



UNIVERSITAT DE
BARCELONA

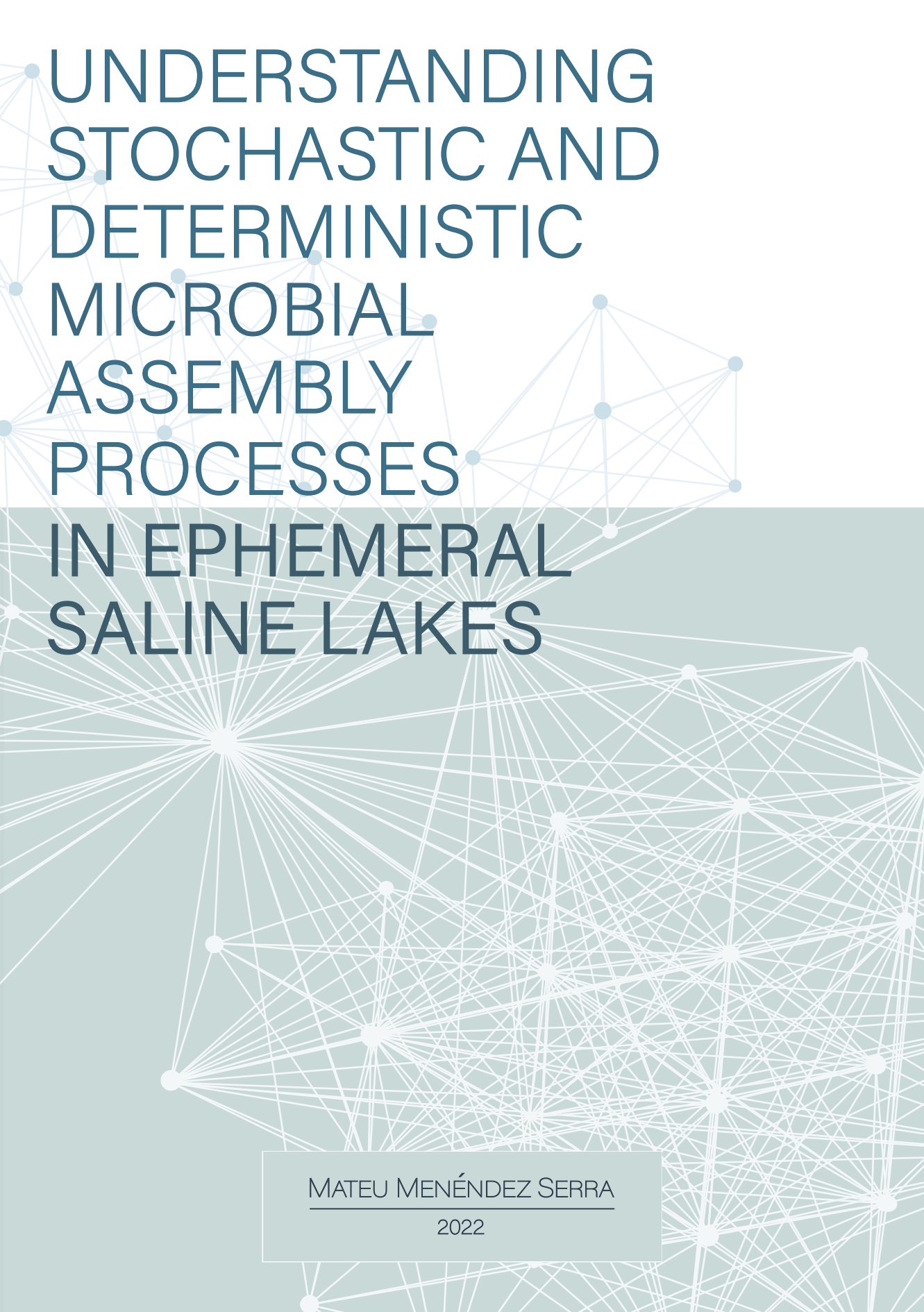
Understanding stochastic and deterministic microbial assembly processes in ephemeral saline lakes

Mateu Menéndez Serra

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



UNDERSTANDING STOCHASTIC AND DETERMINISTIC MICROBIAL ASSEMBLY PROCESSES IN EPHEMERAL SALINE LAKES

MATEU MENÉNDEZ SERRA

2022



UNIVERSITAT_{DE}
BARCELONA

Understanding stochastic and deterministic
microbial assembly processes in
ephemeral saline lakes

Mateu Menéndez Serra

2022

Menéndez-Serra, M, 2022. Understanding stochastic and deterministic microbial assembly processes in ephemeral saline lakes. PhD thesis. Universitat de Barcelona. 321 p.

Tesis Doctoral



Facultat de Biologia, Universitat de Barcelona
Programa de Doctorat d'Ecologia, Ciències Ambientals i Fisiologia Vegetal

Understanding stochastic and deterministic microbial assembly processes in ephemeral saline lakes

Memòria presentada per Mateu Menéndez Serra per obtenir al
Grau de Doctor per la Universitat de Barcelona

Mateu Menéndez Serra
Departament d'Ecologia integrativa d'aigües continentals.
Centre d'Estudis Avançats de Blanes (CEAB)
Consejo Superior de Investigaciones Científicas (CSIC)
Març 2022

Director de tesi
Dr. Emilio Ortega
Casamayor
Centre d'Estudis
Avançats de Blanes
(CEAB)

Tutora de tesi
Dr. Isabel Muñoz
Gracia
Universitat de Barcelona

Co-Director de tesi
Dr. David Alonso
Giménez
Centre d'Estudis
Avançats de Blanes
(CEAB)

*A madre, padre, Pau,
a Elena,
a los amigos de siempre,
i als que aquesta tesi m'ha permès conèixer.*

Agradecimientos i agraïments

Casi cinco años en esta burbuja de las tesis doctorales es tiempo más que suficiente para saber que esta sección será sin duda la más leída. Y por eso mismo intentaré no extenderme mucho, para que os queden ganas de leer ni que sea el resumen.

Como es tradición, empiezo dando las gracias al director de esta tesis, Emilio. En estos más de cuatro años creo que me has dado un buen equilibrio de tutela y libertad que me ha permitido explorar mis propios caminos sin dejar de sentirme respaldado y bien guiado. De ti he aprendido mucho del funcionamiento de dos sistemas complejos: las comunidades microbianas y el mundo de la ciencia (y su a-veces-complejo-a-veces-inquietante funcionamiento interno). En definitiva, gracias por darme la oportunidad de hacer esta tesis y todo lo que ello ha conllevado y seguro conllevará. També al co-director d'aquesta tesi, David, qui m'ha permes ficar el cap (sempre desde la vorera i amb manguitos) a un mon de models, *cuentas* i teoritzacions, a priori hostile però d'inacabables possibilitats.

Com no podia ser d'altra manera, he de continuar amb els que em vau ajudar a fer els primers passos del camí que m'han portat fins aquí. Xevi, no se m'acudeixi millor persona per rebre algú a un grup i fer-lo sentir com a casa des de el primer dia. Gràcies per totes les hores que has invertit (vull pensar que no perdut) en ensenyar-me com funcionava tot, per la teva tendresa i per estar sempre disponible i disposat pel resoldre els milers de dubtes que anaven sorgint. Per mi ets co-director *de facto* d'aquesta tesi. Rudi, *a.k.a. plot brother*, tú hiciste esa labor de acogida dentro pero también fuera de los muros del CEAB. No solo me introdujiste en la comunidad ceabina, hiciste que el piso fuese un hogar desde el primer día. Vicente, compañero de despacho y compañero, en estos más de cuatro años siempre has sabido resolver las muchas dudas que me han ido surgiendo y has sido esos manguitos que mencionaba un párrafo atrás. Sin duda, tus ideas han hecho mejorar con creces esta tesis. Gerard, tot i que el nostre període de solapament va ser més breu del que m'hagués agradat, tu em vas ensenyar la màgia negra al darrere dels loops d'R. També la diferència entre gavina i

gavià, però això ja és un altre tema.

També hi ha la gent que, a part d'ensenyar-me moltes coses, ha fet que tingués ganes de pujar al CEAB cada matí. Al Joan li he dit moltes vegades –i no és broma– que jo de gran vull ser com ell. Mil gràcies per estar sempre disposat a escoltar els dubtes, ja siguin sobre el *SeqEnv* o el futur incert, aquests últims cada cop més recurrents. El fet que tu siguis una de les primeres persones amb qui vulgui parlar quan apareixen deu significar algo. La Marta sempre ha sigut el nucli vertebrador del grup. Sense ella tots hauríem begut molts menys litres de cafè i hauríem treballat moltes més hores, però no tinc cap dubte que sense ella aquests anys no haguessin sigut ni la meitat d'especials i bonics. Amb el Xavi vam arribar alhora al CEAB, i no se m'acut millor company de tesi, cafès o pàdels. Ara, quan acabi tot el que envolta aquest totxo (i que sembla no acabar mai), ens posem en serio a entrenar pel World Padel Tour. El Roger és el millor exponent del bon hipisme, per les ulleres rodones i *open source software*, però sobretot pel bon rollo que transmet en tot moment. Fins i tot quan canvien l'hora i no li senti bé, compartir estones amb ell sempre és un plaer. I fins i tot he sentit a dir que vas fer un *ironman*, això m'ho has d'explicar! Jorge es en todo momento la combinación perfecta de respetable señor Salmantino que nació con un traje puesto y la persona más cariñosas y dispuestas a ayudar que me he cruzado. El día que me des trabajo como catador de *Caulerpa* ya no se te podrá pedir más. La Jana és llum (i sobretot veu) que enlluerna (i ensordeix) allà on va. No tinc cap dubte que a hores d'ara deu ser la persona més popular de Niça, però tot i això: torna Jana, et trobem a faltar. Sempre he pensat que amb la Cris hem coincidit menys del que m'hagués agradat, però el fet que sobrevisqués a la tortura d'una de les persones més cruels que he conegut (òbviament parlo de la Marta García, a qui així i tot li tinc un carinyo especial) va fer que tingués tots els meus respectes des del primer minut. El pas d'altres persones pel CEAB ha sigut més breu, però gent com l'Arnau i el Jan estaran sempre presents quan recordi aquesta etapa. Una de les imatges més idíl·liques d'aquests anys són les cerveses prenent el sol a la coberta del *lapu-lapu*. Si no fos perquè, com diu Sabina "*al lugar donde has sido feliz no debieras tratar de volver*", em doctoraria de tres o quatre coses més si fos al vostre costat.

Gracias también a los que habéis hecho de Blanes un sitio que echar de menos al marchar, aunque muchos hayáis marchado ya. Carmen ya habitaba el piso cuando yo llegué, y aunque nos costó un tiempo echar a andar y el camino fue escarpado, para ella solo tengo cariño y buenos recuerdos. Bueno, y algunos videos cantando, así que ándate con ojo y no me hagas utilizarlos en tu contra. De la mano de Carmen venía Sergio, quien la quiere tanto que primero se vino a Blanes y ahora se ha ido Dinamarca (ojo eh, que se dice pronto... a Blanes!). Para ti también tengo nada más que cariño. En el piso también estaba Mirco, de fuerza de voluntad y bondad difícil de igualar, y más tarde llegó Montagut, a quien describiría como... indescriptible. Por Blanes siempre estuvieron Ibor y Hector, Hector e Ibor, siempre dispuestos a abarrotar su salón y a darle a play a la playlist más extraña del mundo. Y aunque muchos ya se han ido, también hay otros que llegaron más tarde y que se han convertido en indispensables. Como Laura y Mario. Con vosotros he conocido la vida del jubilado que queda "para salir a andar", compara tortillas de patatas y come pollo al ast los domingos, y no me avergüenza decir que me ha gustado. Igual la compañía ha tenido algo que ver. Y en esa compañía también estaba Marta, quien me ha convertido en el gran tenista que soy a día de hoy pero a quien no le perdono que se llevara a Lola lejos de nosotros, e Itzi, un torbellino de energía que ha venido a presidir el CSIC pero todavía no lo sabe.

En el CEAB también he coincidido, quizás de forma más fugaz o tangencial, con infinidad de gente a la que me gustaría mencionar. Por supuesto, hay que aprender de los que van un paso por delante y con su ejemplo nos muestran los caminos a seguir (habéis visto que elegante lo he hecho, que ni he mencionado vuestra avanzada edad?): Eneko, Ribot, María, Sara, Candela (y Merete!), Jordi Boada i Jordi Pagés... y en general a todos los que han compartido alguna o todas las etapas de este camino: Nerea, Célia, Adrià, Pol, Noemí, Martina, Dani, Capi, Gustavo, Xavi Torras, Sara, Mar, Anna, el maravilloso equipo de administración, los valientes IPs que hicieron sus pinitos en el mundo del pádel, los miembros del grupo de ecología teórica... (y tú, persona que está leyendo esto y cuyo nombre he olvidado incluir porque mi memoria es un desastre).

