon-Smits, E., Piombino, P., Pointing, S.B., Pollock, F.J., Potter, C., Prithiviraj, B., Quince, C., Rani, A., Ranjan, R., Rao, S., Rees, A.P., Richardson, M., Riebesell, U., Robinson, C., Rockne, K.J., Rodriguez, S.M., Rohwer, F., Roundstone, W., Safran, R.J., Sangwan, N., Sanz, V., Schrenk, M., Schrenzel, M.D., Scott, N.M., Seger, R.L., Seguinorlando, A., Seldin, L., Seyler, L.M., Shakhsher, B., Sheets, G.M., Shen, C., Shi, Y., Shin, H., Shogan, B.D., Shutler, D., Siegel, J., Simmons, S., Sjöling, S., Smith, D.P., Soler, J.J., Sperling, M., Steinberg, P.D., Stephens, B., Stevens, M.A., Taghavi, S., Tai, V., Tait, K., Tan, C.L., Taş, N., Taylor, D.L., Thomas, T., Timling, I., Turner, B.L., Urich, T., Ursell, L.K., Van Der Lelie, D., Van Treuren, W., Van Zwieten, L., Vargas-Robles, D., Thurber, R.V., Vitaglione, P., Walker, D.A., Walters, W.A., Wang, S., Wang, T., Weaver, T., Webster, N.S., Wehrle, B., Weisenhorn, P., Weiss, S., Werner, J.J., West, K., Whitehead, A., Whitehead, S.R., Whittingham, L.A., Willerslev, E., Williams, A.E., Wood, S.A., Woodhams, D.C., Yang, Y., Zaneveld, J., Zarragonandia, I., Zhang, Q., Zhao, H., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 551, 457–463. <https://doi.org/10.1038/nature24621>
- Timperio, A.M., Gorrasi, S., Zolla, L., Fenice, M., 2017. Evaluation of MALDI-TOF mass spectrometry and MALDI BioTyper in comparison to 16S rDNA sequencing for the identification of bacteria isolated from Arctic sea water. *PLoS One* 12, 1–15. <https://doi.org/10.1371/journal.pone.0181860>
- Tsuchida, S., Umemura, H., Nakayama, T., Mauri, P.L., Marchetti-Deschmann, M., Canetti, D., 2020. Current Status of Matrix-Assisted Laser Desorption/Ionization–Time-of-Flight Mass Spectrometry (MALDI-TOF MS) in Clinical Diagnostic Microbiology. *Mol.* 2020, Vol. 25, Page 4775–4775. <https://doi.org/10.3390/MOLECULES25204775>
- UNESCO, 2022. GROUNDWATER Making the invisible visible The United Nations World Water Development Report 2022.
- United Nations, 2020. Monitoring water and sanitation in the 2030. Agenda for Sustainable Development Integrated Monitoring Initiative for SDG 6. <https://www.unwater.org/publications/monitoring-water-and-sanitation-in-the-2030-agenda-for-sustainable-development/>.
- US Water Alliance, 2016. One Water Roadmap: The Sustainable Management of Life's Most Essential Resource.

- van der Wielen, P.W.J.J., van der Kooij, D., 2013. Nontuberculous Mycobacteria, Fungi, and Opportunistic Pathogens in Unchlorinated Drinking Water in the Netherlands. *Appl. Environ. Microbiol.* 79, 825. <https://doi.org/10.1128/AEM.02748-12>
- Vaz-Moreira, I., Nunes, O.C., Manaia, M., 2014. Bacterial diversity and antibiotic resistance in water habitats: searching the links with the human microbiome. *FEMS Microbiol Rev* 38, 761–78. <https://doi.org/10.1111/1574-6976.12062>
- Vidal, L.M.R., Venas, T.M., Gonçalves, A.R.P., Mattsson, H.K., Silva, R.V.P., Nóbrega, M.S., Azevedo, G.P.R., Garcia, G.D., Tschoeke, D.A., Vieira, V. V., Thompson, F.L., Thompson, C.C., 2020. Rapid screening of marine bacterial symbionts using MALDI-TOF MS. *Arch. Microbiol.* 202, 2329–2336. <https://doi.org/10.1007/S00203-020-01917-9>
- Vitorino, L.C., Bessa, L.A., 2018. Microbial Diversity: The Gap between the Estimated and the Known. *Divers.* 2018, Vol. 10, Page 46 10, 46. <https://doi.org/10.3390/D10020046>
- Von Gunten, U., 2003. Ozonation of drinking water: Part II. Disinfection and by-product formation in presence of bromide, iodide or chlorine. *Water Res.* 37, 1469–1487. [https://doi.org/10.1016/S0043-1354\(02\)00458-X](https://doi.org/10.1016/S0043-1354(02)00458-X)
- Wallis, C., Melnick, J.L., 1967. Concentration of enteroviruses on membrane filters. *J. Virol.* 1, 472–477. <https://doi.org/10.1128/JVI.1.3.472-477.1967>
- Walters, K.E., Martiny, J.B.H., 2020. Alpha-, beta-, and gamma-diversity of bacteria varies across habitats. *PLoS One* 15. <https://doi.org/10.1371/JOURNAL.PONE.0233872>
- Wang, F., Li, W., Li, Y., Zhang, J., Chen, J., Zhang, W., Wu, X., 2018. Molecular analysis of bacterial community in the tap water with different water ages of a drinking water distribution system. *Front. Environ. Sci. Eng.* 12. <https://doi.org/10.1007/s11783-018-1020-4>
- Wang, H., Edwards, M.A., Falkinham, J.O., Pruden, A., 2013. Probiotic approach to pathogen control in premise plumbing systems? A review. *Environ. Sci. Technol.* 47, 10117–10128. <https://doi.org/10.1021/ES402455R>
- Whitehead, P.G., Wilby, R.L., Battarbee, R.W., Kernan, M., Wade, A.J., 2009. A review of the potential impacts of climate change on surface water quality. *Hydrol. Sci. J.* 54, 101–123. <https://doi.org/10.1623/hysj.54.1.101>
- WHO, 2022. Drinking water [WWW Document]. URL <https://www.who.int/news-room/fact-sheets/detail/drinking-water> (accessed 5.17.22).
- WHO, 2018. Guidelines on sanitation and health., World Health Organization.
- WHO, 2008. Guidelines for Drinking-water Quality THIRD EDITION INCORPORATING THE FIRST AND SECOND ADDENDA Volume 1 Recommendations Geneva 2008 WHO Library Cataloguing-in-Publication Data.
- Wolf-Baca, M., Piekarska, K., 2020. Biodiversity of organisms inhabiting the water supply network of Wrocław. Detection of pathogenic organisms constituting a threat for drinking water recipients. *Sci. Total Environ.* 715, 136732. <https://doi.org/10.1016/j.scitotenv.2020.136732>
- Yates, M. V., Nakatsu, C.H., Miller, R. V., Pillai, S.D., 2016. Manual of environmental microbiology, Ed.4. ed. American Society for Microbiology (ASM).
- Zhang, W., DiGiano, F.A., 2002. Comparison of bacterial regrowth in distribution systems using free chlorine and chloramine: A statistical study of causative factors. *Water Res.* 36, 1469–1482. [https://doi.org/10.1016/S0043-1354\(01\)00361-X](https://doi.org/10.1016/S0043-1354(01)00361-X)

Zhang, Y., Liu, W.T., 2019. The application of molecular tools to study the drinking water microbiome—Current understanding and future needs. *Crit. Rev. Environ. Sci. Technol.* 49, 1188–1235. <https://doi.org/10.1080/10643389.2019.1571351>

ANNEXOS

1. Altres publicacions

A continuació es detalla la següent publicació no inclosa en la tesi doctoral però realitzada durant el període de doctorat:

Serra-Compte, A., González, S., Arnaldos, M., Berlendis, S., Courtois, S., Loret, J.F., Schlosser, O., Yáñez, A.M., Soria-Soria, E., Fittipaldi, M., Saucedo, G., **Pinar-Méndez, A.**, Paraira, M., Galofré, B., Lema, J.M., Balboa, S., Mauricio-Iglesias, M., Bosch, A., Pintó, R.M., Bertrand, I., Gantzer, C., Montero, C., Litrico, X., 2021. Elimination of SARS-CoV-2 along wastewater and sludge treatment processes. *Water Res.* 202. <https://doi.org/10.1016/j.watres.2021.117435>

“Elimination of SARS-CoV-2 along wastewater and sludge treatment processes” va ser publicat l’any 2021 a la revista *Water Research*. La revista es troba al primer decil (D1) i primer quartil (Q1) a totes les seves categories i disposa d’un factor d’impacte el 2020 de 11,236.

L’estudi ha format part del projecte internacional REVEAL (co-fundat per Suez i fons europeu de desenvolupament regional ERDF (EU) 2020) en el que la doctoranda ha participat activament en la coordinació dels mostreigs, el disseny experimental, desenvolupament i posta a punt dels protocols de bioseguretat, concentració de mostres, extracció d’àcids nucleics i detecció del virus per RT-qPCR, i ha col·laborat en la redacció i en les revisions de l’article.

Eliminació del SARS-CoV-2 al llarg dels processos de tractament de les aigües residuals i els fangs

“Elimination of SARS-CoV-2 along wastewater and sludge treatment processes”

Albert Serra-Compte, Susana González, Marina Arnaldos, Sabrina Berlendis, Sophie Courtois, Jean Francois Loret, Olivier Schlosser, Adela M. Yáñez, Elena Soria-Soria, Mariana Fittipaldi, Gemma Saucedo, **Anna Pinar-Méndez**, Miquel Paraira, Belén Galofré, Juan M. Lema, Sabela Balboa, Miguel Mauricio-Iglesias, Albert Bosch, Rosa M. Pintó, Isabelle Bertrand, Christophe Gantzer, Carlos Montero, Xavier Litrico

Water Research (2021); 202, 117435, DOI: 10.1016/j.watres.2021.117435

El virus del SARS-CoV-2 que ha causat la pandèmia mundial de la COVID-19, a banda d'aïllar-se de mostres respiratòries de pacients infectats, s'ha detectat també en femtes de pacients simptomàtics i asimptomàtics, fins el moment, en forma no infectiva. La presència del material genètic del virus a les aigües residuals ha estat observada arreu del món i ha posat de manifest la rellevància de la seva detecció com a sistema d'alerta primerenca en la vigilància epidemiològica de la COVID-19.

Tot i que la via de transmissió fecal-oral és molt improbable, no hi havia evidències reportades sobre el risc d'infecció per SARS-CoV-2 als treballadors de les estacions depuradores d'aigües residuals (EDAR), així com el risc que suposava l'abocament de les aigües depurades tractades a l'ambient. L'efectivitat dels processos de depuració s'han mostrat efectius per l'eliminació de molts virus, no obstant la informació que es tenia en el moment sobre la reducció de SARS-CoV-2 a les diferents etapes de les EDAR, era molt limitada.

L'objectiu d'aquest estudi es va centrar en el rol de les EDAR per a l'eliminació del SARS-CoV-2 i la protecció dels rius on s'aboquen els efluents tractats.

Per aquest treball 16 EDAR d'Espanya i França, amb incidència de SARS-CoV-2 a les respectives poblacions, es van mostrejar fins a un total de 164 mostres d'aigua residual no tractada (afluents entrada; mostra integrada 24h) i aigua tractada (efluents primari i secundari; 0,5L, i terciari; 20 L) i 107 mostres de fangs no tractats (líquids: primari, secundari i mixt; 0,4L) i fangs tractats (sòlids: fangs deshidratats, digerit i digerit amb hidròlisi tèrmica; 100g).

La concentració de mostres d'aigua residual i fangs líquids van seguir tres protocols diferents segons el laboratori que processava les mostres: ultrafiltració amb Centricon 70 (Bertrand et al., 2021) o amb Amicon Ultra (Balboa et al., 2021), o precipitació amb PEG 8000 (Hjelmsø et al., 2017). En canvi, un únic protocol es va seguir per a l'elució dels fangs sòlids adaptant el protocol de la EPA/625/R-92/013 i Hjelmsø et al., 2017. L'extracció de l'ARN de les mostres es va realitzar mitjançant l'ús de kits d'extracció manuals o bé automatitzats.