Gràcies Albert pel temps que em vas acollir a la University of Arizona

i a Tucson. De tu he après moltíssim de ciència, però també del món que l'envolta i moltes altres coses vàries. Espero que els camins es tornin a creuar. And also, of course, to all the amazing people of your lab, with a special mention to Mery, Yongyian and Gabriele.

Por último, y pese a ello quizás lo más importante, a la gente externa a este mundillo que es *la siensia*. Aunque pasen los años y los encuentros se vayan limitando a bodas y eventos importantes, los compañeros ambientalólogos siempre están ahí. Como ahí están los amigos y amigas que siguen en el pedazo de roca que sobresale del mediterráneo y a quienes veo menos de lo que me gustaría. Pero si hay alguien que siempre ha estado, es la familia. La cercana, la de *madre, padre* y Pau. Y quizás sea porque yo soy el pequeño y por tanto el último en llegar, pero el caso es que no recuerdo momento en que no estuviesen, y eso es importante. Lo mas importante, me atrevería a decir. Elena apareció poco después de que empezase esta etapa y, por lo tanto, hemos hecho juntos todo este camino. Ella ha hecho que los poquísimos momentos malos lo fuesen menos, y ha insistido en que los buenos podían mejorar con una cena en el Akuma. Con ella llegó una familia numerosa que me acogió a base de cariño y tupperes –incluso hay alguien muy especial que ahora me llama *tito*– así como un grupo de amigos que abrió su hermético círculo para incluirme. Se vino a vivir a Blanes, y si eso no es amor yo ya no entiendo nada. Ahora dice que se quiere venir allá donde me lleve el camino, y ahí sí que ya no me cabe ninguna duda. Con todo esto, solo puedo decirte que lo del amor es mutuo, y que gracias.

Informe dels directors

El Dr. Emilio Ortega Casamayor i el Dr. David Alonso Giménez, directors de la Tesi Doctoral elaborada pel candidat Mateu Menéndez Serra i que porta per títol “Understanding stochastic and deterministic microbial assembly processes in ephemeral saline lakes”,

INFORMEN

Que els treballs de recerca duts a terme per Mateu Menéndez Serra com a part de la seva formació pre-doctoral i inclosos a la seva Tesi Doctoral han donat lloc a sis capítols, tres dels quals ja estàn publicats i tres que es troben en process de ser enviats a revistes d'àmbit internacional. A continuació es detalla la llista d'articles, així com els índex d'impacte (segons la ISI Web of Science) de les revistes on han estat publicats.

1. **Mateu Menéndez-Serra**, Vicente J. Ontiveros, Xavier Triadó-Margarit, David Alonso and Emilio O. Casamayor (2020) Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain). *FEMS Microbiology Ecology*, 96, 2020, faa019. DOI:10.1093/femsec/faa019. Microbiology (Rank 50/137) **Q2. Impact Factor: 4.19**
2. **Mateu Menéndez-Serra**, Xavier Triadó-Margarit, Emilio O. Casamayor (2021) Ecological and metabolic thresholds in the bacterial, protist and fungal microbiome of ephemeral saline lakes (Monegros Desert, Spain) *Microbial Ecology*, DOI:10.1007/s00248-021-01732-9. Ecology (Rank 33/166) **Q1. Impact Factor: 4.55**
3. Xavier Triadó-Margarit, José A. Capitán, **Mateu Menéndez-Serra**, Rudiger Ortíz-Álvarez, Vicente J. Ontiveros, Emilio O Casamayor and David Alonso (2019) A Randomized Trait Community Clustering approach to unveil consistent environmental threshold in community assembly. *ISME J*, 13, 2681-2689, DOI:10.1038/s41396-019-0454-4. Ecology (Rank 5/166) **Q1. Impact Factor: 9.180**

Addicionalment, es preveu la publicació dels articles corresponents a l'estudi dels processos d'assemblatge de les comunitats, l'anàlisi de les interaccions biològiques, i les possibles relacions entre les comunitats aquàtica i terrestre a les llacunes estudiades, corresponents als capítols cinquè, sisè i setè, respectivament:

4. **Mateu Menéndez-Serra**, Vicente J. Ontiveros, Joan Cáliz, David Alonso and Emilio O. Casamayor. Understanding stochastic and deterministic assembly processes in microbial communities along temporal and environmental gradients. Enviat per la seva publicació a *Molecular Ecology*. Ecology (Rank 16/166) **Q1. Impact Factor: 6.185**
5. **Mateu Menéndez-Serra**, Vicente J. Ontiveros, Albert Barberán, Emilio O. Casamayor. Absence of stress-promoted facilitation coupled with a competition decrease in the microbiome of ephemeral saline lakes. Enviat per la seva publicació a *Ecology*. Ecology (Rank 21/166) **Q1. Impact Factor: 5.499**
6. **Mateu Menéndez-Serra**, Joan Cáliz, Xavier Triadó-Margarit, Emilio O. Casamayor. Increasing saline stress reduces allochthonous microbial colonization in ephemeral ponds, driving the transition from habitat generalists to specialists. En procés de preparació per ser enviat a una revista encara per determinar.

Alhora CERTIFIQUEN

Que Mateu Menéndez Serra ha participat activament en el desenvolupament del treball de recerca associat a cadascun d'aquests articles, així com en la seva elaboració. En concret, la seva participació en cadascun dels articles ha estat la següent:

- Plantejament dels objectius.
- Processat i anàlisi de mostres, a més dels anàlisis de dades obtingudes.
- Redacció dels articles i seguiment del procés de revisió dels mateixos.

Finalment, certifiquen que cap dels coautors dels articles presentats a continuació i que formen part de la Tesi Doctoral de Mateu Menéndez Serra utilitzarà implícitament o explícitament aquests treballs per a l'elaboració d'una Tesi Doctoral.

ORTEGA
CASAMAYOR
EMILIO - DNI
39172805X

Firmado digitalmente
por ORTEGA
CASAMAYOR EMILIO -
DNI 39172805X
Fecha: 2022.03.23
14:02:23 +01'00'

Firmado por ALONSO GIMENEZ DAVID
- DNI 36981503X el día 23/03/2022
con un certificado emitido por AC
Sector Público

Director de tesi
Dr. Emilio Ortega Casamayor
Centre d'Estudis Avançats de Blanes
(CEAB)

Co-Director de tesi
Dr. David Alonso Giménez
Centre d'Estudis Avançats de Blanes
(CEAB)

Abstract

Understanding the mechanisms and processes driving biological communities assembly, i.e. determining why a specific community is found under certain circumstances, is one of the fundamental questions in ecology. The potential effect of the different assembly processes (selection, dispersal and ecological drift, together with speciation which operates at evolutionary timescales) has been studied over multiple systems and types of organisms, proposing a complex equilibrium with a variable dominance of either stochastic or deterministic assembly processes. In this PhD thesis, we aimed to reach a deep understanding on the processes and mechanisms driving microbial communities assembly using a highly dynamic system of ephemeral saline lakes potentially interconnected by dispersal processes as study model. Ephemeral saline lakes are a perfect scenario to assess the effect of variable selective pressure (promoted by wide salinity fluctuations) and the potential effect on complementary assembly processes as dispersal and ecological drift. The study of these processes across environmental, spatial and temporal scales may offer a mechanistic understanding on how assembly processes interact to determine communities diversity.

Along three consecutive hydrological cycles, we characterized the monthly dynamics of the aquatic microbiome inhabiting 14 ephemeral salt lakes in the Monegros Desert (NE Spain). These systems experienced severe variations in water volume along the hydrological cycle, which promote wide salinity fluctuations from freshwater to salt-saturation. The inhabiting bacterial, archaeal and microeukaryal communities were characterized by high-throughput sequencing of the 16S and 18S rRNA genes, respectively. We assessed the effect of increasing salinity concentrations in the diversity and composition of microbial communities, together with ecological and metabolic effects. Next, we analyzed the ephemeral lakes from a metacommunity perspective to investigate the role of the different assembly processes along the hydrological cycle and their interactions to shape the observed dynamics.

As expected, we noticed that increasing salinity gradients strongly determined microbial communities compositions. While a taxonomic replacement was observed in the bacterial communities driven by mass extinctions followed by colonization by better-adapted halophilic groups, eukaryal communities suffered a severe environmental filtering which produced a simplified assembly dominated by a few halophilic groups, mainly the photosynthetic algae *Dunaliella* and *Navicula*. Conversely, the rich archaeal assemblage of Monegros become more abundant as lakes reached hypersaline concentrations. Through a novel genomic trait-based approach we set up the transition from stochasticity to deterministic assembly describing a consistent threshold at salinities c.a. 5%. Beyond this threshold, increasing selective pressure reduced the potential colonization success among lakes, but also the potential colonization by allochthonous microbes from other biomes as sediments and soils. Potential interspecies interactions along the salinity gradient were addressed by a network analysis that showed that increasing selective pressure, and the associated transition from generalists to specialists groups, reduced competition and did not promote facilitative associations. Overall, the multiscale assessment of the deterministic and stochastic assembly processes in the Monegros ephemeral lacustrine system, together with the previous vast knowledge on both the ecology and microbiology of saline system and the physiological effects of salinity gradients, allowed us to describe the biological assembly from an unusual combination of mechanistic and integrative points of view.

Resumen

Comprender los mecanismos y los procesos que impulsan el ensamblaje de las comunidades biológicas, es decir, determinar por qué se encuentra una comunidad específica en determinadas circunstancias, es una de las cuestiones fundamentales en ecología. El efecto potencial de los diferentes procesos de ensamblaje (selección, dispersión y deriva ecológica, junto con la especiación que opera a escalas de tiempo evolutivas) se ha estudiado en múltiples sistemas y organismos, proponiendo un equilibrio complejo con un predominio variable de los procesos de ensamblaje estocásticos y deterministas. En esta tesis doctoral, nuestro objetivo es comprender los procesos y mecanismos que impulsan el ensamblaje de las comunidades microbianas utilizando un sistema de lagos salinos efímeros altamente dinámicos y potencialmente interconectados por procesos de dispersión como modelo de estudio. Los lagos salinos efímeros son un escenario perfecto para evaluar el efecto de una presión selectiva variable (promovida por las amplias fluctuaciones de salinidad) y el efecto potencial sobre los procesos de ensamblaje complementarios como la dispersión y la deriva ecológica. El estudio de estos procesos a través de múltiples escalas ambientales, espaciales y temporales puede ofrecer una comprensión mecanicista de cómo los diferentes procesos interactúan para determinar el ensamblaje de las comunidades.

A lo largo de tres ciclos hidrológicos consecutivos, caracterizamos la dinámica mensual del microbioma acuático de 14 lagos salinos efímeros del desierto de Monegros (NE de España). Estos sistemas experimentaron severas variaciones en el volumen de agua a lo largo del ciclo hidrológico, los cuales promueven amplias fluctuaciones de salinidad que van desde agua dulce hasta la saturación de sales. Las comunidades de bacterias, arqueas y microeucariotas que habitan estos lagos se caracterizaron mediante la secuenciación masiva de los genes 16S y 18S rRNA, respectivamente. Evaluamos el efecto del aumento de la salinidad en la diversidad y composición de las comunidades microbianas, junto con los efectos ecológicos y metabólicos. A continuación, analizamos los lagos efímeros desde una perspectiva de

metacomunidad para evaluar el papel de los diferentes procesos de ensamblaje, así como sus interacciones, para originar las dinámicas observadas a lo largo del ciclo hidrológico.