La detecció i quantificació del SARS-CoV-2 es va realitzar per RT-qPCR dirigida a regions genòmiques, com ORF1ab i RdRp-IP4 (part de ORF1b), i subgenòmiques, com N i N1 (nucleocàpsida), S (espícula) i E (envolta), analitzats conjuntament amb controls interns i externs d'amplificació per validar l'assaig, com l'ARNasa P humana, el virus de la gastroenteritis transmissible (TGEV, coronavirus porcí) o els fangs RNA F específics del genogrup II.

Adicionalment, es va realitzar un assaig interlaboratori entre els 5 laboratoris participants, que va permetre confirmar la fiabilitat dels resultats obtinguts i la idoneïtat dels mètodes analítics emprats per a la quantificació de l'ARN de SARS-CoV-2 en les mostres d'aigües residuals.

Els resultats de les mostres dels afluents d'entrada a les EDAR es van comparar amb la taxa d'incidència de la COVID-19 a les àrees sanitàries corresponents. Es va observar una positivitats del 50,5% de les mostres amb incidència a la població entre 1 i 100 casos per 100.000 habitants, i amb un percentatge superior (89,5%) en les àrees que presentaven una incidència més gran a 100 casos per 100.000 habitants, mentre que cap positivitats va ser detectada en l'escenari d'incidència inferior a 1 cas per 100.000 habitants.

Respecte la detecció del SARS-CoV-2 als efluents dels tractaments de les EDAR, primari i secundari van eliminar l'ARN de virus al 76,7% de les mostres positives a l'afluent d'entrada, que suposa una reducció de $1,50 \pm 0,80$ logs, essent el tractament secundari per MBR el que va mostrar la millor efectivitat en quan la total eliminació, així com amb l'addició de clor al tractament terciari on tampoc es va detectar cap positivitats (reducció de $1,97 \pm 0,08$ logs).

En alguns casos, les concentracions en els fangs van ser més elevades que en les aigües. Els fangs no tractats van mostrar un 83% de mostres positives en els fangs primaris, i un 57% als secundaris. Respecte els fangs tractats, es va observar una elevada positivitats i superior als fangs no tractats, amb un 69% de mostres positives als fangs deshidratats, i un 71% als fangs digerits, indicant que l'eliminació de l'aigua en aquestes etapes contribueix a una major persistència de l'ARN del virus. L'únic tipus de fang on no es va detectar l'ARN de SARS-CoV-2, va correspondre al fang de digestió anaeròbica seguida d'una hidròlisi tèrmica, degut a les elevades temperatures

aplicades durant aquest tractament, permetent l'eliminació completa de l'ARN viral i per tant garantint la seguretat dels fangs per aplicacions d'ús agrícola.

Addicionalment en una de les EDAR es va analitzar en paral·lel bacteriòfags F-específics d'ARN com a indicadors del SARS-CoV-2 a l'afluent, fang secundari i efluent del secundari, i es va observar una prevalença del fag superior al SARS-CoV-2 on fins i tot a l'efluent només es va detectar el fag, fet que fa que aquests virus siguin indicadors virals adequats per determinar la seva reducció al llarg dels processos de tractament.

Com a conclusions principals, les EDAR van reduir significativament la concentració d'ARN de SARS-CoV-2 als processos de depuració, no obstant no es va produir una eliminació completa en tots els casos després del tractament secundari (excepte per MBR), ni en els fangs no tractats ni tractats (essessits i digerits), però sí en el terciari així com en els fangs de digestió anaeròbica seguida d'una hidròlisi tèrmica. La presència del material genètic del virus no implica infectivitat però caldrien altres estudis per determinar l'estat infecció en els efluent secundaris per verificar la seva nocivitat.



Contents lists available at ScienceDirect

Water Research

journal homepage: www.elsevier.com/locate/watres

Elimination of SARS-CoV-2 along wastewater and sludge treatment processes

Albert Serra-Compte^{1,*}, Susana González¹, Marina Arnaldos¹, Sabrina Berlendis², Sophie Courtois², Jean Francois Loret², Olivier Schlosser², Adela M. Yáñez³, Elena Soria-Soria³, Mariana Fittipaldi³, Gemma Saucedo⁴, Anna Pinar-Méndez⁴, Miquel Paraira⁴, Belén Galofré⁴, Juan M. Lema⁵, Sabela Balboa⁵, Miguel Mauricio-Iglesias⁵, Albert Bosch⁶, Rosa M. Pintó⁶, Isabelle Bertrand⁷, Christophe Gantzer⁷, Carlos Montero¹, Xavier Litrico⁸

¹ Cetaqua, Water Technology Centre, Cornellà de Llobregat, Spain

² Suez, Centre International de Recherche sur l'Eau et l'Environnement (CIRSEE), Le Pecq, France

³ Labaqua, Alicante, Spain

⁴ Aigües de Barcelona, Barcelona, Spain

⁵ CRETUS Institute, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

⁶ University of Barcelona, Barcelona, Spain

⁷ University of Lorraine Lorraine, CNRS, LCPME, Nancy, France

⁸ Suez Groupe, Paris, France

ARTICLE INFO

Keywords:
SARS-CoV-2
COVID-19
Wastewater
WWTP
Sludge

ABSTRACT

The Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) is shed in the feces of infected people. As a consequence, genomic RNA of the virus can be detected in wastewater. Although the presence of viral RNA does not inform on the infectivity of the virus, this presence of genetic material raised the question of the effectiveness of treatment processes in reducing the virus in wastewater and sludge. In this work, treatment lines of 16 wastewater treatment plants were monitored to evaluate the removal of SARS-CoV-2 RNA in raw, processed waters and sludge, from March to May 2020. Viral RNA copies were enumerated using reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) in 5 different laboratories. These laboratories participated in proficiency testing scheme and their results demonstrated the reliability and comparability of the results obtained for each one. SARS-CoV-2 RNA was found in 50.5% of the 101 influent wastewater samples characterized. Positive results were detected more frequently in those regions with a COVID-19 incidence higher than 100 cases per 100,000 inhabitants. Wastewater treatment plants (WWTPs) significantly reduced the occurrence of virus RNA along the water treatment lines. Secondary treatment effluents showed an occurrence of SARS-CoV-2 RNA in 23.3% of the samples and no positive results were found after MBR and chlorination. Non-treated sludge (from primary and secondary treatments) presented a higher occurrence of SARS-CoV-2 RNA than the corresponding water samples, demonstrating the affinity of virus particles for solids. Furthermore, SARS-CoV-2 RNA was detected in treated sludge after thickening and anaerobic digestion, whereas viral RNA was completely eliminated from sludge only when thermal hydrolysis was applied. Finally, co-analysis of SARS-CoV-2 and F-specific RNA bacteriophages was done in the same water and sludge samples in order to investigate the potential use of these bacteriophages as indicators of SARS-CoV-2 fate and reduction along the wastewater treatment.

1. Introduction

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), is

responsible for the COVID-19 disease, a pandemic declared by the World Health Organization (WHO) in March 2020. SARS-CoV-2 belongs to the *Coronaviridae* family. Coronaviruses are enveloped, RNA single-stranded

* Corresponding author at: Ctra. d'Esplugues, 75, 08940 Cornellà de Llobregat, Barcelona, , Spain.
E-mail address: albert.serra@cetaqua.com (A. Serra-Compte).

<https://doi.org/10.1016/j.watres.2021.117435>

Received 26 February 2021; Received in revised form 26 June 2021; Accepted 9 July 2021

Available online 15 July 2021

0043-1354/© 2021 Elsevier Ltd. All rights reserved.

and positively charged viruses (at neutral pH). This family includes other human pathogenic viruses, such as SARS-CoV and MERS-CoV.

COVID-19 induces respiratory symptoms leading to a critical disease such as pneumonia in 14–17% of the cases (Jones et al., 2020). Besides, other symptoms include fever, cough, dyspnea and gastrointestinal symptoms (nausea, vomit, diarrhea). Accordingly, SARS-CoV-2 RNA has been detected in the respiratory tract, blood and feces of infected people (Jones et al., 2020). SARS-CoV-2 shedding by infected people has been observed in asymptomatic individuals (Park et al., 2020), patients during clinical symptomatology and after patient's recovery for a mean of 12.5 days and up to 33 days after respiratory samples reported negative results (van Doorn et al., 2020). However, the viral infectivity in these samples is still not well understood. For instance, Wölfel et al., (2020), suggested that no infective viral particles can be found in patients feces while they were present in the respiratory tract. In addition, Zang et al. (2020) demonstrated that SARS-CoV-2 released into the intestinal lumen were inactivated by simulated human colonic fluid, and infectious virus was not recovered from the stool specimens of COVID-19 patients. On the contrary, (Wang et al., 2020) and (Xiao et al., 2020) recovered infectious virus from stool samples tested positive for SARS-CoV-2 RNA.

The shedding of SARS-CoV-2 RNA in COVID-19 patients feces leads to the occurrence of viral genetic material in wastewater. The presence of SARS-CoV-2 RNA in untreated wastewater has been observed in different studies worldwide (Medema et al., 2020; Wurtzer et al., 2020; La Rosa et al., 2020; Nemudryi et al., 2020; Chavarria-Miró et al., 2021; Ahmed et al., 2020a)). The measurement of SARS-CoV-2 RNA in wastewater may provide helpful information regarding the COVID-19 epidemic trends in particular areas (Lodder and De Roda-Husman, 2020), through the approach named wastewater-based epidemiology, whose application has been suggested for surveillance of previous virus outbreak such as poliovirus (Bancroft, 1957; Wiley, 1962). A good correlation between SARS-CoV-2 RNA occurrence in wastewater and the reported COVID-19 incidence has been shown in different areas (Medema et al., 2020; Weidhaas et al. 2021; Peccia et al. 2020; Zahedi et al. 2020). Furthermore, SARS-CoV-2 RNA in wastewater has been detected in areas with low COVID-19 incidence (Randazzo et al., 2020) or in wastewater samples collected before the first cases were reported (Chavarria-Miró et al., 2021; Medema et al., 2020), highlighting the potential SARS-CoV-2 sewer surveillance as an early warning system (Peccia et al., 2020).

Although preliminary studies performed by Westhaus et al., (2021) suggest that wastewater is unlikely to be a major route of transmission for SARS-CoV-2 to humans, fecal transmission of SARS-CoV-2 has been suggested by two epidemiological studies (Kang et al., 2020; Yuan et al., 2020). However, the body of evidence is still low and to the best of the authors' knowledge, there is to date no reported excess risk of SARS-CoV-2 infection in sewage workers. Improving the knowledge on the fate of SARS-CoV-2 genetic material in WWTPs would be useful to better characterize the risk related with discharges to the environment or wastewater reuse (Lesimple et al., 2020). The effectiveness of wastewater treatment technologies in reducing viruses has been previously documented (Oakley and Mihelcic, 2019), but information on the reduction of specific virus types such as SARS-CoV-2 is still scarce. Randazzo et al., (2020) determined the occurrence of SARS-CoV-2 RNA after secondary treatment, but it was not detected after tertiary treatment. Similarly, other authors highlighted the efficiency of disinfection treatment such as chlorination in the inactivation of SARS-CoV-2 (Collivignarelli et al., 2020).

Considerably less information is available concerning SARS-CoV-2 RNA presence in sludge. Few studies reported the occurrence of SARS-CoV-2 RNA in primary sludge (Balboa et al., 2021; Peccia et al., 2020). Besides, SARS-CoV-2 RNA has been also detected in both, primary and activated sludge in different WWTPs from Turkey (Alpaslan Kocamemi et al., 2020).