Como se esperaba, observamos que la composición de las comunidades microbianas está fuertemente determinada por el aumento de los niveles de salinidad. Mientras que se observó un reemplazo taxonómico en las comunidades bacterianas, impulsado por extinciones masivas seguidas de la colonización por grupos halófilos mejor adaptados, las comunidades de eucariotas sufrieron un severo filtrado ambiental que dio lugar a un ensamblaje simplificado dominado por unos pocos grupos halófilos, principalmente las algas fotosintéticas *Dunaliella* y *Navicula*. Por el contrario, el rico y diverso ensamblaje de arqueas presente en los Monegros se hizo más abundante a medida que los lagos alcanzaban concentraciones hipersalinas. Mediante un novedoso enfoque basado en los rasgos genómicos, identificamos la transición de un ensamblaje dominado por procesos estocásticos a un ensamblaje determinista, describiendo un umbral consistente a salinidades c.a. 5‰. Más allá de este umbral, el aumento de la presión selectiva redujo el éxito potencial de los procesos de colonización entre los lagos, pero también la colonización potencial por parte de microorganismos alóctonos provenientes de otros biomas como sedimentos y suelos. Las potenciales interacciones entre especies a lo largo del gradiente de salinidad se abordaron mediante un análisis de redes que mostró que el aumento de la presión selectiva, y la transición asociada de grupos generalistas a especialistas, redujo la competencia y no promovió la facilitación entre especies. En general, el análisis multiescala de los procesos de ensamblaje deterministas y estocásticos en el sistema de lagos efímeros de Monegros, junto con los amplios conocimientos previos sobre la ecología y microbiología de los sistemas salinos y los efectos fisiológicos de los gradientes de salinidad, nos permitieron describir el ensamblaje de las comunidades desde una inusual combinación de puntos de vista mecanicistas e integradores.

Contents

1	General Introduction	1
1.1	Biological communities assembly	1
1.1.1	Local selection	1
1.1.2	Dispersal	2
1.1.3	Speciation	3
1.1.4	Ecological drift	4
1.1.5	Stochastic vs deterministic assembly	4
1.2	Defining a microbial metacommunity	7
1.2.1	Spatial scale	9
1.2.2	Temporal scale	10
1.2.3	Phylogenetic scale	12
1.2.4	Environmental scale	16
1.3	Ephemeral saline aquatic environments as model systems for evolutionary and ecological studies	18
1.3.1	Salinity in the microbial world	19
1.3.2	Osmoadaptation mechanisms	20
1.3.3	The Monegros desert lacustrine system	22
2	Objectives	25
3	Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain)	31
3.1	Introduction	32
3.2	Experimental Procedures	34
3.2.1	DNA Extraction and sequencing	34
3.2.2	Bioinformatic processing and genetic analyses	35
3.2.3	Statistical analysis and calculation of colonization and extinction rates	36
3.3	Results	38
3.3.1	Archaeal abundance and diversity	38
3.3.2	Archaeal assembly along the salinity gradient	39
3.3.3	Archaeal ecology and dynamic behaviour	42

3.4	Discussion	44
4	Ecological and metabolic thresholds in the bacterial, pro- tist and fungal microbiome of ephemeral saline lakes (Mon- egros Desert, Spain)	49
4.1	Introduction	50
4.2	Materials and methods	52
4.2.1	Study site and sampling	52
4.2.2	DNA Extraction and sequencing	53
4.2.3	16S rRNA and 18S rRNA genes sequence analysis	54
4.2.4	Statistical analysis	55
4.3	Results	56
4.3.1	Alpha diversity descriptors along the salinity gradient	56
4.3.2	Taxa replacement along the salinity gradient	57
4.3.3	Indicator bacterial taxa (IndVal) along the salt gradient	57
4.3.4	Indicator eukaryal taxa (IndVal) along the salt gradient	59
4.3.5	Metabolic and trophic thresholds along the salt gradi- ent in ephemeral lakes	61
4.4	Discussion	63
5	A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly)	69
5.1	Introduction	70
5.2	Materials and methods	72
5.2.1	Null models of community assembly	73
5.2.2	Statistics	76
5.3	Results	76
5.4	Discussion	80
6	Understanding stochastic and deterministic assembly pro- cesses in microbial communities along temporal and envi- ronmental gradients	87
6.1	Introduction	88
6.2	Materials and methods	90

6.2.1	Study site, sampling and genetic analyses	90
6.2.2	Environmental and spatial analyses	91
6.2.3	Dynamic analysis	93
6.3	Results	94
6.3.1	Selection by environmental filtering	94
6.3.2	Dispersal	95
6.3.3	Relative weight and combined effect of selection and dispersal	96
6.3.4	Dynamic behavior	99
6.4	Discussion	100
7	Absence of stress-promoted facilitation coupled with a competition decrease in the microbiome of ephemeral saline lakes	107
7.1	Introduction	108
7.2	Methods	109
7.2.1	Study site and molecular dataset	109
7.2.2	Network construction	110
7.2.3	Network analysis and null network model simulation	111
7.3	Results	112
7.3.1	Pairwise co-occurrences and co-exclusions	112
7.3.2	Networks and subnetworks composition and topology	112
7.4	Discussion	115
7.4.1	Inference of potential microbial interactions and eco- logical interpretation	115
7.4.2	The stress gradient hypothesis in the microbial world	116
7.4.3	Concluding remarks	118
8	Increasing saline stress reduces allochthonous microbial colonization in ephemeral ponds, driving the transition from habitat generalists to specialists.	119
8.1	Introduction	120
8.2	Materials and Methods	122
8.2.1	Water and sediments molecular datasets	122
8.2.2	Environmental Ontology sequences annotation . . .	123

8.2.3	Specificity index	124
8.2.4	Additional statistical analysis	124
8.3	Results	125
8.3.1	Water–sediment communities relationship	125
8.3.2	Predicted sources for pond microorganisms	127
8.3.3	Local and global specificity and generalism	129
8.4	Discussion	131
9	General Discussion	137
9.1	The hydrological cycle through the microbial metacommunity perspective	137
9.2	Assembly processes and the physiological explanation of emerging ecological patterns	141
9.3	Multiscale community assembly	144
9.3.1	Temporal scale	144
9.3.2	Spatial scale	145
10	Conclusions	147
	Bibliography	151
	APPENDIXES	181
A	Chapter 2 Supporting Information	183
A.1	Supplementary Tables	183
A.2	Supplementary Figures	187
B	Chapter 3 Supporting Information	189
B.1	Supplementary Tables	189
B.2	Supplementary Figures	197
C	Chapter 4 Supporting Information	201
C.1	Supplementary Information	201
C.2	Supplementary Figures	209
D	Chapter 5 Supporting Information	217
D.1	Supplementary Tables	217

D.2	Supplementary Figures	218
E	Chapter 6 Supporting Information	221
E.1	Supplementary Information	221
E.2	Supplementary Figures	222
F	Chapter 7 Supporting Information	227
F.1	Supplementary Figures	227
G	Published Chapters	229
H	Related works	263

GENERAL INTRODUCTION



General Introduction

1.1 Biological communities assembly

Biological communities encompass the sum of all organisms living and interacting within a particular geographical area. Every single community is a unique and complex system characterized not only by organisms identities but also their relative abundances. As MacArthur stated, “*patterns of species diversity exist*” (MacArthur, 1965), and one of the main goals in ecology is to reach a comprehensive understanding of the forces and mechanisms driving communities assembly, i.e. being able to determine why an specific community and not a different one is found in a certain place and time (Legendre et al., 2005; Leibold et al., 2004; Shmida and Wilson, 1985). Factors involved in community assembly have been largely proposed in the literature to the extent that, a quarter of a century ago, Palmer (1994) summarized more than one hundred different hypotheses to explain the arise and maintenance of species diversity. With the high number of described factors and mechanisms arose the need of a conceptual framework able to encompass and summarize all factors potentially shaping biological communities. In this context, Vellend (2010) proposed that all previously described factors can be included in four distinct kind of processes: selection, dispersal, drift and speciation. In nature, these processes acquire a changing importance according the scale at which community assembly is considered.

1.1.1 Local selection

Originally, biological communities were mainly considered as closed and isolated systems, and therefore mechanisms behind its assembly were only assessed at a local scale. At this point, selection was historically considered to rule communities assembly (Lawton, 1999). Selection is defined as the

sum of processes by which organisms relative abundances are determined by a differential functionality among species (Vellend, 2010). These differences imply a differential deterministic fitness in terms of interaction with the environment and among species. Environmental selection acts firstly by determining the capability of organisms to successfully establish and develop under specific environmental conditions. Additionally, idiosyncratic deterministic fitness differences among the organisms being present in a community may lead to an unbalanced distribution of relative abundances due to competitive advantage of some species in relation to others. Under changing environmental conditions, organisms deterministic fitness may change leading to variable equilibria which can even led to local extinctions. As stated above, selection also encompasses the effect of biological interactions on the community assembly. For macroorganisms, biological interactions have been historically considered key to determining the equilibrium among species, e.g. predator-prey dynamics (Berryman, 1992), plant-pollinators interactions (Soberon and Martinez Del Rio, 1981). However, biological interactions have been historically neglected when studying microbial communities assembly due methodological limitations for assessing potential microbial interactions at the community scale.

1.1.2 Dispersal

As a counterpoint to the local perspective for communities assembly where selective processes clearly prevailed over other factors, a more inclusive framework considers processes acting at a broader spatial and temporal scales (Ricklefs, 1987). In this context, the way in which communities interact with each other, i.e. the dispersal of organisms across space, enters into the equation at the time that it is recognized that factors taking place at a larger spatial scale have an effect on local communities. In this context several theories were developed considering colonization and extinction dynamics as main determinant of local patches (or "*islands*") composition. The island biogeography theory (MacArthur and Wilson, 1967) proposes dispersal from a main source ("*continent*") and no effect of local scale processes over community assembly. Later theories proposed a multidirectional network of local patches interconnected by dispersal processes, whether considering

single species and no effect of local scale processes (metapopulation, Hanski 1994) or multispecies assemblies in which local processes are also considered (metacommunity, Leibold et al. 2004).

Dispersal processes between local communities originate spatial patterns, the prevalence and weight of which are associated to the variable influence of dispersal in communities assembly. For instance, mass effects (Shmida and Wilson, 1985) and rescue effects (Brown and Kodric-Brown, 1977) may appear at high dispersal rates. This processes occur when a species with low fitness maintain its population within a given community by the arrival of individuals from areas with higher success. At high dispersal rates other processes as homogenizing dispersal, i.e. the decrease of the strength of selection by the effect of dispersal processes (Fodelianakis et al., 2019), may occur. By contrast, when dispersal rates between specific communities are low or null, i.e. under dispersal limitation, local patches with similar environmental characteristics may present differentiated communities due the limitation for potential colonizers to arrive and establish.

1.1.3 Speciation

If wider temporal and spatial scales are considered, other processes may arise as determinant of communities assembly. Speciation, defined as the appearance of new species by random mutations, was firstly proposed as a mechanism with influence in determining communities composition by Charles Darwin when, in 1845, described the Galapagos islands finches communities. What Darwin described as “*the first appearance of new beings on this earth*” was indeed the result of a speciation process given within islands local populations. Speciation can be usually neglected when assessing differences within interconnected local communities since it occurs at evolutionary timescales, but it should be considered to be a key component in the configuration of the regional pool. However, it has been proposed that the high speciation rates of microbial organism might potentially led to a convergence of both ecological and evolutionary timescales (Barberán and Casamayor, 2014).

1.1.4 Ecological drift

In parallel to these processes, stochastic demographic events (births, deaths) affect local populations and modify their relative abundances, which is known as ecological drift. Relative abundances fluctuations can lead to destabilize the established balances between local species and even lead to local extinctions. Since drift is a stochastic process, the probability of one species whether to get extinct or to reach a mono-dominance within the community is linked to its initial frequency when only drift is acting. Also, the rate at which the mono-dominance would be reached is related to community size. Therefore, the potential effect of drift over communities composition would increase as community size decreases (Vellend, 2010).

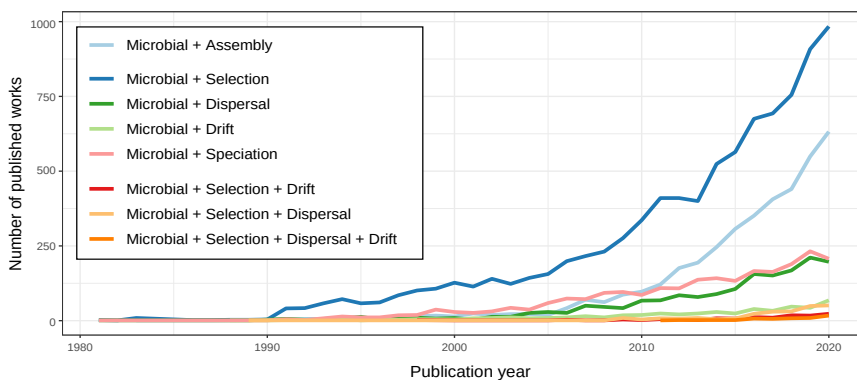
1.1.5 Stochastic vs deterministic assembly

Whether communities are assembled by deterministic or stochastic processes is one of the main current debates in community ecology. However, this dichotomy has been present in the field since quite long. Already in 1859, Darwin made his philosophical position clear by stating that there was no place for stochasticity or chance in biological communities assembly (Darwin, 1859). In this line, ecological communities have been traditionally understood as high-dimensional complex systems, being result of differences in the interaction of every species with the physical environment and other species. Consequently, the traditional approach to approximate its understanding has been the consideration of as many dimensions as possible, and the incorporation of the greatest number of details into complex models. As a counterpoint to this maximalist approach which try to explain communities assembly in the most accurate way, neutral theory adopt a reductionist approach in which, from the simplest model, incorporates the minimum complexity that allows to reproduce observations (Hubbell, 2001). Therefore, the fundamental key lies on determining what is the minimum necessary information to characterize a community at the desired level of precision and realism. With this, the initial model requires of a series of simplifications about communities, some of which created great controversy among ecologists (Alonso et al., 2006). Undoubtedly, the most criticized has

been the functional equivalence of species on the same trophic level. This simplification neglect species ecological niches and assume that trophically similar species are demographically identical, also equivalent in their dispersal capability (Hubbell, 2001). Although this simplification has been widely questioned, the neutral theory have been proved to be useful for yielding insights into the functioning of multiple systems (Etienne and Alonso, 2007; Rosindell et al., 2012). Indeed, what have been considered as two contrasting ecological theories are, in practice, two complementary tools to better understand the processes mediating communities assembly. Furthermore, studies independently supporting the dominance of both frameworks (Fierer and Jackson, 2006; Logue and Lindström, 2010; Van der Gucht et al., 2007) pointed to an equilibrium between deterministic and stochastic factors which may vary depending on the idiosyncrasy of the studied system (Gravel et al., 2006).

Box 1: Microbial assembly processes on the literature

The evolution of the published papers assessing microbial community assembly processes reflex the evolution produced in the field, moving to a more inclusive perspective of the assembly process. Also, advances on the study of microbial communities are also linked to methodological advances as the rise of novel sequencing technologies (Mardis, 2013). As in general ecology, local selective processes accounted for most of the attention since the beginning of microbial communities exploration. At this point, selection was mainly focused on environmental selection as a consequence of a lack of robust methodologies to assess microbial interactions at the community level. However, only a few works yet evaluated the role of specific interactions as predation (Arndt, 1993) or facilitation (Bever, 1994), both within microbial communities and between microbes and plants. It was not until the further development of next generation sequencing techniques (Mardis, 2013) and the development of more inclusive frameworks (e.g. metacommunity, Leibold et al. (2004)) when dispersal and speciation processes gained importance on the microbial community literature. Likewise, ecological drift increased its prevalence in the literature from 2010, coinciding in time with the establishment of the conceptual framework proposed by Vellend (2010). However, its prevalence on the current literature remains still lower than previously reported processes due the higher conceptual and methodological complexity of its assessment. Also, the simultaneous assessment of several processes has been underexplored until very recently. Only in the last few years a robust core of studies assessing simultaneously dispersal and selection mechanisms has started to be built, while those incorporating drift and selection together with other processes are still scarce.



Number of yearly published papers as reported through a search in “web of science” (WoS) according to the different keywords.

1.2 Defining a microbial metacommunity

Quantitatively –or pseudo-quantitatively– assessing the role of each of the factors taking place on the assembly of a community requires, necessarily, to initially establish the scale at which community assembly is considered. It is broadly recognized that ecological processes are, at least to some extent, scale dependent (Belmaker and Jetz, 2011; Belmaker et al., 2015; Levin, 1992; Maurer and Taper, 2002; McGill, 2010; Schneider, 2001). Therefore, selected scales strongly determine the differential effect of those processes mediating communities assembly and, with it, any result and conclusion regarding assembly processes. Although the spatial scale has been undoubtedly the most widely studied (Belmaker and Jetz, 2011; Belmaker et al., 2015), also phylogenetic (Graham et al., 2018) and temporal (Debray et al., 2021) scales are key determinants on community assembly. The established scales not only determine the variable role of the different processes, but are also key for assessing different processes and patterns which are scale specific. Indeed, one of the main challenges for general –and therefore also microbial– ecologists is to successfully match (i) the question, (ii) the theoretical framework, (iii) the study system and (iv) the scale at which the system will be considered (Prosser et al., 2007). However, some questions cannot be fully addressed from a single scale perspective and require of a multiscale approach. Thus, it has been suggested that the combined study of both local and short-term mechanisms as well as regional processes occurring over longer timescales may be crucial for a complete understanding of ecosystem assembly and function (HilleRisLambers et al., 2012). In this context, Leibold and collaborators proposed the metacommunity concept, i.e. a set of local communities linked by dispersal of multiple interacting species (Leibold et al., 2004), as a multi-scale framework to the study of processes shaping local communities. In this framework, both the local and regional (metacommunity) scales are considered, with different processes acting at each scale.

Additionally, other processes acting at a broader spatial and temporal scale have an impact on the configuration of the regional pool and, therefore, of local communities composition. Indeed, one of the main advances of broadening the consideration of communities assembly, previously assessed

from a local perspective, was the recognition that composition and diversity of local communities was fundamentally dependent on the composition and diversity of the regional species pool (Vellend, 2010). In this framework, local communities composition results from a selected subset of species present on a regional pool or metacommunity, whose specific composition and diversity is controlled by both stochastic and deterministic processes. The composition of the regional pool of species is conditioned by processes acting at a broader spatial and temporal scales. With this, the regional pool is progressively configured by the effect of environmental selection, present and past dispersal events, and speciation (Barberán and Casamayor, 2014). Moving from this conceptual framework to establish the actual scale at which assembly processes operate is not trivial and it is strongly dependent on the studied systems.

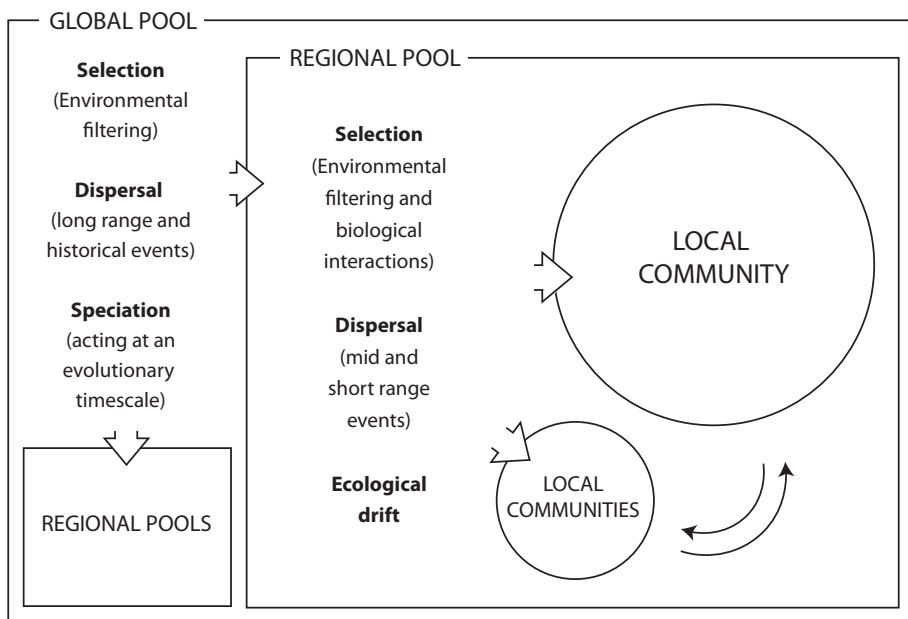


Figure 1.1: Assembly processes determining the species pool of the global, regional and local pools. Adapted from Barberán et al. (2014).

1.2.1 Spatial scale

The spatial scale at which assembly is considered has a strong influence on the prevalence of the different assembly processes (Viana and Chase, 2019). Variable scales modulate processes linked to dispersal and environmental selection (Viana and Chase, 2019), but also processes that are apparently less directly related as biological interactions (Araújo and Rozenfeld, 2014; Belmaker et al., 2015). Establishing the spatial scale of a metacommunity implies the determination of local communities which are potentially linked by dispersal processes (Leibold et al., 2004). However, the scale at which establishing the potential dispersal between local communities is highly dependent on numerous factors. The environment typology (aquatic, terrestrial, aerial), its continuity (continuous vs discrete landscape) or the potential limitations to dispersal (e.g. geologic barriers) are the main determinant environmental factors. Furthermore, the idiosyncrasy of the targeted organisms is also crucial for its potential dispersal. In this sense, the definition of a microbial metacommunity is particularly complex as the limits to dispersal are not clear for microbes. For a long time, the prominent statement “*everything is everywhere, but, the environment selects*” (Baas Becking, 1934) condensed the widespread belief that microbes were ubiquitous and lacked limitations to dispersal due to their small size, short reproduction time and elevated survival. Although this did not imply that microbes do not exhibit biogeographic patterns, those were considered to be negligible compared to processes acting locally as selection. However, later descriptions of consistent biogeographical patterns in microbial communities, linked to the progress in resolution of communities composition and large-scale studies (Battle et al., 2018), highlighted the importance of dispersal processes (Telford et al., 2006). Nonetheless, microbial biogeographical patterns are usually weaker than the ones presented by macroorganisms in equivalent conditions, which has been attributed to the lower limitations to dispersal of microbes (Barberán and Casamayor, 2014) and differences on sampling effort derived from the imperfect detectability inherent to microbial surveys (Meyer et al., 2018).

The observed idiosyncrasy of microorganisms suggested that ecological frameworks historically developed for animals and plants might present lim-

itations for its application to microbes. This put microbial ecologists in the diatribe of either building a novel core of ecological knowledge to explain the specific processes observed in the microbial world, or incorporating microbial particularities into pre-existing ecological theories originally conceived for macroorganisms (Barberán and Casamayor, 2014; Prosser et al., 2007). In this context, the already mentioned lower limitation to dispersal, together with a higher colonization success and lower extinction, drove Barberán and collaborators (2014) to propose an adaptation of Hubbell’s neutral model (Hubbell, 2001) in which the relative weight of the global pool in local communities assembly increased in detriment of the regional pool. Since this adapted framework requires of several assumptions regarding the lack of limitation to dispersal, this might not apply to dispersal limiting microorganisms. This limitation, given by intrinsic characteristics of the organisms, might be linked to its habitat breath or lack of resistance forms, among others. Nevertheless, although the boundaries between global and regional pools might be blurred for microbes, the definition of a metacomunity and its regional pool might be still useful and conceptually valid for the study of local assembly processes. The potential lack of limitations to dispersal does not imply and equal probability to dispersal across space, and therefore local communities will be more affected by dispersal events from close communities than from distant ones. Then, the challenge is to establish the most consistent and meaningful limit, without overlooking that the assembly cannot be fully assessed within the metacomunity context without considering also processes acting at a broader spatial scale.

1.2.2 Temporal scale

Temporal scales and dynamics have been considered in ecology for long time (Magurran, 2008). Time scales were first explicitly considered in the study of ecological communities by Preston (1960) when describing the pattern of species-time relationship, i.e. the link between species number and time. Since then, variations in communities composition or species richness have been studied linked to temporal dynamics, with focus on patterns as species temporal turnover (Hillebrand et al., 2018) or time-decay of similarity in communities composition (Dornelas et al., 2014). These patterns have been

also reported for microbial communities since the appearance of temporal series on the literature, mostly linked to the development of next generation sequencing technologies. Many studies considering the temporal dimension in microbial communities describe beta-diversity patterns as a time-decay of similarity in communities composition (Barberán and Casamayor, 2014; Cáliz et al., 2018; Guo et al., 2018) or seasonality patterns (Eiler et al., 2012; Höfle et al., 1999; Shade et al., 2007). However, many of the described patterns are linked to environmental changes occurred over time rather than to pure temporal processes, hindering the analysis of dynamic processes leading to temporal turnover without the influence of environmental fluctuations. In this context, some studies override seasonal environmental fluctuations by describing interannual trends (Cáliz et al., 2018) or in environmentally stable experimental setups (Guo et al., 2018). Beyond the descriptions of alpha- and beta- diversity patterns along time, MacArthur and Wilson (1967) developed the equilibrium theory of island biogeography as the first attempt to unveil community dynamics from a mechanistic point of view (Losos and Ricklefs, 2010). This neutral theory focuses on the dynamic equilibrium between colonization and extinction events as determinant of communities species richness and temporal turnovers (Simberloff, 1969).