Given the widespread detection of SARS-CoV-2 RNA in raw wastewaters, there is the need for comprehensive studies to understand the

role of wastewater treatment in the elimination of SARS-CoV-2 RNA and the protection of water bodies receiving the treated effluents. The present work reports the occurrence and quantification of SARS-CoV-2 RNA along the treatment lines of 16 WWTPs in Spain and France receiving wastewaters from populations presenting different COVID-19 incidence rates. Data are provided on the efficiency of different water treatment technologies in eliminating SARS-CoV-2 RNA, also the fate of SARS-CoV-2 RNA in sludge lines is discussed.

2. Material and methods

2.1. Sample collection

Sixteen WWTPs (8 wastewater treatment facilities in France and 8 in Spain) were sampled. Considering all WWTPs jointly, their water treatment capacity is higher than 10 million population equivalents. Samples were taken from mid-March until the end of May 2020, at different steps of the water treatment process. The characteristics of the monitored WWTPs are specified in the supporting information; Table S1 specifies the water and sludge treatment technology for each monitored WWTP as well as the samples collected at each site. Overall, considering all WWTPs investigated in the study, 164 water samples were analyzed including 101 non-treated and 63 treated wastewater samples, originating from primary, secondary and tertiary effluents. In addition, 107 sludge samples were analyzed, including 56 non-treated sludge and 51 treated sludge samples (thickened sludge, digested sludge and digested sludge plus thermal hydrolysis).

For water analysis, 0.5 L was collected from the WWTP influents and primary and secondary treatment effluents, whereas 20 L were obtained from tertiary treatment effluents. Influent water was collected with automatic samplers as 24 h integrated samples, and the rest of the water samples were collected using grab sampling techniques. Regarding sludge, 0.4 L was sampled from primary, secondary and mixed sludge, whereas 100 g were collected for solid sludge analysis. Both, water and sludge samples were transported under refrigerated conditions and stored at $5 \pm 3^\circ\text{C}$ until sample processing. Samples were processed within 24–48 hours after collection.

2.2. Sample processing

Sample concentration and RNA extraction were performed in five different laboratories depending on the sample's origin. The samples obtained from the WWTPs located in France were analyzed at the LCPME (CNRS and University of Lorraine, Nancy, France) and the CIRSEE (Suez, Le Pecq, France) laboratories while those from Spanish plants were analyzed in three different laboratories, i) University of Barcelona (Spain) for samples from the Barcelona area; ii) University of Santiago de Compostela (Spain), for samples from Ourense and iii) Labaqua (Alicante, Spain) for samples from Alicante, Sabadell and Murcia regions.

Slightly different protocols were followed for water and liquid sludge concentration, depending on the laboratory. Wastewater and liquid sludge samples from the French WWTPs were processed according to the paper by Bertrand et al. (2021), Spanish samples (except Ourense) were processed according to Hjelmsø et al., (2017). Finally, Ourense samples were processed as reported in Balboa et al., (2021). Details on the different processing protocols for wastewater, liquid and solid sludge are reported as supporting information.

RNA extraction from concentrated wastewater and sludge samples was performed in each laboratory according to the corresponding equipment manufacturer's instructions. Briefly, RNA extraction was done using NucliSENS® (BioMérieux) kit for those water samples obtained from French WWTPs. The RNA extraction from the Barcelona area WWTPs samples was done using a Maxwell® RSC PureFood GMO and Authentication kit (Promega). For the rest of Spanish water samples RNA was extracted using QIAmp Viral RNA mini kit by (QIAGEN),

except for Ourense WWTP samples that were extracted using STARMag 96 × 4 Universal Cartridge Kit (Seegene, Seoul, South Korea).

2.3. RT-qPCR analysis

SARS-CoV-2 detection and quantification was done through RT-qPCR analysis at the five different laboratories depending on samples origin (as mentioned in section 2.2).

For the French WWTPs, viral RNA samples were processed with the RNA UltraSense™ One-Step Quantitative RT-PCR System kit (Invitrogen) and the SuperScript III One-Step RT-PCR System (Invitrogen) on a Stratagene MX 3005P real time PCR system (Agilent). The used SARS-CoV-2-related genes primer set included, “E” set developed by Corman et al. (2020) targeting the envelope protein E and “RdRp-IP4” set targeting part of the ORF1ab developed by Pasteur Institute (Paris, France; WHO, 2020). Both genes were quantified according to the method described in Bertrand et al. (2021). Samples from the WWTPs at the Barcelona area were analyzed using the IP4, E and N1 targets with the RNA UltraSense™ One-Step Quantitative RT-PCR System kit (Invitrogen). Samples obtained from Murcia, Sabadell and Alicante WWTPs were processed by an Applied Biosystems 7500 PCR system using the reaction mix Taqman Virus 1-step Master Mix (ThermoFisher; 4444436) combined with TaqMan 2019-nCoV Kit (ThermoFisher; A47532) for the analysis of three target genes, ORF1ab (RNApol), S gene, N gene. Finally, samples obtained from Ourense WWTP were characterized for viral RNA determination by a one-step multiplex RT-qPCR Allplex system™ 2019-nCoV (Seegene, Seoul, South Korea); the target genes were RdRp, N gene and E gene. The specificity of the target genes used in each laboratory has been previously reported in bibliography (references are described at table 1 for each target gene). All target genes had a high specificity for SARS-CoV-2, except for E gene that is specific for SARS-Coronavirus (Barra et al., 2020). Therefore, to minimize the risk of false positives results, a minimum of two out of all target genes used in each laboratory should give positive results to consider the sample as positive. For quantification purposes of SARS-CoV-2 genomic copies, ten-fold serial dilutions of a quantified SARS-CoV-2 RNA reference

material (Twist synthetic SARS-CoV-2 RNA control 2 (MN908947.3; Ref.: 102024; Twist bioscience) was used to prepare the standard curves to compare the Cycle threshold (Ct) values obtained in the analyzed samples. The highest concentration detected from the different targeted genes was reported for each positive sample.

To account for role of potential inhibitors blocking amplification, Amplification Controls, internal (IAC) or external (EAC) were included in each sample. Briefly, in the case of Labaqua samples a fragment of the human RNase P was amplified simultaneously with each specific RNA target (TaqMan™ 2019nCoV Assay Kit v1). A mean value for RNase P was obtained by analyzing at least 30 negative samples. Then an acceptance interval was established as the media value of ± 3 times the standard deviation. Samples with RNase P outside this interval were considered inhibited (ISO 22174 2005). Regarding the University of Santiago de Compostela, internal control provided by the RT-qPCR kit (Allplex 2019-nCoV Assay test) was used in raw and diluted samples. In the University of Barcelona laboratory, TGV virus was used, as well as logarithmic dilutions (1/10). For French samples the estimation of RT-qPCR inhibition in the wastewater samples was based on the concentrations of F-specific RNA phages of genogroup II (FRNAPH GGII) detected by using the primers set “VTB4-F-GII” published by (Wolf et al., 2010). For raw wastewater, logarithmic dilutions (1/10 and 1/100) were performed in PCR grade water following viral RNA extraction. The RT-qPCR assay was then carried out on both undiluted and diluted RNA extracts. The percentage of inhibition was estimated for undiluted and 1/10 samples by taking the concentration obtained from the 1/100 samples as a reference, due to the high dilution of potential inhibitors. For sludges, values of concentrations of FRNAPH GGII phages were used as quality control to take in account role of potential inhibitors of the sample matrix.

In addition to SARS-CoV-2 RNA, F-specific bacteriophages RNA was characterized in some samples. Briefly, for the FRNAPH GGII genome, the VTB4-Fph GII set published by Wolf et al. (2010) was used with the SuperScript III One-Step RT-PCR System with 1 μ M of each primer, 0.275 μ M of probe in 20 μ l final reaction volume with 2 μ l of RNA and 0, 75 μ l of Superscript III RT/Platinum Taq mix. On the Stratagene system,

Table 1

Analytical performance parameters from the five laboratories involved in the study. Recoveries, detection and quantification limits, standard curves R^2 and efficiency are reported as mean values per laboratory.

Laboratory	Target genes (reference)	Standard curvesSlope	Inhibition assessment	Recovery (%)	Reference material	Detection limit complete process
Labagua	ORF1ab (Lue et al., 2020) S gene (Lu et al., 2020) N gene (Lue et al., 2020)	-3.22 -3.13 -3.49	Human RNaseP	11-55% (n=158)	Accuplex SARS-CoV-2 Reference material kit	$7,2 \times 10^2$ gc/L
University of Barcelona	RdRp IP4, (Institut Pasteur, Paris) E (Charité, Berlin) N1 gene (CDC, Atlanta)	-3.40 -3.44 -3.51	Dilution analysis*	2.53% \pm 0.17% (n=70)	SARS-CoV-2 synthetic RNA (MN908947.3, Twist Bioscience)	1×10^2 gc/L for IP4, E and N1 genes.
LCPME (University of Lorraine)	RdRp IP4 (Pasteur Institute, Paris, France) E gene (Corman et al., 2020)	-3.61 -3.28	FRNAPH GII	55.8-64.0 % (n=4)	RdRp-IP4: RNA extracted from tested positive patients Gene E: eurofins genomics	$2 \times 10^2 - 2 \times 10^3$ for RdRp IP4 genes. 2×10^2 gc/L for E genes.
CIRSEE	RdRp IP4 (Pasteur Institute, Paris, France) E gene (Corman et al., 2020)	-3.21 -3.47	FRNAPH GII	12-116% (n=7)	RdRp -IP4 and E amplicons in genscript plasmid from in vitro transcribed RNA derived from strain BetaCov_Wuhan_WIV04_2019 (NC_004718.3) (Pasteur institute donation). (Pezzi et al., 2020)	$2 \times 10^2 - 2 \times 10^3$ for RdRp IP4 genes. 2×10^2 gc/L for E genes.
Univ. Santiago de Compostela	E, RdRp and N (not provided by kit)	-3.24		33.3 % \pm 15.6% (n=50)	RT-qPCR Allplex system™ 2019-nCoV Reference material kit,	1×10^3 gc/L

* Testing two replicates of the direct sample and two replicates of the one tenth dilution

the RT was conducted at 55°C during 20 min followed by an initial denaturation at 94°C during 3 min and 45 cycles including denaturation step at 94°C during 15s, hybridization step at 58°C during 30s. Final cooling is done at 40°C during 30s.

2.4. Proficiency testing scheme

To evaluate the performance of the laboratories involved in the present work and to assess the reliability of the different methods used, a Proficiency Testing (PT) scheme was performed. The test was coordinated by ielab, a company accredited by Entidad Nacional de Acreditación (ENAC) for organizing PT schemes based on ISO/IEC17043 standard and the (CGA-ENAC-PPI 2020) guide. In order to perform the test, the supplied material included two homogeneous and stable samples in liquid format (1 mL vial each) containing natural SARS-CoV-2 and one plastic bottle containing the matrix (500 mL of secondary wastewater effluent). Regarding the analysis of the provided sample vials, one of them (vial A) had to be analyzed directly, to evaluate RNA extraction and RT-qPCR steps, while the other vial (B) had to be used to inoculate the provided matrix. In this case, the evaluation comprised the concentration of the inoculated sample matrix, RNA extraction and RT-qPCR steps.

All laboratories analyzed the two materials in triplicate providing qualitative and quantitative results. The PT scheme data pre-treatment was performed according to Murtula et al., 2012, including, (i) eliminate laboratories reporting false positive or negative results; (ii) make a logarithmic conversion (log10) of the results, which allows obtaining a Gaussian distribution (normal), and (iii) detect the outlier laboratories. The performance of the laboratories was assessed through a Z-score approach, as required by the (ISO 13528 2015) standard. The z-scores were calculated from the laboratory results as follows:

$$z = \frac{x - X}{\sigma_{pt}}$$

where x is the average of the logarithmic values of the results of a laboratory; X is the assigned value; and σ_{pt} is the Standard Deviation for Proficiency Assessment (SDPA). To establish the assigned value the participant's PT results were used (Murtula et al., 2012). SDPA defines the acceptable range variation of the laboratories on each assay. In this PT scheme, the SDPA value has been fixed to 0.75 Log by ielab Steering Committee to adequate it as much as possible to the fitness for purpose. This value was determined considering the standard deviation that is found in the concentration step and in the RT-qPCR/RT-PCR step.