The consideration of the temporal scale is also key for understanding the potential influence of past conditions –or events– on the present state of the system. The impacts that past conditions have on the current properties of the community are referred as legacy effects (Monger et al., 2015). For instance, it is well known that current primary production of plants is strongly affected by previous-year precipitation (Sala et al., 2012) or that land-use history is a major factor determining its current ecological state (Foster et al., 2003). Legacy effects can encompass both biotic (priority effects) and abiotic (species sorting by past environmental conditions) factors (Vass and Langenheder, 2017).

Priority effects are produced when the history of colonization events, i.e. the arrival sequence of colonizers, determine the arrival and establishment of later colonizers (Fukami et al., 2007). There are three main mechanism by which priority effects can operate, which can be classified according to whether they facilitate or inhibit the establishment of potential colonizers

arriving later. Inhibition of later arrivers in theis establishment in the community can be produced by a niche pre-emption, depleting habitat resources (including space), or modification in a way that prevents the establishment of other groups. Conversely, facilitation can be produced by a niche modification in a way that benefits the establishment of later arriving organisms. Facilitation can occur by several mechanisms including stress amelioration (Monier and Lindow, 2005) or resource availability by the degradation of complex substrates into more accessible ones (Enke et al., 2019), among others. Although these effects are mostly detected during the period following the initial colonization, or the recovery following a perturbation, it can condition the final composition of the community and therefore propagate over time indefinitely (Debray et al., 2021).

The effect of past abiotic factors may also play a major role on the current composition and structure of the community. Species sorting by past environmental conditions, also referred as time-lags in environmental selection, occur when current populations are totally or partially adapted to previous environmental conditions rather than present ones (Blanquart and Gandon, 2013; Fox and Harder, 2015). The presence and duration of this lag depends on several factors including the adaptive potential of local communities and the rate at which better adapted groups arrive or resuscitate from dormancy and grow (Fukami, 2015). While the latter factor is strongly linked to present and past dispersal events, the former depends on the inhabiting community and the nature of the environmental change. Although time-lags in environmental selection are expected to be low in microbial communities due to high dispersal rates and short replication times, previous studies have described the effect of past environmental conditions over inhabiting microbial communities up to several days (Andersson et al., 2014; Berga et al., 2012). In this case, and contrary to priority effects whose effects can propagate indefinitely over time, the effects of past abiotic environmental conditions will be transitory.

1.2.3 Phylogenetic scale

Many of the historically reported patterns and processes vary across the phylogeny and are phylogenetically scale dependent (Buckley et al., 2010;

Damuth, 1981; Šizling et al., 2011; Storch and Šizling, 2008). Although several works had already described this dependency, the phylogenetic scale and its role on the study of community ecology, macroecology or biogeography patterns and processes has not been formally conceptualized until very recently (Graham et al., 2018). The phylogenetic scale can be assessed in terms of extent and grain (Ladau and Eloe-Fadrosh, 2019). The grain refers to the elementary unit of analysis which can be defined in terms of clade size or taxonomic rank, whereas the extent accounts for the total phylogeny encompassing the organisms under study (Graham et al., 2018).

A narrow phylogenetic extent will only consider closely related taxa (e.g. within the same order, genus or even species) while wide extents will consider more evolutionarily distant organisms. As stated above, many historically reported ecological patterns are consistent across a given phylogenetic scale. For instance, historically reported trends as the negative correlation between body mass and population abundance (Damuth, 1981; Isaac et al., 2011) or the latitudinal decline of species richness (Buckley et al., 2010) might not be consistent across variables extents. The correlation between body mass and population abundance arise when evolutionarily distant organisms

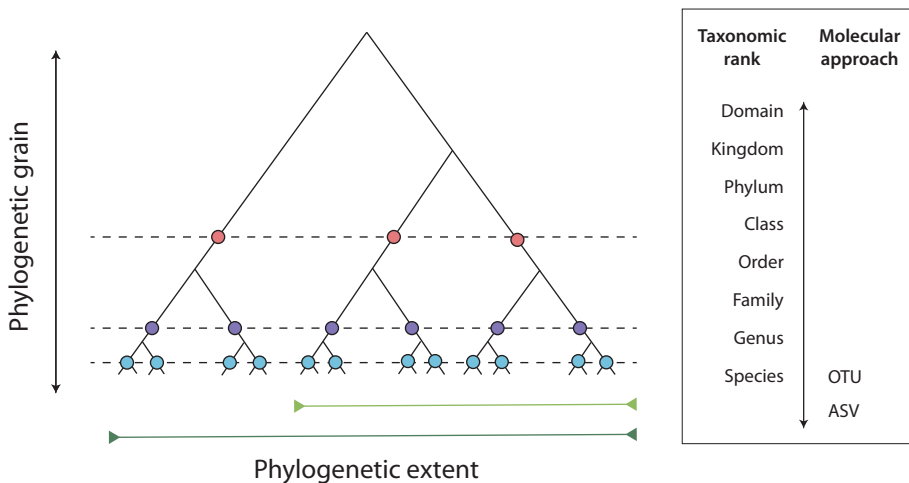


Figure 1.2: Conceptual definition of the phylogenetic extent and grain.

are considered but tends to disappear within ecologically similar species (Cotgreave, 1994; Damuth, 1981; Isaac et al., 2011). Likewise, the latitudinal diversity gradient observed for mammals disappears across small clades within this group at the time that other patterns arise (Buckley et al., 2010).

The phylogenetic extent under consideration also have an impact on the study of community assembly processes. By only considering phylogenetically close groups, which most likely will be also ecologically similar (Losos, 2008), differences on abiotic (environmental fitness) and biotic (biological interactions) preferences are expected to be low. Therefore, the effect of selection on determining the assembly and distribution of those group is as well expected to be low. This way, the establishment of a phylogenetic extent narrow enough to make ecological differences between groups negligible would meet to an analogous scenario to the equivalence assumption of Hubbell neutral theory (Hubbell, 2001) in which community assembly is driven by stochastic processes. Conversely, a wide phylogenetic extent considering ecologically distant groups will emphasize the role of selective processes over stochastic ones. Additionally, different groups might be differentially assembled by different processes, which will determine the prevalence of specific community assembly processes based on the composition of the studied communities. These differences have been clearly reported between bacterial and microeukaryal communities inhabiting the same systems (Casamayor et al., 2013; Logares et al., 2018), but also within microbial domains among groups presenting ecological differences as habitat specialists and generalists (Liao et al., 2016; Pandit et al., 2009; Székely and Langenheder, 2014). Furthermore, the phylogenetic extent is also key for understanding the potential role of biological interactions on the community assembly. This is especially important for microbes due the relevance of host-associated microbial groups, but also for free-living microbes involved on complex biogeochemical cycles or food webs (Milici et al., 2016; Paerl and Pinckney, 1996).

The phylogenetic grain, i.e. the resolution at which the studied organisms are defined, is also a fundamental issue when evaluating the different processes mediating communities assembly and specific patterns within microbial communities. One of the patterns which has been reported to be

grain-dependent is nestedness of microbial communities (Thompson et al., 2017), which occurs when less diverse communities are conformed by a subset of more diverse ones. For instance, marine communities are generally nested within soil ones at the phylum level. However, this pattern progressively disappears with the increase of taxonomic resolution to the extent that only a small proportion of OTUs are shared between both environments (Thompson et al., 2017). With this, determining whether nestedness (subsetting) or turnover (tax replacement) drives diversity patterns among communities (beta-diversity) is strongly dependent on the phylogenetic grain under consideration. At the same time, the potential contribution of assembly processes as speciation, or the strength of biogeographic patterns (Bay et al., 2020) will increase at fine phylogenetic resolution. Additionally, the explicit consideration of the phylogenetic grain and the embracement of multi-scale approaches have been proposed as a useful tool for the study of the processes driving the assembly of microbial communities (Chalmandrier et al., 2019; Graham et al., 2018). For instance, some authors propose that, while the prevalence of environmental filtering is expected to determine beta-diversity patterns which are detectable at a coarser phylogenetic grain, the effect of ecological drift derived from a stochastic assembly should be mainly detectable at finer grains (Chalmandrier et al., 2019).

The phylogenetic grain and the definition of the elementary unit of biological diversity, the specie (Hohenegger, 2014), is specially controversial for prokaryotes (Cohan, 2002; Rosselló-Mora and Amann, 2001). While in animals and plants the genetic cohesion sustaining the species definition is easily characterized by sexual compatibility and population structure, which allows an ecologically and biologically relevant definition of species, this issue is still challenging for prokaryotes (Bobay, 2020). A pragmatic approach to overcome this limitation was the definition of operational taxonomic units (OTUs) rather than species as a fundamental unit of study, which are based on sequence similarity thresholds. The identification of the 16S rRNA subunit as a universal gene shared by all protists (Woese and Fox, 1977) allowed the assessment of the whole prokaryotic diversity with a single marker gene. The correspondence to the species level previously defined by the 70% identity based on DNA-DNA hybridization essays

(Brenner et al., 2015) was then established at the 97% of identity for the 16S rRNA subunit (Stackebrandt and Goebel, 1994), and the definition of 97% identity OTUs has been the standard in microbial community analysis for the last decades. Recently, novel bioinformatic methods have developed sequence error-correction (or denoising) strategies to identify real biological sequences, which allows to define operational taxonomic units at a single nucleotide resolution (Patil, 2019). These units receive different names as Amplicon Sequence Variants (ASV, Callahan et al. 2016), sub-OTUs (sOTU, Amnon et al. 2017) or zero-radius OTUs (zOTU, Edgar 2016) according to the different bioinformatic software. The application of these methods improves phylogenetic resolution compared to previous clustering approaches, which might have an impact on the study of the microbial community assembly processes emphasizing the role of dispersal limitation and speciation processes.

1.2.4 Environmental scale

The environmental scale can be defined as the variability or amplitude regarding one or several descriptive parameters of the environment. This variability can encompass a multidimensional landscape including physical, chemical and biological characteristics of the environments. The environmental scale is usually referred as environmental heterogeneity and is rarely considered by itself on the ecological theory but as the result of the defined spatial and temporal scales. However, many ecological patterns reported along space and time are indeed driven by the environmental changes occurred along those scales. For instance, historically described latitudinal gradients in species diversity (Hillebrand, 2004) are the result to environmental changes linked to a spatial configuration. Likewise, seasonal patterns (White and Hastings, 2020) which are usually described in terms of temporal fluctuations are the result of environmental changes linked to time. Nonetheless, the environmental scale considered determines, to a large extent, the contribution of the different assembly processes. Studies describing differences in dominant community assembly processes along variable spatial scales relate the observed patterns to the increase in environmental heterogeneity with spatial extent (Garzon-Lopez et al., 2013; Viana and Chase, 2019).

As previously reported, the assembly process shows higher contribution of neutral assembly, i.e. driven by stochastic processes, when studied at lower scales and therefore lower environmental heterogeneity. Conversely, a wide environmental scale, i.e. high environmental heterogeneity, will most likely result on a dominance of niche assembly processes. With this, and based on the results of the study by Garzon-Lopez and collaborators (2013), Chase (2014) postulated that the spatial scale “*resolves the niche versus neutral theory debate*”, despite that it is the environmental heterogeneity linked to the spatial extent which primarily modulates the contribution of the different assembly processes.

The effect that environmental heterogeneity has on community assembly processes and therefore on the resulting beta-diversity patterns will depend on the niche breadth of the inhabiting organisms and the nature of the environmental variability. The niche breadth encompass the environments tolerated by an organism and the diversity of resources it can use, defined as an “n-dimensional” set of biotic and abiotic factors or conditions, i.e. niche axis, for each organism (Carscadden et al., 2020). Organisms exhibiting a narrow breadth along one or several niche axis are considered as “*specialists*”, while those presenting wide breadths are considered as “*generalists*”. Both specialists and generalists can be in terms of resource (Poisot et al., 2012), host (Price, 1977) or habitat (Monard et al., 2016). The differential contribution of assembly processes over habitat generalist and specialists groups have been reported for microbes (Liao et al., 2016; Luo et al., 2019; Székely and Langenheder, 2014) but also for other organisms as invertebrates (Pandit et al., 2009). Most studies support a dominance of environmental selection on habitat specialists distribution, while habitat generalists are found to be affected by selective but also stochastic processes (Liao et al., 2016; Luo et al., 2019; Pandit et al., 2009). Among all the physicochemical variables that affect the characterization of a habitat, some factors have a greater effect on the composition of the community with the potential to inhabit it. For instance, strong differences are expected to be found among microbial communities inhabiting environments with variable levels of water (aquatic or terrestrial organisms) oxygen (aerobic/anaerobic groups) or light (presence of photosynthetic groups). Additionally, other

physicochemical variables as pH (Cao et al., 2016; Steenwerth et al., 2002; Zhalnina et al., 2015), temperature (Skirnisdottir et al., 2000; Wang et al., 2013; Wu et al., 2019) or salinity (Dupont et al., 2014; Filker et al., 2017) have been described to have an strong influence on determining the microbial communities composition. Indeed, salinity has been repeatedly reported as one of the main factors determining microbial composition at a global scale both for bacterial (Barberán and Casamayor, 2010; Lozupone and Knight, 2007), archaeal (Auguet et al., 2010) and microeukaryal (Mo et al., 2021) groups.

The study of strong gradients linked to these environmental variables allow to explore how microbial communities respond to a variable selective pressure and the associated environmental filtering, as well as its effect on other potential assembly processes (He et al., 2021; Rousk et al., 2010; Teng et al., 2017; Tripathi et al., 2018; Wang et al., 2017; Xu et al., 2021). However, strong environmental gradients are usually registered across wide spatial scales or over long periods of time, making it difficult to study their potential effects alone.

1.3 Ephemeral saline aquatic environments as model systems for evolutionary and ecological studies

Ephemeral saline aquatic environments present specific hydrogeomorphological conditions which drive to severe fluctuations on its hydrographic regime over time. These fluctuations are in turn translated into wide salinity variations in site and over short time periods, being therefore a perfect scenario for the study of ecological processes across variable selective pressure. Ephemeral saline aquatic environments have been already considered for the study of community dynamics for microbes (Martin-Cuadrado et al., 2019) but also for other organisms as crustacea, (McCulloch et al., 2008) other invertebrates (Carver et al., 2009) or phytoplankton (Álvarez et al., 2006). Furthermore, the variable selective pressure exerted by changing salinity levels allow to evaluate not only assembly but also evolutionary processes (Wiggins et al., 1980).

Salinity levels and salts composition of saline aquatic environments are strongly related to its origin. Seawater present variable salts concentrations ranging between 3.1 and 3.7% with an average concentration of 3.5% (w/v, i.e. 35 g/l or 35 ppt), and with Cl^- (55%), Na^+ (30.6%) and SO_4^{2-} (7.7%) as main ions. Likewise, hypersaline environments which are derived from seawater, and subject to evaporation which increases its salts concentration, present similar ionic compositions (Rodriguez-Valera, 1993). Contrary, athalassohaline aquatic systems are continental-derived saline environments which present a differentiated ionic composition, weather by the lack of an specific ion or changes on its proportion. The vast majority of these systems are located in arid and semi-arid areas (Okin et al., 2006; Saccò et al., 2021), where hydrological imbalances promotes a dominance of evaporation processes which induce the progressive accumulation of salts (Yechieli and Wood, 2002).

1.3.1 Salinity in the microbial world

Organism (and microorganisms) can be classified according to its tolerance or requirement of salts for its successful development, ranging from non-halophile (or freshwater) to extreme halophile. Non-halophile organisms develop under non-saline conditions and lack of osmoadaptation mechanisms, not being therefore able to adapt to saline concentrations in the media. Halotolerant organisms can also develop in freshwater environments as they do not require of elevated salts concentrations. However, halotolerant can also grow under variable saline condition, from marine to hypersaline, due to the presence of osmoadaptation mechanisms (Imhoff, 2001; Oren, 2013). Different halotolerant groups present variable tolerance levels and are therefore able to develop up to different salinity concentrations. Finally, halophile organisms not only can tolerate saline conditions but require of them for its successful development, not being therefore able to grow below a specific threshold. Halophile organisms can be classified as slight, moderate, or extreme, according the range of salts concentration they require for growing (Caumette, 1993; Imhoff, 2001; Kushner, 1978, 1985). Within halotolerant and halophile organisms, differences in tolerance to specific salt ranges or salinity requirements are mainly linked to the nature of different

osmoadaptation mechanisms.

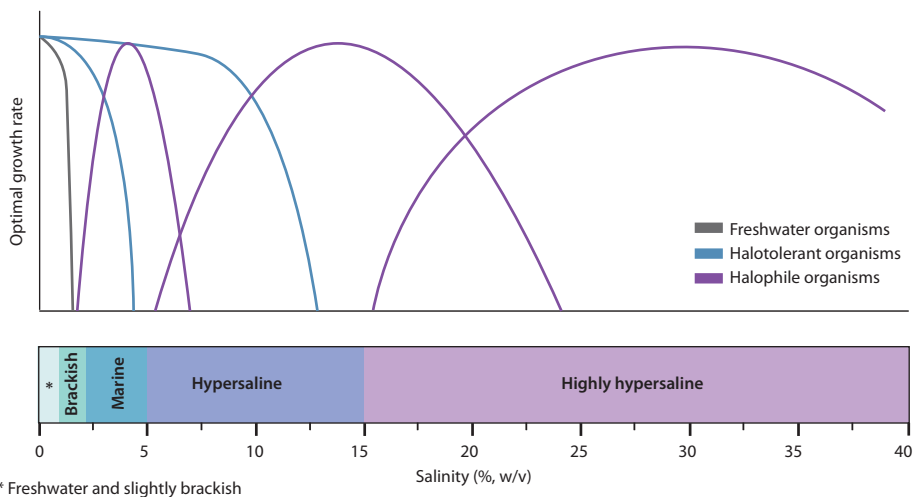


Figure 1.3: Different water masses according its salinity, as classified in the present thesis, and schematic representation of optimal growth rates for freshwater, halotolerant and halophilic organisms.

1.3.2 Osmoadaptation mechanisms

The challenge for organisms to grow in environments with high concentrations of salt is, ultimately, a physical question. Usually, cells maintain a slightly higher intracellular concentration of solutes on the cytoplasm compared the extracellular media. This way, the semipermeable cellular membrane allows a flux of water into the cell, driven by the osmotic pressure originated by the differential salts concentration, which maintain the required cellular turgor pressure (Sleator and Hill, 2002). However, when salts concentration on the environment are elevated, osmotic pressure forces the water flow towards the extracellular medium which, in absence of osmoadaptation mechanisms, can led to a loss of turgor and plasmolysis, i.e. the contraction of the cytoplasm due the loss of water.

Osmoadaptation mechanisms focus on maintaining an elevated intracellular solutes concentrations in order to avoid the water flow outside the cell, maintain therefore cellular turgor (Brown, 1976). For this porpoise, two main mechanisms have been described in nature to deal with the potential

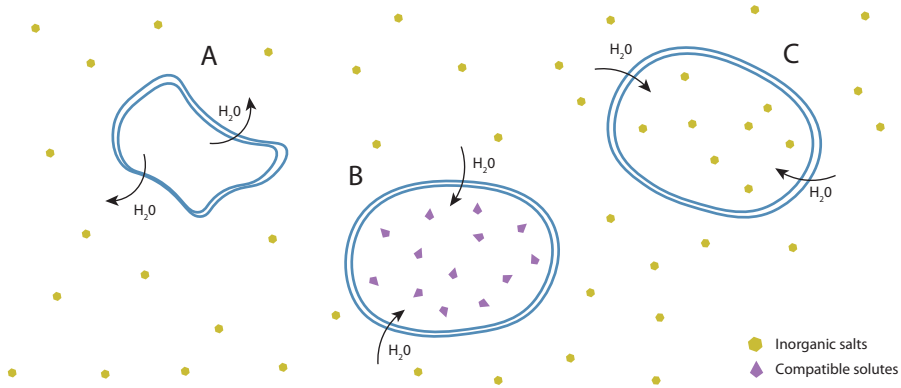


Figure 1.4: Cellular response of different groups under elevated salts concentrations. A: plasmolysis due a lack of osmoadaptation mechanisms. B: synthesis of incorporation of compatible solutes. C: accumulation of inorganic salts by active pumping from the extracellular medium

osmotic pressure: the accumulation of salts and the accumulation of compatible solutes within the cell (Galand et al., 2010; Trüper and Galinski, 1990). The accumulation of inorganic salts within the intracellular medium, known as "salt-in" strategy, is produced by an active pumping of salts (mostly Cl^- and K^+) from the surrounding environment (Galinski and Trüper, 1994). As elevated salt concentrations within the cell exposes the cytoplasm to an elevated ionic strength, specific adaptations off the intracellular enzymatic machinery are required (Sleator and Hill, 2002). This permanent adaptations, based on proteins amino acid composition and its structural configuration (Dennis and Shimmin, 1997; Sleator and Hill, 2002), limit its adaptation to a certain range of elevated salts concentrations. This strategy has been reported for halophilic archaea (Oren, 1999, 2008) and a limited group of halophilic bacteria (Oren, 2002). Conversely, the accumulation of compatible solutes, known as "salt-out" strategy, is widespread across bacterial and eukaryal halotolerant and halophilic groups. This strategy is based on a two-phase response to increases in salts concentration. The primary response consist on an initial accumulation of intracellular K^+ , followed by a second response in which the organisms either synthesize the compatible solutes or actively transport them from the extracellular medium (Sleator and Hill, 2002). Once reached stable osmolarity levels, intracellular

K^+ remains at low concentrations while organic molecules represent the main solutes maintaining positive osmotic pressure within the cell. Different organic compounds can be either synthesized or accumulated depending on the idiosyncratic characteristics of the environment and the organism. These can be either polyols and its derivatives, amino acids, sugars or betaine and ectoines, among others (Galinski, 1995). Although the nature of these compounds is highly variable, its common characteristic is its compatibility at high concentrations in the intracellular environment with the correct functioning of the internal cellular machinery.

1.3.3 The Monegros desert lacustrine system

The Monegros Desert is located in north-east Spain and covers about 276,440 hectares (Mor et al., 2010). Within it, the Monegros desert lacustrine system is located within the Central Ebro Basin ($41^{\circ}42'$ N, $0^{\circ}20'$ W), with the Ebro river running c. 5 km south. The area is an endorheic platform of c. 400 km² with 149 closed depressions, some of which host ephemeral saline lakes, or playa-lakes, with brines of Cl-SO₄-Na-(Mg) type (Gutiérrez and Gutiérrez, 2014). The depressions, which range in size from <2 to 239 ha, were originated by the dissolution of the subsurface gypsiferous bedrock, with contribution of aeolian deflation caused by the strong winds registered on the area (Gutiérrez and Gutiérrez, 2014). The morphology of the approximately 20 depressions hosting playa-lakes, presenting flat and moist bottoms, are indicative of a topography controlled by the water table (Rosen, 1994) as a result of a close connection with two main aquifers in the area (Samper-Calvete and García-Vera, 1998). This set of inland saline lakes is one of the largest in Europe and its scientific value has been repeatedly described at geological (Mees et al., 2011; Pueyo Mur and Inglés Urpinell, 1987) hydrological (Samper-Calvete and García-Vera, 1998) and ecological (Baltanás, 2001; Conesa et al., 2011; Pedrocchi, 1998) level. Previous studies on the area already reported a highly diverse assembly for the three microbial domains (Casamayor et al., 2013) and described its food-web structure (Alcorlo, 1999; Alcorlo et al., 2001).