The Z-score indicates how far away the score of each participant is from the mean in units of standard deviations and in which direction (positive or negative). The interpretation of the z-score is the following: $|Z| \leq 2 =$ satisfactory results; $2 < |Z| \leq 3 =$ questionable results; and $|Z| \geq 3 =$ unsatisfactory results (ISO 13528:2015).

2.5. Data treatment

Log removal values were calculated as follows:

$$\text{Log removal} = \log_{10} \frac{\text{Influent concentration}}{\text{Effluent concentration}}$$

Influent concentration refers to the SARS-CoV-2 genetic copies / L in the influent of each WWTP. Only those samples showing positive results for SARS-CoV-2 RNA in the influent were considered for the log removal calculation. Effluent concentration refers to SARS-CoV-2 genetic copies / L in the effluent of each corresponding treatment step. For solid sludge samples, the RNA concentration was expressed in genetic copies / Kg. Log removal was calculated for each day of analysis and the values are presented as mean \pm standard deviation for each treatment analyzed. When virus RNA was not detected at the effluent of a given treatment, a value of 2 logs was used for log removal calculation corresponding to the

average limit of quantification of all methods.

3. Results and discussion

3.1. Quantitative detection of SARS-CoV-2 in wastewater samples

The performance of the different RT-qPCR methodologies applied in the study is reported in Table 1. All methodologies provided good performance in terms of the slope of the standards, recoveries and detection limits.

3.2. Proficiency testing for the quantification of virus RNA in wastewater

The PT scheme included a total of 36 laboratories using a variety of different SARS-CoV-2 RNA detection and quantification RT-qPCR methods. The 5 laboratories participating in the present work, correctly reported qualitative results (positive/negative) values. Besides, from the 36 laboratories participating in the PT scheme, 19 reported quantitative results. In order to evaluate the performance of the participants regarding the quantified results reported, the Z-score was estimated. The obtained absolute Z-score value for all the laboratories involved in the present work was lower than 2 for vial A, which is intended to evaluate the RNA extraction and RT-qPCR steps. The same result (a Z-score lower than 2) was obtained also for all the laboratories participating in the present work when analyzing and quantifying vial B, which evaluates the concentration process, RNA extraction and RT-qPCR analysis. Therefore, the reliability of the results obtained for all the laboratories and the suitability of their analytical methods for SARS-CoV-2 RNA quantification in wastewater was confirmed as established in the (ISO 13528 2015) standard.

3.3. SARS-CoV-2 RNA in untreated wastewater and relationship with COVID-19 cases

The occurrence and concentration of SARS-CoV-2 RNA in wastewater are expected to be related with the COVID-19 incidence rate in the corresponding area. It is hypothesized that the amount of SARS-CoV-2 RNA in wastewater is depending on the viral load shed in the corresponding area due to both, symptomatic and asymptomatic cases. Figure 1 shows all the influent wastewater samples collected from the different WWTPs in which SARS-CoV-2 RNA was either detected (filled dots) or not detected (empty dots) and the Y-axis indicates COVID-19 prevalence rate in the corresponding area at the time of sampling. Besides, Figure S1 shows the evolution of SARS-CoV-2 RNA quantification at the influent of each WWTP over time with the changes in the reported COVID-19 incidence values.

From the 101 wastewater influent samples collected, half of them (50.5%) showed positive results for SARS-CoV-2 RNA occurrence. Most of the wastewater influent samples (89.5%) collected from WWTPs located in areas with a COVID-19 incidence higher than 100 cases per 100,000 inhabitants presented positive results of SARS-CoV-2 RNA. The number of positive samples from regions with a COVID-19 incidence higher than 100 cases per 100,000 inhabitants was significantly higher than the ones reporting lower COVID-19 incidence, between 1 to 100 cases per 100,000 inhabitants (p-value < 0.05, Chi-squared test). Samples collected in these areas (COVID-19 incidence rate between 1 to 100 cases per 100,000 inhabitants) showed high variability, being about half of the samples collected positive for SARS-CoV-2 RNA. Finally, no positive samples were obtained when the incidence was lower than 1 case per 100,000 inhabitants, however only three samples were characterized under this COVID-19 prevalence.

The evolution of SARS-CoV-2 RNA concentration at the influent is represented for each WWTP with the corresponding incidence value at supporting information, Figure S1. Most of the plants showed a decreasing number of SARS-CoV-2 RNA copies along the sampling time. This decreasing trend could be related in most cases with a decrease in

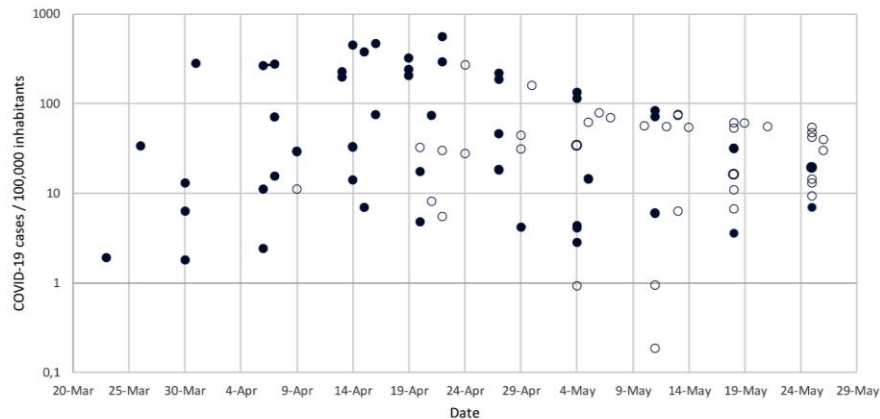


Fig. 1. Influent wastewater samples from the different targeted WWTP. Y-axis represents the COVID-19 prevalence rate in the corresponding area at the time of sampling. Filled dots represent SARS-CoV-2 RNA positive samples whereas empty dots represent no SARS-CoV-2 RNA occurrence.

the reported COVID-19 cases in the corresponding regions along the sampling period. However, this trend did not apply to all target sites. It is important to note that the testing capacity to report the COVID-19 incidence, was more limited at the beginning of the pandemic (when samples were taken in France and Spain for the present study) than in the later months. For instance, the reported COVID-19 cases in French regions are explained at the department level, which in some cases, the

disease trend may not be representative of the local variation of the WWTP sanitation area. Besides, some factors could provoke variability in the reported SARS-CoV-2 RNA concentration in wastewater. For instance, in a thorough assessment of uncertainty in SARS-CoV-2 prevalence estimation by wastewater-based epidemiology, [Li et al \(2021\)](#) pinpointed precipitation and the decay in sewers as major obstacles to prevent higher accuracy. Correlation between higher COVID-19

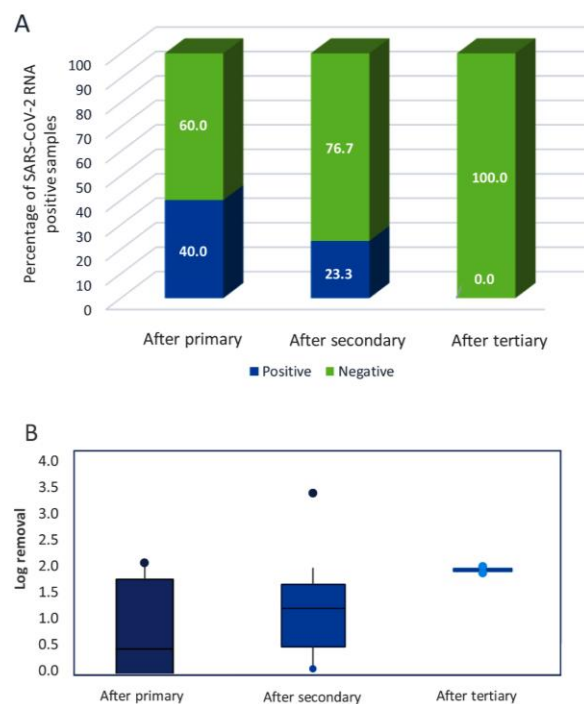


Fig. 2. The graph shows integrated results of all WWTPs monitored. Primary (n=5), secondary (n=30), tertiary (n=2). A) Percentage of positive and negative samples for SARS-CoV-2 RNA occurrence in each water treatment step and B) SARS-CoV-2 RNA log removal in each treatment step comparing influent and effluent concentration in each treatment.

incidence and positive samples in wastewater has been observed in previous studies, although high variability in detection has been reported. Hart and Halden (2020) performed a detailed computational analysis, showing that the detection limits would range from 1 in 114 individuals to 1 in 2,000,000 individuals (spanning 5 orders of magnitude). Despite the fact that wastewater based epidemiology is out of the scope of this study, the obtained results are in line with previous works that indicated an association between health data and viral RNA quantification in wastewater (Medema et al., 2020; Randazzo et al., 2020; Weidhaas, 2021)).

To further evaluate the evolution of SARS-CoV-2 RNA in wastewater treatment, only those WWTPs showing positive results in the influent were considered in the following sections.

3.4. SARS-CoV-2 RNA in WWTP water lines

In order to evaluate the removal of SARS-CoV-2 RNA in WWTPs water lines, samples were taken at the effluent of different WWTPs secondary treatment (which considers the removal of primary plus secondary treatment) as it is the most common configuration for WWTPs worldwide. Besides, to understand the role of each treatment step, few samples were obtained after primary treatment only, whereas some of them were obtained after tertiary treatment (which considers primary, secondary and tertiary treatment). Figure 2 shows the occurrence of SARS-CoV-2 RNA at the different wastewater treatment steps, as well as the log removal.

The common configuration of WWTPs, primary plus secondary treatment, eliminated SARS-CoV-2 RNA from the 76,7% of the positive samples occurring at the influent (only 7 out of the 30 secondary treatment samples analyzed showed positive results) with a virus RNA removal up to 1.50 ± 0.80 logs (Figure 2, table S2). From this removal, a significant part might correspond to the primary treatment which allowed to eliminate SARS-CoV-2 RNA from the 60% of the positive samples analyzed (only 2 out of the 5 samples analyzed showed presence of SARS-CoV-2 RNA) with a mean log removal of 0.48 ± 1.17 log of magnitude (figure 2, table S2). Finally, no occurrence of virus RNA was determined after tertiary treatments (2 samples analyzed) achieving more than 1.97 ± 0.08 logs of reduction.

SARS-CoV-2 RNA occurrence was significantly reduced after the water treatment steps. However, our study showed that about 23% of the effluent samples from secondary treatments were tested positive for SARS-CoV-2 RNA. Previous studies also reported positive samples of SARS-CoV-2 RNA after secondary treatments, including two WWTPs from Israel (Abu Ali et al., 2020), two WWTPs in Spain (Randazzo et al., 2020) and one WWTP in Japan (Haramoto et al., 2020). On the contrary, other studies reported no occurrence of SARS-CoV-2 RNA after secondary treatment, despite positive samples measured in the inflow (Kumar et al., 2020; Sherchan et al., 2020). These results suggest that secondary treatments are not able to eliminate SARS-CoV-2 RNA in all cases. However, virus detection after treatment can be strongly influenced by the influent viral concentration, as well as by the nature of the secondary treatment applied. Abu Ali et al., (2020) detected SARS-CoV-2 RNA at the effluent of secondary treatment when the virus influent concentration was close to 6 log copies / L. A similar influent concentration was reported in one WWTP in Spain when positive samples for SARS-CoV-2 RNA in the effluent were detected. However, an additional positive effluent sample was reported in Spain when the influent concentration was lower than 2 log copies / L (Randazzo et al., 2020). Therefore, the presence of RNA in effluent from secondary treatment has been reported in the literature at various levels of RNA concentration at the influent (between 2 and 6 log copies / L). It is likely that the secondary treatment plays an important role in the final SARS-CoV-2 RNA elimination rate. However, the scarcity of reported data on secondary treatment inlet/outlet measurements hinders a more accurate analysis of the fate of SARS-CoV-2 RNA in this section, which would be needed to cope with the potential variability due to the

different analytical methods applied or the variation of the wastewater characteristics from the different studies.