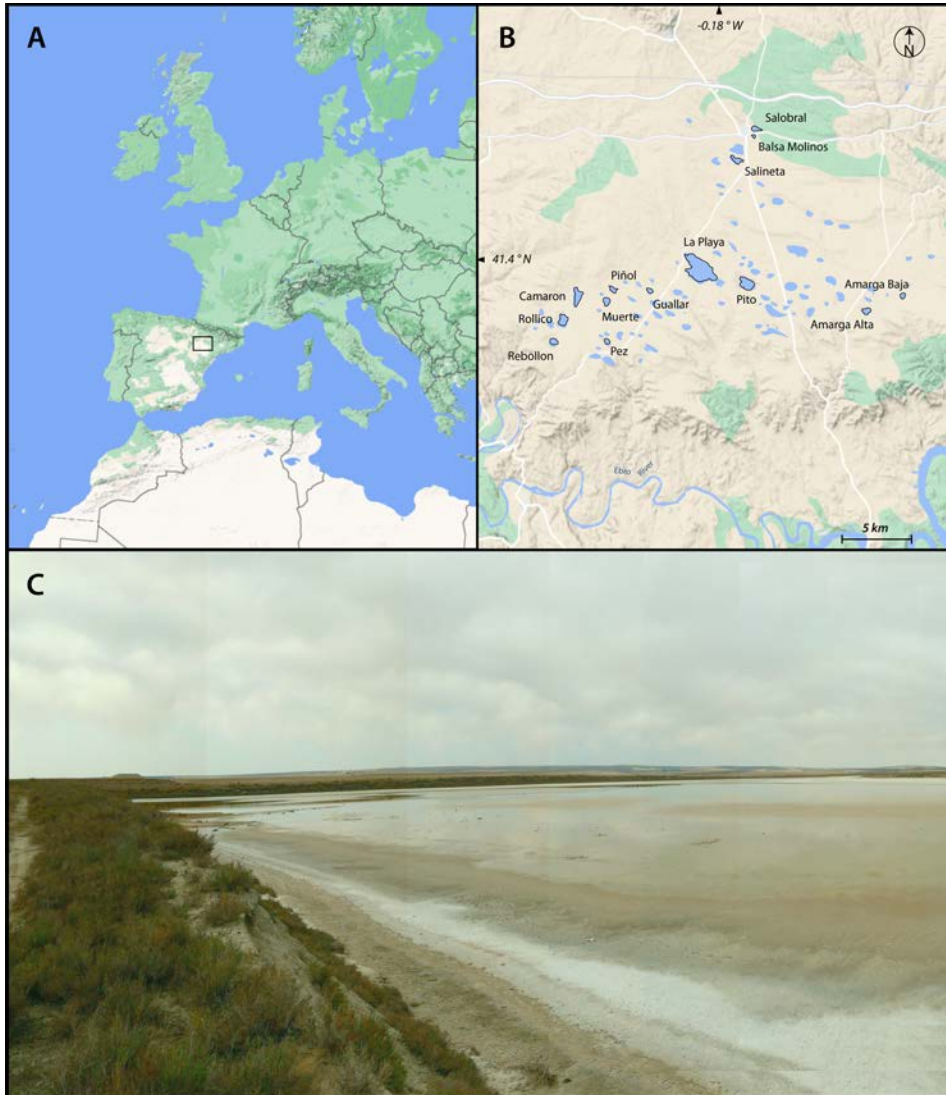


Figure 1.5: (A) Situation of the Monegros Desert lacustrine system within the European context. (B) Sampled lagoons within the Monegros desert lacustrine system. (C) La Playa, one of the 14 sampled lagoon, as seen in November 2017 © Xavier Triadó Margarit.

Objectives

This thesis aims to provide a comprehensive understanding of the processes driving community assembly on a highly dynamic environment in which environmental, spatial and temporal scales potentially play a prominent role. For this, we studied the Monegros desert lacustrine complex as a model system due the elevate environmental dynamisms registered on the area and its spatial configuration, which allow to hypothesize a metacommunity dynamic. The sampling scheme covered all microbial compartments (archaea, bacteria and microeukaryotes) inhabiting 14 lagoons along three hydroperiods and capturing salinity fluctuations from freshwater to near saturation, providing a multidimensional perspective of assembly processes. The initial objective lies on describing the inhabiting communities and linking its structure, composition and functioning to the variable selective pressure associated to the studied salinity gradient. Additionally, lagoons and their inhabiting communities will be considered from the metacommunity perspective in order to evaluate potential dispersal dynamics and their influence on determining local communities over space and time, both among lagoons but also considering inter-compartments dispersal.

The specific objectives of each chapter are detailed below:

- **Chapter 2:** *Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain):* The aim of this chapter is to provide a detailed description of the archaeal communities inhabiting the Monegros Desert lacustrine system, with special focus on the evolution of the archaeal groups along the salinity gradient. Although the low archaeal abundance at low salinities does not allow the assessment of potential spatial processes influencing their assembly, the temporal patterns of the studied groups will be assessed in terms of colonization and extinction dynamics.

- **Chapter 3:** *Ecological and metabolic thresholds in the bacterial, protist and fungal microbiome of ephemeral saline lakes (Monegros Desert, Spain):* This chapter aims to provide a detailed description of the bacterial and microeukaryal microbiome inhabiting at the different salt ranges on the studied system. The characterization of the microbial groups inhabiting each salt range will potentially allow to unveil ecological and metabolic thresholds associated to the studied salt range, at the time that allow to test the potentially differentiated response according to its osmoadaptation mechanisms.
- **Chapter 4:** *A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly:* This chapter has two complementary objectives. First is to develop a randomized approach to determine the presence of trait clustering and overdispersion patterns along environmental gradients, which can be linked to the effect of different assembly mechanisms at the metacommunity level. The second objective is to apply the clustering approach to the molecular dataset of prokaryotes inhabiting the studied system in order to determine consistent thresholds along the studied salt gradient, which potentially determine a change on the dominant community assembly mechanism based on the genomic traits of the inhabiting communities.
- **Chapter 5:** *Understanding stochastic and deterministic assembly processes in microbial communities along temporal and environmental gradients:* The aim of this chapters is to reach a mechanistic understanding of how deterministic and stochastic assembly processes interact to determine communities composition on the studied system. The prevalence of both processes and their interaction will be assessed, along the salt gradient but also from a dynamic perspective, in order to link the variable influence of both processes on determining the observed alpha- and beta- diversity patterns for both bacterial and microeukaryal groups.
- **Chapter 6:** *Absence of stress-gradient promoted facilitation and a competition decrease in the microbial communities of ephemeral saline*

lakes: This chapter aims to explore interspecies biological interactions within the studied metacommunity (including the three microbial domains), with special focus on their behavior along a stress gradient associated to the increasing saline conditions. Specifically, the main objective of this chapter is to test whether the stress-gradient hypothesis, i.e. the increase in the frequency of positive interactions over negative ones as the stress increases, applies to microbial communities reproducing previously observed dynamics in plant communities and experimental setups.

- **Chapter 7:** *Increasing saline stress reduces allochthonous microbial colonization in ephemeral ponds, driving the transition from habitat generalists to specialists*: The main objective of this chapter is to test the potential interaction by dispersal processes between the microbial communities inhabiting water and sediment compartments within the studied area. For this, we performed a joint analysis of the microbiome of both compartments in order to determine the cross contribution between both compartments and the potential influence of the salinity gradient in modulating this interaction.

A network diagram consisting of white circular nodes connected by thin white lines, forming a complex web. The nodes are distributed across the upper half of the image, with a higher density of connections in the top right corner. The background is a solid teal color.

CHAPTER TWO

Mateu Menéndez-Serra, Vicente J. Ontiveros, Xavier Triadó-Margarit, David Alonso and Emilio O. Casamayor (2020). Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain). *FEMS Microbiology Ecology*, 96, fiae019. DOI:10.1093/femsec/fiae019.

This chapter can be found in its publication format in the *Published Chapter* section of the present thesis. Likewise, the bibliography of this chapter can be found on the unified *Bibliography* section.

Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain)

Abstract: We characterized the rich Archaea microbiome of shallow inland lakes (Monegros Desert, NE Spain) by 16S rRNA gene tag sequencing covering a wide salinity range (0.1%–40% w/v) along 3 years. Up to 990 operational taxonomic units (OTUs; >97% identity) were detected allocated in 14 major archaeal phyla and heterogeneously distributed along the salt gradient. Dynamics and idiosyncratic ecological distributions were uncovered for the different phyla. A high genetic richness was observed for Woesearchaeota and Pacearchaeota (>370 OTUs each), followed by Halobacteria (105), Nanohaloarchaeota (62) and Thermoplasmata (19). Overall, the distribution of genetic richness was strongly correlated with environmental niche amplitude, but not with occurrence. We unveiled high occurrence for a very rich Woesearchaeota assemblage, and an unexpected positive correlation of Pacearchaeota abundance with salinity at >15% dissolved salt content. The estimated dynamic behaviour (temporal ‘turnover’ rates of presence/absence data) unveiled Thaumarchaeota and Halobacteria as the most dynamic groups, and Aenigmarchaeota and Thermoplasmata as the most stable. The DPANN Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate rates, suggesting higher resilience to environmental perturbations. A rich and dynamic Archaea microbiome was unveiled, including unseen ecological traits for relevant members of the still largely unknown DPANN group, supporting a strong ecological differentiation between Pacearchaeota and Woesearchaeota.

3.1 Introduction

Inland lakes in endorheic areas are very dynamic ecosystems globally distribute in semi-arid and arid regions of the world (Williams, 1996). Most shallow inland lakes show wider ranges of salinities, environmental conditions, and salt compositions than coastal areas (Williams, 1996). Numerous microbial ecology studies have been carried out on diverse terrestrial saline environments in all the continents (North- and South-America (e.g. Demergasso et al. 2008; Dillon et al. 2009), Australia (Narasingarao et al., 2012), European Mediterranean areas (Gasol et al., 2004; Ventosa et al., 2014a), Asia (Wu et al., 2006), Africa (Cavalazzi et al., 2019), Antarctica (Williams et al., 2014), just to mention a few among many others). The role of salinity as a physiological stressor has also promoted the study of idiosyncratic ecophysiological adaptations along salinity gradients in simplified ecological communities (Pedrós-Alió et al., 2000). Terrestrial saline aquatic environments (both inland and coastal) are very good natural model systems and have been traditionally associated to highly specialized planktonic organisms. At the most saline ends, dominance of a few groups has been reported, i.e., *Haloquadratum walsbyi*, *Salinibacter ruber* and the recently uncovered Nanohaloarchaea (Ghai et al., 2011; Narasingarao et al., 2012; Ventosa et al., 2015). At intermediate salinities, different studies show high levels of heterotrophic activity (Gasol et al., 2004), planktonic algae accumulation (Estrada et al., 2004), and presence of largely unknown organisms (Demergasso et al., 2008; Triadó-Margarit and Casamayor, 2013). Among them, inland lakes are prone to contain new groups of Halobacteria and poorly characterized uncultured Archaea (Casamayor et al., 2013; Narasingarao et al., 2012).

Several studies in the last decade have improved the picture on the Archaea domain, with higher diversity and environmental distribution than historically considered (Auguet et al., 2010; Rinke et al., 2013). There are however important gaps still on the current knowledge of the Archaea ecology even in apparently well studied habitats where archaea were predominant organisms, such as high saline systems. Recently, the monophyletic superphylum DPANN has been shown to contain a large number of widespread taxa most of them missing cultured counterparts, such as Nanohaloar-

chaeota, Pacearchaeota, Woesearchaeota, Aenigmarchaeota, Micrarchaeota, Diapherotrites, Parvarchaeota, and Nanoarchaeota (Castelle et al., 2015; Ortiz-Alvarez and Casamayor, 2016; Rinke et al., 2013). Because of the difficulties in culturing and the still limited application of both powerful single cell genomic analysis (Dal Molin and Di Camillo, 2018) and metagenomics approaches (Narasimarao et al., 2012) in archaea, the biology and ecology of many of these groups are still unknown. In addition, the detailed study of temporal dynamics in archaeal ecological communities has been rarely carried out (Galand et al., 2010; Hugoni et al., 2013) missing substantial information on the ecology of the third domain of life.

Within this context, we explored community patterns and temporal dynamics of the aquatic archaeal microbiome in the Monegros Desert (NE Spain). The Monegros ecosystem contains one of the largest sets of inland saline lakes in Europe and constitutes a unique landscape of great ecological value, easy to reach and sample (Casamayor et al., 2013). Physico-chemical parameters strongly change with time in the same aquatic spot offering a wide repertoire of heterogeneous environmental conditions within a very narrow geographic distance of a few kilometers. A preliminary study already evidenced the high microbial novelty present in Monegros, including novel archaea (Casamayor et al., 2013). In the present study, we aimed to unveil temporal dynamics, ecological patterns and environmental preferences for the most unknown Archaea covering different inter-annual dry and wet periods. We added a quantitative temporal dimension to the analysis by applying a community ecology method based on temporal changes in the absence/presence of local populations. This method applies a mathematical estimation of ‘colonization’ and ‘extinction’ rates (Ontiveros et al., 2019) showing different ecological dynamics for the different groups. The range of habitat gradients occupied and the estimated dynamical behavior, unveiled community level traits and patterns unseen to date for some of the most intriguing uncultured archaeal groups.