In order to evaluate potential differences in the elimination of SARS-CoV-2 RNA within different secondary treatments, the three secondary treatment technologies covered in the study, namely activated sludge, activated sludge plus nutrient removal and membrane bioreactor (MBR) were compared. For each of them, the detection occurrence and the log removal of SARS-CoV-2 RNA are shown in Figure 3.

SARS-CoV-2 RNA was not detected in 63.6% of the samples after activated sludge followed by clarification (4 out of the 11 activated sludge effluent samples showed positive results). Detection in the effluent was less frequent in activated sludge plus nutrient removal followed by clarification where SARS-CoV-2 RNA was not detected in 81% of the samples analyzed after this treatment (2 out of the 11 activated sludge plus nutrient removal effluent samples showed positive results). Finally, there were no SARS-CoV-2 RNA positive samples in the effluent from the MBR treatment (8 samples analyzed) (Figure 3A). The same trend was observed when the log removals were calculated for each treatment; samples obtained after activated sludge and clarification, showed a mean value of 1.03 ± 0.59 , followed by activated sludge plus nutrient removal and clarification, with 1.37 ± 0.72 and the highest log removal was obtained after MBR treatment being more than 1.97 ± 0.93 log (Figure 3, Table S2). Our results suggested MBR as the most effective secondary treatment in the elimination of SARS-CoV-2 RNA from the water line. However, a high variability was observed in the measures (presented as standard deviation) which may difficult to determine significant differences concerning the efficiency in the elimination of SARS-CoV-2 RNA within the different treatments.

Scarce information is available regarding the efficiency of water treatment technologies in the elimination of SARS-CoV-2. For instance, positive samples of SARS-CoV-2 RNA were observed after conventional activated sludge treatment in three WWTPs from Germany (Westhaus et al., 2021). Primary settling followed by activated sludge treatment provided an elimination rate close to 2 logs for SARS-CoV-2 (Abu Ali et al., 2020), which is slightly higher than the one reported in the present work. Previous studies showed a higher efficiency of MBR treatment compared to activated sludge process in the elimination of adenovirus (Amoah et al. 2020). Removal for different types of viruses including adenovirus, norovirus and F+ coliphages during MBR treatment has been reported as being between 2 and 7 logs (Amoah et al. 2020). These results of other viruses' elimination rate agree with the ones found for SARS-CoV-2 removal in the present work; MBR treatment has been shown to be more effective than conventional sludge with a log reduction for SARS-CoV-2 in a similar range than the ones reported for other type of viruses. Secondary treatments including activated sludge and MBR can be effective in reducing SARS-CoV-2 and other type of viruses from the water line, due to the adverse conditions that viruses are exposed to in these treatments (Lesimple et al., 2020) and because of virus sorption to organic particles and further elimination by settling (Bogler et al., 2020). Further virus removal can be achieved by MBR through virus particle retention by membrane and cake layer (Bhatt et al. 2020).

In addition to the secondary treatment evaluation, our study provided insights on the role of primary treatment only in the removal of SARS-CoV-2 RNA, as well as the addition of a tertiary process in the water line. The obtained results on the primary treatment SARS-CoV-2 RNA removal of 0.48 ± 1.17 log are in the range of previously reported elimination for WWTPs primary treatment (Abu Ali et al. 2021) that observed an elimination close to 1 log. Whereas no detection of SARS-CoV-2 RNA was determined after the addition of chlorination, applied as tertiary treatment. There is a general agreement that tertiary treatments are effective in the elimination of SARS-CoV-2 from the water lines (Bogler et al., 2020; Lesimple et al., 2020; Randazzo et al., 2020). Specifically, chlorination has been previously determined as an effective treatment for the elimination of SARS-CoV-2 (Abu Ali et al., 2020). Nevertheless, further characterization of SARS-CoV-2 RNA

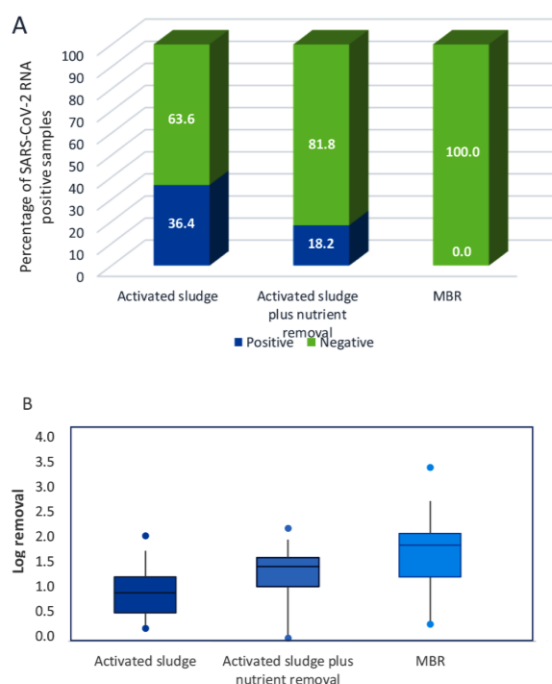


Fig. 3. The graph shows integrated results of all WWTPs monitored. Activated sludge (n=11), activated sludge plus nutrient removal (n=11), MBR (n=8). A) Percentage of positive and negative samples for SARS-CoV-2 RNA occurrence in each treatment process and B) SARS-CoV-2 RNA log removal in each treatment process comparing influent and effluent concentrations after each treatment.

occurrence after primary treatment only and the addition of a tertiary treatment should be performed to confirm the insight provided in the present work.

3.5. SARS-CoV-2 RNA in WWTP sludge lines

Sludge samples were obtained from those WWTPs showing SARS-CoV-2 RNA positive results in the water influent samples. Firstly, non-treated sludge was obtained from primary and secondary treatments and secondly, treated sludge was collected after sludge thickening, anaerobic digestion and anaerobic digestion plus thermal hydrolysis (Figure 4).

Non-treated sludge samples presented a high percentage of SARS-CoV-2 RNA occurrence, being 83% of the samples positive. However, a small percentage of positive samples at the influent were negative for SARS-CoV-2 RNA occurrence in primary sludge, that might be related to heavy rainfall and low solids content in the influent, as no inhibition was detected in those samples. In the case of secondary sludge, the 57% of the samples were positive (Figure 4A). These percentages of occurrence are consistent with the elimination observed in water treatments (with a detection in only 40% of the effluents after primary treatment and 23% after secondary treatment). This can be explained by the fact that virus particles can adsorb to solids and colloidal material because of the lipid bilayer surrounding the SARS-CoV-2 protein capsid (Balboa et al., 2021). In accordance, previous studies have already reported an important occurrence of SARS-CoV-2 RNA in primary and secondary sludge (Alpaslan-Kocameci et al., 2020; Balboa et al., 2021; Peccia et al., 2020), and in some cases with concentrations higher than in water samples, which highlights SARS-CoV-2 affinity for solids.

Similarly to non-treated sludge, thickened and digested sludge also presented a high percentage of SARS-CoV-2 RNA positive samples, being

69% and 71%, respectively. In both, thickened and digested sludge, the reported viral log removal was negative, being mean values of -0.47 ± 1.20 and -0.17 ± 0.89 , respectively (table S2). A negative log removal indicates higher viral RNA concentration after the treatment when compared to the entrance. The elimination of water at these steps certainly contributes to the increase in viral RNA concentration in these matrices. Previous studies reported an important occurrence of SARS-CoV-2 in thickened sludge and pinpointed it as a hotspot of virus RNA in the plant (Balboa et al., 2021). Besides, digested sludge showed an occurrence of virus RNA in a similar proportion than thickened sludge. To the best of the authors' knowledge, no previous information regarding SARS-CoV-2 RNA in digested sludge is available. Previous studies with other human coronaviruses showed the occurrence of these types of viruses in sludge samples after anaerobic digestion (Bibby et al., 2011; Bibby and Peccia, 2013). These results pinpoint that human coronavirus genome can be found after digestive phases, in line with the results of the present work with SARS-CoV-2.

The only group of sludge samples with no detection of SARS-CoV-2 RNA was after anaerobic digestion followed by thermal hydrolysis. The high temperatures applied during this treatment ($150\text{--}160^\circ\text{C}$ saturated steam) allowed the complete elimination of viral RNA with a log removal higher than 1.69 ± 0.27 . Therefore, sludge thermal hydrolysis enabled a complete inactivation of SARS-CoV-2 and guarantees sludge safety if it is applied for agricultural purposes.

3.6. SARS-CoV-2 and F-specific RNA bacteriophages load along WWTP processes

Bacteriophages are commonly used as viral indicators to determine the efficiency of water treatments and guaranty safe water use and reuse (Withy et al., 2005). Figure 5 shows the concentration of genome of

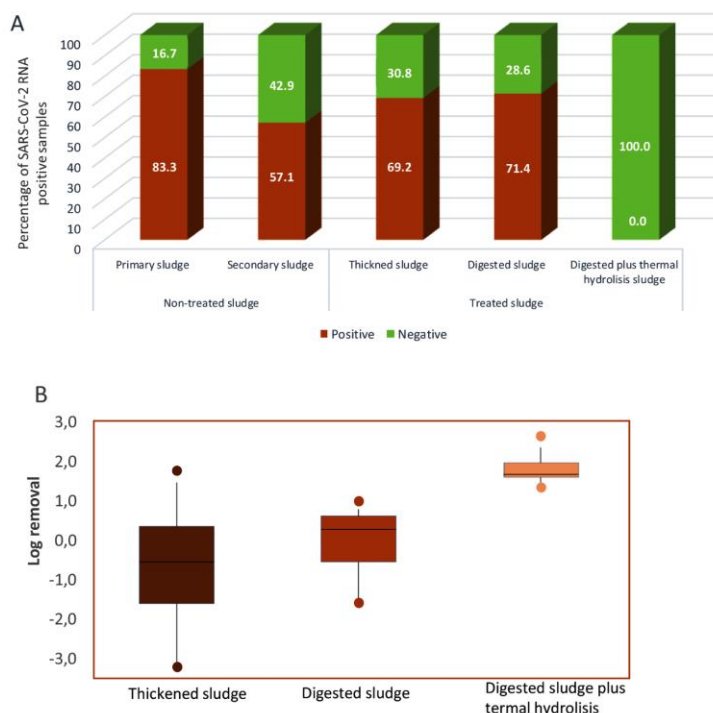


Fig. 4. The graph shows integrated results of all WWTPs monitored. Primary sludge (n=6), secondary sludge (n=14), thickened sludge (n=13), digested sludge (n=7), digested sludge plus thermal hydrolysis (n=5) A) Percentage of positive and negative samples for SARS-CoV-2 RNA in sludge samples and B) Boxplot showing first and third quartiles and the median value of SARS-CoV-2 RNA log removal in each sludge matrix comparing influent and effluent samples after each treatment.

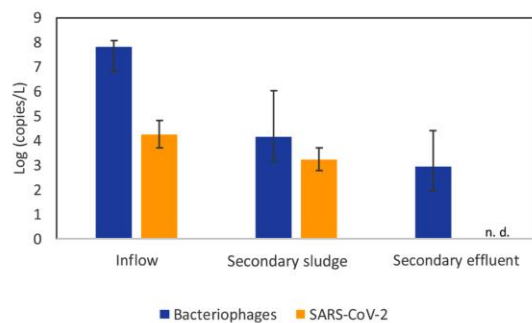


Fig. 5. SARS-CoV-2 RNA and F-specific RNA bacteriophages concentration Log (copies/L) along the different steps of a selected WWTP. n. d. = not detected. Bars represent mean values (n=4) with the corresponding standard deviation.

both, SARS-CoV-2 and F-specific RNA bacteriophages in the WWTP14. This WWTP has a secondary treatment based on activated sludge. Water was sampled in the influent of the plant and in the effluent (after activated sludge and clarification). Secondary sludge was collected from the activated sludge process (Figure 5). Paired samples were taken from 4 different days in the WWTP target processes.

The concentration of bacteriophages in the influent of the WWTP was higher than the concentration of SARS-CoV-2, with mean values of 7.80 ± 0.26 log copies/L and 4.26 ± 0.56 log copies/L, respectively. Bacteriophages presented also a higher concentrations than SARS-CoV-2

in activated sludge, although the difference in concentration between both viruses was reduced at this step, bacteriophages and SARS-CoV-2 presenting respectively a mean concentration of 4.19 ± 1.9 log copies/L and 3.23 ± 0.46 log copies/L in this matrix (Figure 5, Table S3). At the effluent of the plant (after activated sludge and clarification), SARS-CoV-2 RNA was not detected in any of the samples characterized, whereas bacteriophages were still detected in all of them with a mean concentration of 2.94 ± 1.52 log copies/L. The wastewater treatment line was able to eliminate SARS-CoV-2 down to below detection limits, but not bacteriophages. However, bacteriophages presented an initial concentration about three orders of magnitude higher than SARS-CoV-2. F-specific RNA bacteriophages may be valid indicators of viral elimination in water treatment due to its high initial amounts and measurable concentrations after the treatment process. Nevertheless, further studies would be needed to determine the sensitivity of both virus along the wastewater treatment process; considering if possible, higher concentrations of SARS-CoV-2 in the influent, enabling a detection of the virus also in the effluent, in order to compare more precisely the reduction of both viruses in the same conditions.

4. Conclusions

WWTPs significantly reduced the occurrence and concentration of SARS-CoV-2 RNA in the water lines. However, a complete removal was not observed in all cases in this study after secondary treatment, and viral RNA could be still detected after activated sludge and activated sludge plus nutrient removal followed by clarification. On the contrary, SARS-CoV-2 RNA was not detected after MBR and chlorination applied as a tertiary treatment.

SARS-CoV-2 RNA was detected in large proportions of all types of sludge, including sludge from primary and secondary treatments and treated sludge after thickening and anaerobic digestion. These results highlight that these treatments are not effective in eliminating SARS-CoV-2 RNA from sludge, while, only after applying thermal hydrolysis SARS-CoV-2 RNA was not detected.

The presence of virus RNA in the analyzed samples (wastewater and sludge) does not imply infectivity nor integrity. Further studies are required to estimate the infectious state of the virus in the effluents from secondary treatments and after sludge thickening and digestion to verify whether additional treatments are needed to protect the receiving environments.

Finally, F-specific RNA bacteriophages presented high concentrations at the influent which makes these viruses suitable viral indicators to determine its reduction along treatment processes. However, further research should be done in order to determine the potential of bacteriophages as indicators of SARS-CoV-2 elimination in water and sludge treatments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was financed by SUEZ company. The authors would like to thank the Àrea Metropolitana de Barcelona (AMB), the city halls of Sabadell, Alicante and Ourense. Also, thanks to Entidad de Saneamiento y Depuración de Aguas Residuales de la Región de Murcia (ESAMUR) and to the water authorities from the regions covering the present work. Authors also acknowledge operators and local authorities (Métropole de Lyon, Eurométropole de Strasbourg, Dijon métropole and Toulouse métropole) for the sampling effort made in the French WWTPS and the commitment of Suez Eau France Technical Direction (Hubert Dupont and Thierry Lebrun). The study was also supported by the French Scientific Group of interest named Obépine. Sabela Balboa, Miguel Mauricio-Iglesias and Juan M. Lema belong to the CRETUS Strategic Partnership (ED431E 2018/01) and to the Galician Competitive Research Group (ED431C2017/029). All these programs are co-funded by ERDF (EU).

Authors contribution

M. A., S. G., J. F. L., O. S., A. M. Y., M. P., B. G., G. S., A. P. M., J. M. L., C. M., and X. L. Designed and validated the experimental strategy and methods. S. C., S. Be., E. S. S., G. S., A. P. M., B. G., S. Ba., M. M., A. B., R. M. P., I. B., C. G. Coordinated sampling and performed extraction and analysis through q-PCR for SARS-COV-2 determination. M. F. and A. M. Y. contributed in interlaboratory assay. I. B., C. G., S. Be., S. C., performed bacteriophages analysis. A. S. C., S. Be., A. M. Y., wrote the manuscript. All authors revised and agreed with the manuscript content.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2021.117435.

References

- Abu Ali, H., Yaniv, K., Bar Zeev, E., Chaudhury, S., Shaga, M., Ronen, Z., Kushmaro, A., Nir, O., 2020. Tracking SARS-CoV-2 RNA through the wastewater 1 treatment process. medRxiv, 2020.10.14.20212837.
- Ahmed, W., Angel, N., Edson, J., Bibby, K., Bivins, A., O'Brien, J.W., Choi, P.M., Kitajima, M., Simpson, S.L., Li, J., Tscharke, B., Verhagen, R., Smith, W.J.M., Zaugg, J., Dierens, L., Hugenholz, P., Thomas, K.V., Mueller, J.F., 2020a. First

- confirmed detection of SARS-CoV-2 in untreated wastewater in Australia: A proof of concept for the wastewater surveillance of COVID-19 in the community. *Sci. Total Environ.* 728, 138764 <https://doi.org/10.1016/j.scitotenv.2020.138764>.
- Alpaslan-Kocememli, B., Kurt, H., Sait, A., Sarac, F., Saatci, A.M., Pakdemirli, B., 2020. SARS-CoV-2 detection in Istanbul wastewater treatment plant sludges. medRxiv. <https://doi.org/10.1101/2020.05.12.20099358>, 2020.05.12.20099358.
- Amoah, D.I., Kumari, S., Bux, F., 2020. Coronaviruses in wastewater processes: Source, fate and potential risks. *Environment International* 143, 105962.
- Balboa, S., Mauricio-Iglesias, M., Rodriguez, S., Martínez-Lamas, L., Vasallo, F.J., Regueiro, B., Lema, J.M., 2021. The fate of SARS-CoV-2 in WWTPS points out the sludge line as a suitable spot for monitoring. *Sci. Tot. Environm.* 772 <https://doi.org/10.1016/j.scitotenv.2021.145268>.
- Bancroft, P.M., Engelhard, W.E., Evans, C.A., 1957. Poliomyelitis in Huskerville (Lincoln) Nebraska; studies indicating a relationship between clinically severe infections and proximate fecal pollution of water. *J Am Med Assoc.* 164, 836–847. <https://doi.org/10.1001/jama.1957.02980080006002>.
- Barra, G.B., Santa Rita, T.H., Mesquita, P.G., Jácome, R.H., Nery, L.F.A., 2020. Analytical Sensitivity and Specificity of Two RT-qPCR Protocols for SARS-CoV-2 Detection Performed in an Automated Workflow. *Genes (Base)* 12 (10), 1183. <https://doi.org/10.3390/genes12101183>, 11.
- Bertrand, I., Challant, J., Jeulin, H., Hartard, C., Mathieu, L., Lopez, S., Interest, S., Obépine, G., Schvoerer, E., Courtois, S., Gantzer, C., 2021. Epidemiological surveillance of SARS-CoV-2 by genome quantification in wastewater applied to a city in the northeast of France: comparison of ultrafiltration- and protein precipitation-based methods. *Int. J. Hyg. Environ. Health* 113692. <https://doi.org/10.1016/j.ijheh.2021.113692>.
- Bhatt, A., Arora, P., Prajapati, S.K., 2020. Occurrence, fates and potential treatment approaches for removal of viruses from wastewater: A review with emphasis on SARS-CoV-2. *Journal of Environmental Chemical Engineering* 8, 104429.
- Bibby, K., Peccia, J., 2013. Identification of viral pathogen diversity in sewage sludge by metagenome analysis. *Environ. Sci. Technol.* 47, 1945–1951.
- Bibby, K., Viau, E., Peccia, J., 2011. Viral metagenome analysis to guide human pathogen monitoring in environmental samples. *Lett Appl Microbiol* 52, 386–392.
- Bogler, A., Packman, A., Furman, A., Gross, A., Kushmaro, A., Ronen, A., Dagot, C., Hill, C., Vaitzel-Ohayon, D., Morgenroth, E., Bertuzzo, E., Wells, G., Kiperwas, H.R., Horn, H., Negev, I., Zuckler, I., Bar-Or, I., Moran-Gilad, J., Balcazar, J.L., Bibby, K., Elimelech, M., Weisbrod, N., Nir, O., Sued, O., Gillor, O., Alvarez, P.J., Cramer, S., Aron, S., Walker, S., Yaron, S., Nguyen, T.H., Berchenko, Y., Hu, Y., Ronen, Z., Bar-Zeev, E., 2020. Rethinking wastewater risks and monitoring in light of the COVID-19 pandemic. *Nat. Sustain.* <https://doi.org/10.1038/s41893-020-00665-2>.
- Chavarría-Miró, G., Anfruns Estrada, E., Martínez-Velázquez, A., Vázquez-Portero, M., Guix, S., Paraira, M., Galofré, B., Sánchez, G., Pinao, R., Bosch, A., 2021. Time evolution of SARS-CoV-2 in wastewater during the first 2 pandemic wave of COVID-19 in the metropolitan area of Barcelona. *Applied Environmental Microbiology*. <https://doi.org/10.1128/AEM.02750-20>. Online.
- Collivignarelli, M.C., Collivignarelli, C., Carnevale Miino, M., Abbà, A., Pedrazzani, R., Bertanza, G., 2020. SARS-CoV-2 in sewer systems and connected facilities. *Process Saf. Environ. Prot.* 143, 196–203. <https://doi.org/10.1016/j.psep.2020.06.049>.
- Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K.W., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., Mulders, D.G.J.C., Haagmans, B.L., Van Der Veer, B., Van Den Brink, S., Wijsman, L., Godekski, G., Ronette, J.L., Ellis, J., Zambon, M., Peiris, M., Goossens, H., Reusken, C., Koopmans, M.P.G., Drosten, C., 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 25, 1–8. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.
- CGA-ENAC-PPI. Entidad Nacional de Acreditación. 2020. Organismos de Control de Instalaciones: requisitos de competencia técnica.
- Haramoto, E., Malla, B., Thakali, O., Kitajima, M., 2020. First environmental surveillance for the presence of SARS-CoV-2 RNA in wastewater and river water in Japan. *Sci. Total Environ.* 737, 140405 <https://doi.org/10.1016/j.scitotenv.2020.140405>.
- Hart, O., Halden, R., 2020. Computational analysis of SARS-CoV-2/COVID-19 surveillance by wastewater-based epidemiology locally and globally: Feasibility, economy, opportunities and challenges. *Science of the Total Environment* 730.
- Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timonedá, N., Lukjancenko, O., Seidel, M., Elsäßer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLoS One* 12, 1–17. <https://doi.org/10.1371/journal.pone.0170199>.
- ISO 13528 (2015) Statistical methods for use in proficiency testing by interlaboratory comparisons.
- ISO 22174: 2005. Microbiology of food and animal feeding stuffs - Polymerase chain reaction (PCR) for the detection of food-borne pathogens - General requirements and definitions.
- Jones, D., Baluja, M.Q., Graham, D.W., Corbishley, A., McDonald, J.E., Malham, S.K., Hillary, L.S., Connor, T.R., Gaze, W.H., Moura, I.B., Wilcox, M.H., Farkas, K., 2020. Shedding of SARS-CoV-2 in feces and urine and its potential role in person-to-person transmission and the environment-based spread of COVID-19. *Sci. Total Environ.* 749, 141364 <https://doi.org/10.1016/j.scitotenv.2020.141364>.
- Kang, M., Wei, J., Yuan, J., Guo, J., Zhang, Y., Hang, J., Qu, Y., Qian, H., Zhuang, Y., Chen, X., Peng, X., Shi, T., Wang, J., Wu, J., Song, T., He, J., Li, Y., Zhong, N., 2020. Probable Evidence of Fecal Aerosol Transmission of SARS-CoV-2 in a High-Rise Building. *Ann. Intern. Med.* 1–8 <https://doi.org/10.7326/m20-0928>.
- Kumar, M., Patel, A.K., Shah, A.V., Raval, J., Rajpara, N., Joshi, M., Joshi, C.G., 2020. First proof of the capability of wastewater surveillance for COVID-19 in India through detection of genetic material of SARS-CoV-2. *Sci. Total Environ.* 746, 141326 <https://doi.org/10.1016/j.scitotenv.2020.141326>.

- La Rosa, G., Iaconelli, M., Mancini, P., Bonanno Ferraro, G., Veneri, C., Bonadonna, L., Lucentini, L., Suffredini, E., 2020. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Sci. Total Environ.* 736, 139652 <https://doi.org/10.1016/j.scitotenv.2020.139652>.
- Lesimple, A., Jasim, S.Y., Johnson, D.J., Hilal, N., 2020. The role of wastewater treatment plants as tools for SARS-CoV-2 early detection and removal. *J. Water Process Eng.* 38, 101544 <https://doi.org/10.1016/j.jwpe.2020.101544>.
- Li, X., Zhang, S., Shi, J., Luy, S.P., Jiang, G., 2021. Uncertainties in estimating SARS-CoV-2 prevalence by wastewater-based epidemiology. *Chemical Engineering Journal* 415, 1. <https://doi.org/10.1016/j.cej.2021.129039>.
- Lodder, W., De Roda-Husman, A.M., 2020. SARS-CoV-2 in wastewater: potential health risk, but also data source. *Lancet Gastroenterol. Hepatol.* 5, 533–534. [https://doi.org/10.1016/S2468-1253\(20\)30087-X](https://doi.org/10.1016/S2468-1253(20)30087-X).
- Lu, R., Zhao, X., Li, J., Niu, P., Yang, B., Wu, H., Wang, W., Song, H., Huang, B., Zhu, N., Bi, Y., Ma, X., Zhang, F., Wang, L., Hu, T., Zhou, H., Hu, Z., Zhou, W., Zhao, L., Chen, J., Meng, Y., Wang, J., Lin, Y., Yuan, J., Xie, Z., Ma, J., Liu, W.J., Wang, D., Xu, W., Holmes, E.C., Gao, G.F., Wu, G., Chen, W., Shi, W., Tan, W., 2020. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* 2020, 19–21.
- Medema, G., Heijnen, L., Elsinga, G., Italiaander, R., Brouwer, A., 2020. Presence of SARS-Coronavirus 2 RNA in Sewage and Correlation with Reported COVID-19 Prevalence in the Early Stage of the Epidemic in The Netherlands. *Environ. Sci. Technol. Lett.* <https://doi.org/10.1021/acs.estlett.0c00357>.
- Murtula, R., Soria, E., Yañez, M.A., Catalán, V., 2012. Proficiency testing schemes for the assessment of *Legionella* PCR methodologies. *Accred Qual Assur.* <https://doi.org/10.1007/s00769-012-0903-5>.
- Nemudryi, A., Nemudraia, A., Wiegand, T., Surya, K., Buyukyoruk, M., Vanderwood, K., Wilkinson, R., Wiedenheft, B., 2020. Temporal Detection and Phylogenetic Assessment of SARS-CoV-2 in Municipal Wastewater. *Cell Reports Med.* 100098. <https://doi.org/10.2139/ssrn.3664367>.
- Oakley, S. and Mihelcic, J.R. 2019. Pathogen Reduction and Survival in Complete Treatment Works. In: J.B. Rose and B. Jiménez-Cisneros, (eds) *Global Water Pathogen Project*. <http://www.waterpathogens.org> (J.R. Mihelcic and M.E. Verbyla) (eds) Part 4 Management Of Risk from Excreta and Wastewater) <http://www.waterpathogens.org/book/pathogen-reduction-and-survival-in-complete-treatment-works> Michigan State University, E. Lansing, MI, UNESCO. <https://doi.org/10.14321/waterpathogens.49>.
- Park, S.-K., Lee, C.-W., Park, D.-L., Woo, H.-Y., Cheong, H.S., Shin, C.S., Ahn, K., Kwon, M.-J., Joo, E.-J., 2020. Detection of SARS-CoV-2 in Fecal Samples From Patients With Asymptomatic and Mild COVID-19 in Korea. *Clin. Gastroenterol. Hepatol.* In press.
- Peccia, J., Zulli, A., Brackney, D.E., Grubaugh, N.D., Kaplan, E.H., Casanovas-massana, A., Ko, A.I., Malik, A.A., Wang, D., Wang, M., Warren, J.L., Weinberger, D. M., Arnold, W., Omer, S.B., 2020. Measurement of SARS-CoV-2 RNA in wastewater tracks community infection dynamics. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-020-0684-z>.
- Randazzo, W., Truchado, P., Cuevas-Ferrando, E., Simón, P., Allende, A., Sánchez, G., 2020. SARS-CoV-2 RNA in wastewater anticipated COVID-19 occurrence in a low prevalence area. *Water Res* 181. <https://doi.org/10.1016/j.watres.2020.115942>.
- Sherchan, S.P., Shahin, S., Ward, L.M., Tandukar, S., Aw, T.G., Schmitz, B., Ahmed, W., Kitajima, M., 2020. First detection of SARS-CoV-2 RNA in wastewater in North America: A study in Louisiana. *USA. Sci. Total Environ.* 15, 140621.
- van Doorn, A.S., Meijer, B., Frampton, C.M.A., Barclay, M.L., de Boer, N.K.H., 2020. Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Aliment. Pharmacol. Ther.* 52, 1276–1288. <https://doi.org/10.1111/apt.16036>.
- Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., Tan, W., 2020. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA - J. Am. Med. Assoc.* 323, 1061–1069. <https://doi.org/10.1001/jama.2020.1585>.
- Weidhaas, Jennifer, Aanderud, Zachary T., Roper, D. Keith, VanDerlice, James, Brown Gaddis, Erica, Ostermiller, Jeff, Hoffman, Ken, Jamal, Rubayat, Heck, Phillip, Zhang, Yue, Torgersen, Kevin, Vander Laan, Jacob, LaCross, Nathan, 2021. Correlation of SARS-CoV-2 RNA in wastewater with COVID-19 disease burden in sewersheds. *Science of The Total Environment* 775. <https://doi.org/10.1016/j.scitotenv.2021.145790>.
- Westhaus, S., Weber, F.A., Schiwy, S., Linnemann, V., Brinkmann, M., Wiedera, M., Greve, C., Janke, A., Hollert, H., Wintgens, T., Ciesek, S., 2021. Detection of SARS-CoV-2 in raw and treated wastewater in Germany – Suitability for COVID-19 surveillance and potential transmission risks. *Sci. Total Environ.* 751, 141750 <https://doi.org/10.1016/j.scitotenv.2020.141750>.
- Wiley, John S., Chin, Tom D. Y., Gravelle, Clifton R., Robinson, Sherry, 1962. Enterovirus in Sewage during a Poliomyelitis Epidemic. *Water Pollution Control Federation* 34, 168–178.
- Withey, S., Cartmell, E., Avery, L.M., Stephenson, T., 2005. Bacteriophages–potential for application in wastewater treatment processes. *Sci. Total Environ.* 339, 1–18. <https://doi.org/10.1016/j.scitotenv.2004.09.021>.
- Wolf, S., Hewitt, J., Greening, G.E., 2010. Viral multiplex quantitative PCR assays for tracking sources of fecal contamination. *Appl. Environ. Microbiol.* 76, 1388–1394. <https://doi.org/10.1128/AEM.02249-09>.
- Wölfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Müller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., Hoelscher, M., Bleicker, T., Brünink, S., Schneider, J., Ehmann, R., Zwirgmaier, K., Drosten, C., Wendtner, C., 2020. Virological assessment of hospitalized patients with COVID-2019. *Nature* 581, 465–469. <https://doi.org/10.1038/s41586-020-2196-x>.
- Wurtzer, S., Marechal, V., Mouchel, J.M., Maday, Y., Teyssou, R., Richard, E., Almayrac, J.L., Moulin, L., 2020. Evaluation of lockdown effect on SARS-CoV-2 dynamics through viral genome quantification in waste water, Greater Paris, France, 5 March to 23 April 2020. *EuroSurveill* 25 (50). <https://doi.org/10.2807/1560-7917.ES.2020.25.50.2000776> pii=2000776.
- Xiao, F., Sun, J., Xu, Y., Li, F., Huang, X., Li, H., Zhao, J., Huang, J., 2020. Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerg. Infect. Dis.* 26 (8), 1920–1922.
- Yuan, J., Chen, Z., Gong, C., Liu, H., Li, B., Li, K., Chen, X., Xu, C., Jing, Q., Liu, G., Qin, P., Liu, Y., Zhong, Y., Huang, L., Zhu, B.-P., Yang, Z., 2020. Coronavirus Disease 2019 Outbreak Likely Caused by Sewage Exposure in a Low-income Community: Guangzhou, China, April 2020. *Lancet*.
- Zahedi, A., Monis, P., Deere, D., Ryan, U., 2020. Wastewater-based epidemiology – surveillance of waterborne pathogens with a focus on SARS-CoV-2, Cryptosporidium and Giardia. *Parasitology Research.* <https://doi.org/10.1007/s00436-020-07023-5>.
- Zang, R., Gomez Castro, M.F., McCune, B.T., Zeng, Q., Rothlauf, P.W., Sonnek, N.M., Liu, Z., Brulois, K.F., Wang, X., Greenberg, H.B., Diamond, M.S., Ciorba, M.A., Whelan, S.P.J., Ding, S., 2020. TMPRSS2 and TMPRSS4 promote SARS-CoV-2 infection of human small intestinal enterocytes. *Sci. Immunol.* 5 (47) eabc3582.

Supporting Information

Elimination of SARS-CoV-2 along wastewater and sludge treatment processes

Albert Serra-Compte^{1*}; Susana González¹; Marina Arnaldos¹; Sabrina Berlendis²; Sophie Courtois²; Jean Francois Loret²; Olivier Schlosser²; Adela M. Yáñez³; Elena Soria-Soria³; Mariana Fittipaldi³; Gemma Saucedo⁴; Anna Pinar-Méndez⁴; Miquel Paraira⁴; Belén Galofré⁴; Juan M. Lema⁵; Sabela Balboa⁵; Miguel Mauricio-Iglesias⁵; Albert Bosch⁶; Rosa M. Pintó⁶; Isabelle Bertrand⁷; Christophe Gantzer⁷; Carlos Montero¹; Xavier Litrico⁸

¹Cetaqua, Water Technology Centre, Cornellà de Llobregat, Spain

²Suez, Centre International de Recherche sur l'Eau et l'Environnement (CIRSEE), Le Pecq, France

³Labaqua, Alicante, Spain

⁴Aigües de Barcelona, Barcelona, Spain

⁵CRETUS Institute, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

⁶University of Barcelona, Barcelona, Spain

⁷University of Lorraine Lorraine, CNRS, LCPME, Nancy, France

⁸Suez Groupe, Paris, France

Material and Methods

Sample processing

Wastewater and liquid sludge samples obtained from French WWTPs were processed according to Bertrand et al. (2021). An ultrafiltration procedure was performed with a Centricon Plus-70 centrifugal ultrafilter with a cut-off of 100 kDa (Merck Millipore). The ultrafilter was washed with 60 ml deionized water by centrifugation at 1,500 x g for 15 min to remove trace amounts of glycine in accordance with the manufacturer's recommendations. Two volumes of 50 mL wastewater were both filtered by centrifugation at 1,500 x g for 15 min. After each centrifugation step, the concentrate was recovered by inverting the system and applying centrifugation (1,000 x g for 2 min). The resulting concentrate's volume was around 1.5 mL. A first washing step of the ultrafilter was done with 3.5 mL of deionized water and two additional washing steps were performed two times 5 min with 5 mL of lysis buffer. The 15 mL resulting lysis solution and sample was then purified with 15 mL phenol-chloroform-isoamyl alcohol (25:24:1 pH 7.8-8.2).

For water sampling from the Spanish WWTPs (except Ourense), 500 mL of water were centrifuged at 8,000 x g for 30 min previous addition of 62.5 mL of glycine buffer (0.05 M glycine, 3% beef extract, pH 9.6). Then, supernatant was filtered through 0.45 µm Polyethersulfone (PES) membrane filters. Subsequently, PEG 8000 (80 g/L) and NaCl (17.5 g/L) were added and incubated overnight at 4°C with agitation. Finally, after centrifugation at 13000 x g at 4°C during 90 min, pellet was resuspended in 1 mL of Phosphate Buffered Saline (PBS) and stored at -80°C prior RNA extraction (Hjelmsø et al., 2017). In each set, a negative and positive sample were included as quality control of the process. The positive sample was prepared by adding non-infectious and replication deficient SARS-CoV-2 reference material (Accuplex™ SARS-CoV-2-SERACARE 0505-0133). The water samples obtained from Ourense WWTP were concentrated by ultrafiltration (Balboa et al., 2020). Briefly, 100 mL were gently centrifuged to remove large particles at 4600 x g during 30 min. Supernatants obtained were further concentrated by Amicon 15 mL 10 K centrifugal devices. A last step of buffer exchange was performed using phosphate buffer saline (PBS) pH 7.4. For solid sludges, the procedure was as follows: 180 mL of glycine/beef extract elute buffer were added to the wet sludge sample (wet sample amount equivalent to 6 g of dry weight), with a subsequent stir for 2h at 4°C. Then, the mixture was centrifuged at 8,000 x g for 30 min and the supernatant was filtered through 0.45µm PES membrane filters. After filtration, PEG 8000 (80 g/L) and NaCl (17.5 g/L) were added and incubated overnight with agitation at 4 °C, followed by a centrifugation at 13,000 x g for 90min at 4°C. The resulting pellet was resuspended in 1 mL of PBS and stored at -80 °C until RNA extraction. Regarding Ourense WWTP, sludge samples were analyzed according to Balboa et al., 2020.

Table S1. Wastewater treatment plant characteristics and sampling point

WWTP Name	Country	P.E.	Secondary treatment	Tertiary treatment	Sludge treatment	Sampling points
WWTP1	Spain	2.800.000	Activated sludge plus clarifier		Thickening + dewatering	Influent, effluent, mixed sludge, thickened sludge
WWTP2	Spain	2.000.000	Activated sludge with nutrient removal plus clarifier	UV + Chlorination	Thickening+ anaerobic digestion + dewatering	Influent, effluent, mixed sludge, thickened sludge, digested sludge
WWTP3	Spain	300.000	Membrane Batch Reactor	Ultrafiltration	Thickening+ anaerobic digestion + dewatering	Effluent, tertiary effluent
WWTP4	Spain	200.000	Membrane Batch Reactor	Chlorination	Thickening+ anaerobic digestion + dewatering	Influent, effluent, tertiary effluent, thickened sludge, digested sludge
WWTP5	Spain	306.592	Activated sludge plus clarifier	Microfiltration/ Ultrafiltration	Thickening+ anaerobic digestion + dewatering	Influent, effluent, tertiary effluent
WWTP6	Spain	204.089	Activated sludge with nutrient removal plus clarifier	Filtration plus UV	Thickening+ anaerobic digestion + dewatering	Influent, effluent, tertiary effluent, mixed sludge, thickened sludge, digested sludge
WWTP7	Spain	833.000	Activated sludge with nutrient removal plus clarifier	Chlorination	Thickening+ anaerobic digestion + dewatering	Influent, effluent, mixed sludge, secondary sludge
WWTP8	Spain	350.000	Activated sludge with nutrient removal plus clarifier	Microfiltration	Thickening+ anaerobic digestion + dewatering	Influent, primary effluent, effluent, primary sludge, secondary sludge, thickened sludge, digested sludge
WWTP9	France	1.000.000	Activated sludge plus clarifier		Thickening + anaerobic digestion + centrifugation	Influent, clarified water, effluent, thickened sludge, primary sludge, digested sludge
WWTP10	France	400.000	Activated sludge plus clarifier		Thickening + Centrifugation	Influent, effluent, activated sludge
WWTP11	France	300.000	Activated sludge plus clarifier		Thickening + anaerobic digestion + centrifugation	Influent, digested sludge, activated sludge
WWTP12	France	600.000	Activated sludge plus clarifier		Flotation + centrifugation	Influent, effluent,
WWTP13	France	180.000	Biofiltration		Anaerobic digestion + centrifugation	
WWTP14	France	950.000	Activated sludge high load and biofiltration		Centrifugation + incineration	Influent, effluent, activated sludge
WWTP15	France	650.000	Activated sludge high load plus clarifier		Thickening + pre-heating plus filter press	Influent, effluent, secondary sludge
WWTP16	France	250.000	Activated sludge plus clarifier MBR		Thickening + anaerobic digestion plus centrifugation	Influent, effluent, activated sludge, MBR sludge

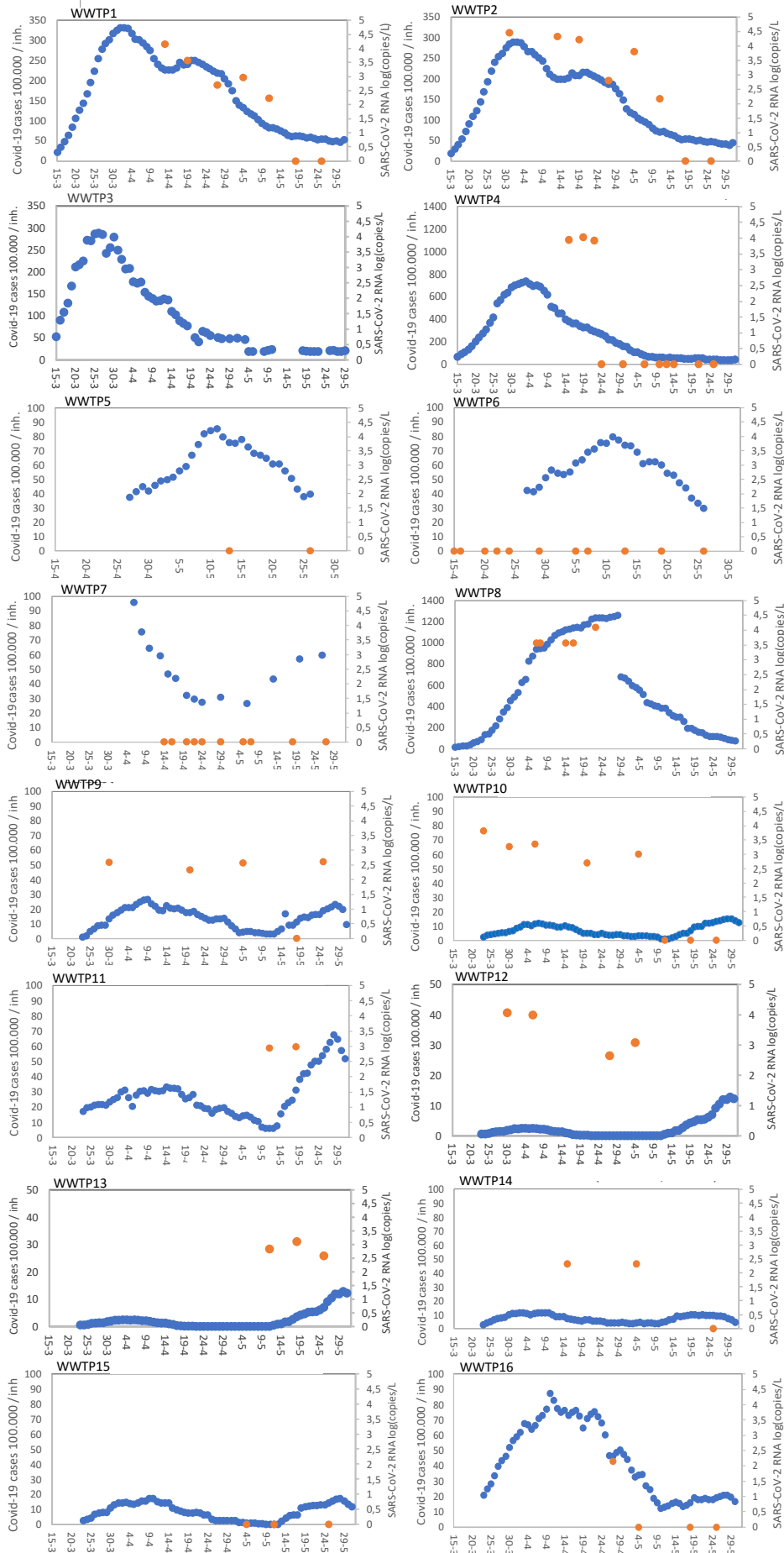


Figure S1. SARS-CoV-2 quantification in the influent of the WWTPs monitored (orange dots) and COVID-19 reported cases in the corresponding area (blue dots) during sampling time. COVID-19 cases are reported per municipality in Spanish WWTPs and at department level for French plants.

Table S2. SARS-CoV-2 concentration mean \pm standard deviation for the different water and sludge treatment steps (cg/L log-transformed values). Log removal (mean \pm standard deviation) is provided for each treatment step.

	Influent concentration (log (copies/L))*	Effluent concentration (log (copies/L))*	log removal
Water line			
Primary treatment	3.46 \pm 0.6	2.98 \pm 1.36	0.48 \pm 1.17
Secondary treatment	3.66 \pm 0.7	2.1 \pm 0.3	1.5 \pm 0.8
Tertiary treatment	3.97 \pm 0.08	<2	>1.97 \pm 0.08
Water treatment technologies			
Activated sludge	3.29 \pm 0.67	2.26 \pm 0.47	1.03 \pm 0.59
Activated sludge plus nutrient removal	3.65 \pm 0.68	2.28 \pm 0.70	1.37 \pm 0.72
MBR	3.89 \pm 0.89	2.13 \pm 0.35	1.96 \pm 0.93
Sludge line			
Thickened sludge	3.27 \pm 0.68	3.74 \pm 1.42	-0.47 \pm 1.20
Digested sludge	3.25 \pm 0.86	3.41 \pm 1.50	-0.16 \pm 0.89
Digested sludge plus thermal hydrolysis	3.69 \pm 0.27	<2	>1.69 \pm 0.27

*For solid sludge samples the concentration is log (copies/Kg)

Table S3 SARS-CoV-2 and F-specific bacteriophages concentration in Log (copies/L) at different steps of the water treatment process (mean \pm standard deviation, log values).

	SARS-CoV-2 concentration log (copies/L)	F-specific bacteriophages concentration log (copies/L)
Influent	4.26 \pm 0.56	7.83 \pm 0.26
Activated sludge	3.23 \pm 0.46	4.19 \pm 1.90
Effluent (after clarification)	n.d.	2.94 \pm 1.52

n.d. not detected

References

- Balboa, S., Mauricio-Iglesias, M., Rodriguez, S., Martínez-Lamas, L., Vasallo, F.J., Regueiro, B., Lema, J.M., 2020. The fate of SARS-CoV-2 in WWTPS points out the sludge line as a suitable spot for monitoring. medRxiv 1–24.
- Bertrand, I., Challant, J., Jeulin, H., Hartard, C., Mathieu, L., Lopez, S., Interest, S., Obépine, G., Schvoerer, E., Courtois, S., Gantzer, C., 2021. Epidemiological surveillance of SARS-CoV-2 by genome quantification in wastewater applied to a city in the northeast of France: comparison of ultrafiltration- and protein precipitation-based methods. *Int. J. Hyg. Environ. Health* 113692. <https://doi.org/10.1016/j.ijheh.2021.113692>
- Hjelmsø, M.H., Hellmér, M., Fernandez-Cassi, X., Timoneda, N., Lukjancenka, O., Seidel, M., Elsässer, D., Aarestrup, F.M., Löfström, C., Bofill-Mas, S., Abril, J.F., Girones, R., Schultz, A.C., 2017. Evaluation of methods for the concentration and extraction of viruses from sewage in the context of metagenomic sequencing. *PLoS One* 12, 1–17. <https://doi.org/10.1371/journal.pone.0170199>