3.2 Experimental Procedures

The Monegros Desert is an endorheic platform (c. 400 km²) located in the Central Ebro Basin (NE Spain, 41°42'N, 0°20'W, 328 m a.s.l.), with 149 scattered depressions ranging in size from <2 to 239 ha, some of them hosting hypersaline wetlands. This is one of the most arid regions in Europe (mean annual rainfall 350 mm, mean annual temperature 14.9 °C, and mean reference evapotranspiration 1255 mm yr⁻¹ (Faci González and Martínez-Cob, 1991)), and unique for the combination of geological, edaphic, mineralogical and hydrological characteristics (Mees et al., 2011; Pueyo Mur and Inglés Urpinell, 1987; Samper-Calvete and García-Vera, 1998). We carried out monthly sampling in 14 different ponds (Table 3.1) covering different dry-wet periods with a total of 128 water samples analysed (see details in Table A.1). Conductivity, temperature, pH, dissolved oxygen and redox potential were measured in situ using an HQ40d multiparameter meter (Hach, Loveland, CO, USA). For chlorophyll a (Chl a) analysis 1L water was filtered through 47-mm-diameter Whatman GF/F filters (0.7 µm nominal particle retention), stored in the dark, and kept frozen. Chl a concentration was determined in acetone extracts by spectrophotometry (UV-2401PC, ultraviolet-visible Spectrometer; Shimadzu) as previously reported (Bartrons et al., 2012). Salinity was measured in situ by a hand salinity refractometer (Atago S-28E, Japan). The water column height was measured at the deepest part of the ponds. Substantial changes in water column height, temperature, salinity and other environmental variables were observed both among ponds and inside the same pond along the hydrological period. Detailed information about environmental parameters and sampling date and sites is available in Supplementary Table A.1.

3.2.1 DNA Extraction and sequencing

For DNA analyses, water samples were pre-filtered in situ through a 50 micron-pore-size net, to retain large zooplankton and algae, and then 100—500 mL were subsequently filtered on 5 and 0.2-micron pore-size polycarbonate membranes (47 mm diameter), respectively. The 0.2 µm PC

Table 3.1: Range of environmental parameters measured in the different salt lakes in Monegros.

Lagoon	Salinity (% w/v)	Cond. (mS/cm)	Temp. (°C)	pH	OD (mg/l)	Redox (mv)	Chla (μ g)
Guallar	2.8 – 23	26.2 – 197.6	2 – 26.5	7.6 – 8.4	8.8 – 12.8	145.5 – 350	0.9 – 10
Pinol	1.2 – 8.4	14.4 – 103	5 – 25.1	8 – 8.8	8.9 – 13.3	187.1 – 315	0.9 – 27.9
Muerte	2 – 32	20.3 – 207.4	7.5 – 32	7.3 – 8.61	7.4 – 12.8	224 – 312	0.9 – 5.2
Camaron	1.2 – 15.4	14.2 – 164.7	2.3 – 29.5	8.0 – 9.0	8.8 – 14.1	220 – 305	0.5 – 46
Pez	1.1 – 13.8	8.5 – 165.5	2 – 31	7.8 – 8.42	8.4 – 13.1	200 – 320	1.1 – 6.7
Rollico	0.8 – 5.8	5.6 – 77.6	8.7 – 21.5	8 – 8.8	9.1 – 12.8	205 – 345	0.7 – 13.7
Rebollon	1 – 2.4	5.1 – 27.2	7.5 – 14.8	8.1 – 9.1	10.4 – 14.6	225 – 340	1 – 27.1
Pito	1.2 – 28	12 – 94.1	1.5 – 34.5	7.5 – 8.3	7.5 – 12.7	240 – 335	1.1 – 6.6
Salineta	6.9 – 40	66.3 – 167.9	7.4 – 34	7.7 – 9.2	7.8 – 15.7	188.6 – 315	2.4 – 237.4
Playa	2.1 – 21	26.3 – 194.5	4 – 3	8 – 8.8	7.3 – 12.4	144 – 312	0.4 – 29.6
Amarga Alta	0.9 – 1.7	3.6 – 15.6	3 – 15	7.9 – 8.4	10.9 – 12.5	248 – 326	0.5 – 6.8
Amarga Baja	0.9 – 1.8	4.3 – 15.6	1.5 – 11.2	7.9 – 9.2	10.5 – 13.1	242 – 339	0.7 – 11.4
Salobral	2.9 – 3.8	31.3 – 47.1	5.9 – 30.3	8.8 – 9.5	8.1 – 14.2	122.6 – 341	1.7 – 15.5
Balsa Molinos	0.1 – 0.3	0.7 – 0.9	8 – 17.4	8.3 – 9.1	9.8 – 12.1	195 – 240	3 – 9.7

membranes were stored in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose), enzymatically digested and phenol extracted (phenol—chloroform—isoamyl alcohol; 25:24:1, v/v/v) as previously reported (Ferrera et al., 2004). DNA purification and concentration was carried out with Amicon®Ultra 4 Centrifugal Filter Units —100 000 NMWL (Millipore). DNA extracts were quantified using Quant-it assays in a Qubit™ fluorometer (Invitrogen). Sequencing with the Illumina 2x250 MiSeq platform was carried out on PCR amplicons of the 16S rRNA gene (V4 region, primers set F515806, Caporaso et al. 2010) according to genomic core facilities, protocols and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>).

3.2.2 Bioinformatic processing and genetic analyses

Raw sequences were processed using UPARSE (Edgar, 2013). The total number of sequences prior to quality filtering was 14,445,263 and 70.6% of them had >Q30. After merging of read pairs, the dataset was filtered by setting a maximum number of expected errors of 0.25 (Edgar and Flyvbjerg, 2015), and chimeric reads were detected and filtered out by UCHIME following both de novo and reference-based analyses (against the ‘Gold’ reference database available for ChimeraSlayer). Filtered sequences were clustered into Operational Taxonomic Units (OTUs) at 97% identity. The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012), using the SILVA 128 database (Gurevich et al., 2013). OTUs were filtered

by their alignment quality score and sequences classified as mitochondria and chloroplasts were excluded. A final filtering by library size keeping only samples with >10,000 reads was carried out. The resulting OTU table consisted of 7,241 prokaryotes OTUs (6251 bacterial OTUs and 990 archaeal OTUs), corresponding to a total of 3,374,654 reads (of which 54,807 were assigned to Archaea) for 128 samples. Data was normalized by rarefying the reads of all samples to a minimum threshold of 10,000 reads/sample. For community related analysis, rarefaction at the same depth was repeated 100 times in order to avoid the loss of less abundant OTUs, and the resulting rarefactions were unified on an average rarefied OTU table. Genetic distances between OTUs were calculated using the 'calc_distmx' command in the USEARCH pipeline. Reference sequences were included to properly differentiate *Candidatus Pacearchaeota* (i.e., GenBank accession numbers MFWP01000008, MNVW01000021, MNVX01000029, and MNVY01000002, Anantharaman et al. 2016) from *Candidatus Woearchaeota* (accession number MNWW01000056, Nawrocki et al. 2018). The resulting distance matrix was visualized by Non-Metric Multi-Dimensional Scaling (NMDS) analysis run with MetaMDS functions in Vegan package. We estimated genetic diversity by calculating the nucleotide diversity using the R package 'Pegas' (<https://cran.r-project.org/web/packages/pegas/index.html>). The phylogenetic tree shown in supplemental Figure S1, is an ARB maximum parsimony tree provide by default (<http://www.arb-home.de/>) with 16S rRNA gene amplicon sequences added using the tool “add species to existed tree / ARB-parsimony tool interactive”, providing only very rough information because of the short read length.

The complete raw data set is available in GenBank under BioProject record ID PRJNA429605, and GenBank accession numbers for highlighted OTUs are shown in Supplemental materials (Table A.2, Figure A.1).

3.2.3 Statistical analysis and calculation of colonization and extinction rates

Statistical analyses were run in R (<http://www.r-project.org>). Species richness, evenness, and diversity indexes were calculated using “Vegan” package version 2.4-5 (Oksanen et al., 2016). The environmental amplitude (range

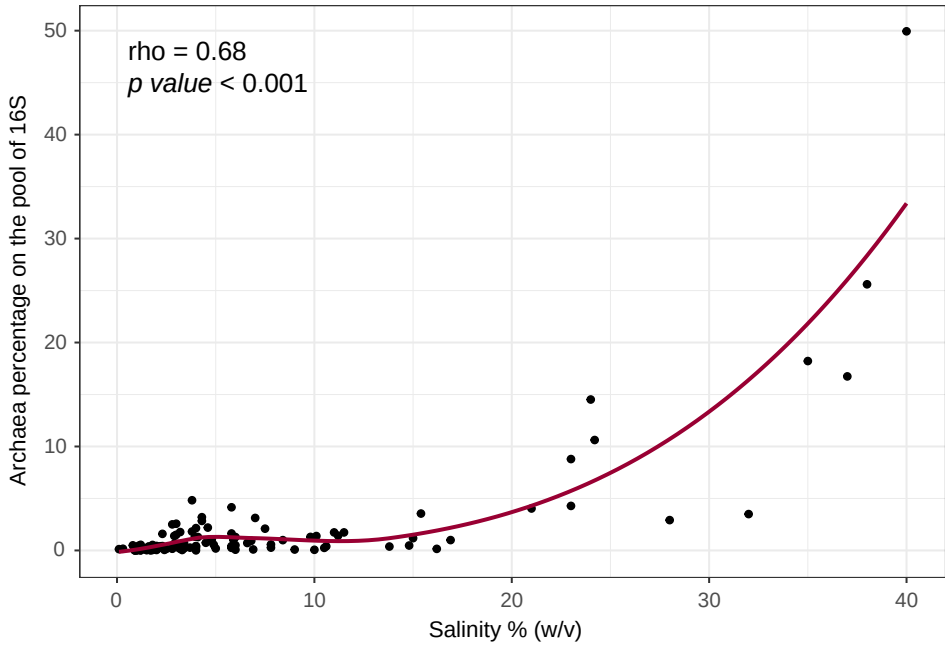


Figure 3.1: Relationship between archaeal percentage in the prokaryotic 16S rRNA amplicon mixture and salinity (% w/v) for the complete dataset analyzed.

of environmental conditions occupied) for the different archaeal lineages, was calculated according to salinity, pH, redox potential, dissolved oxygen, and Chl *a* data, weighting its variance with the relative abundance of the lineage. We avoided the use of variables that co-varied with salinity such as conductivity (Spearman, $\rho = 0.98$ and $p\text{-value} < 0.001$) and temperature (Spearman, $\rho = 0.50$ and $p\text{-value} < 0.001$). Spearman’s Rank order correlation coefficient was used to test monotonic associations between numeric variables. The estimated temporal “turnover” rates were calculated for each archaeal lineage based on the absence/presence of local populations for consecutive sampling times (“extinction / colonization” rates, respectively). “Colonization” and “extinction” rates were calculated using the “island” package on R (<https://CRAN.R-project.org/package=island>; Ontiveros et al. (2019)) and provide an effective view on the dynamic behavior of the different archaeal populations in time (‘turnover’ rate). The ‘turnover’ rates were expressed as day^{-1} .

Table 3.2: OTU richness, occurrence and abundances for the archaeal lineages found in Monegros.

Archaeal lineage	OTUs	Occurence (%)	Max. abundance (%)
EURY: Halobacteria	105	105 (82,03%)	22.08
DPANN: Woesearchaeota	402	123 (96,09%)	3.34
DPANN: Nanohaloarchaeota	62	63 (49,22%)	23.94
DPANN: Pacearchaeota	376	112 (87,5%)	4.98
EURY: Thermoplasmata	19	30 (23,44%)	2.99
DPANN: Aenigmarchaeota	5	24 (18,75%)	0.35
TACK: Thaumarchaeota	8	44 (34,38%)	0.23
Archaea unclass.	6	14 (10,94%)	0.27
Asgard: Lokiarchaeota	1	5 (3,91%)	0.04
EURY: Methanomicrobia	1	3 (2,34%)	0.04
EURY: Unclass.	2	5 (3,91%)	0.01
EURY: WSA2	1	2 (1,56%)	0.02
EURY: Archaeoglobi	1	3 (2,34%)	0.01
EURY: Miscellaneous Group (MEG)	1	1 (0,78%)	0.01

3.3 Results

3.3.1 Archaeal abundance and diversity

The relative abundance of Archaea in the 16S rRNA gene amplicon mixture ranged between 0.01 and 50% of total prokaryotes. Archaea abundance consistently increased along the salinity gradient, mainly at salinities >20% (w/v), reaching the highest relative abundances at the highest salinity end (Figure 3.1). For salts concentrations <20%, relative abundances were <5% in all cases.

The genetic richness of the Archaea planktonic microbiome was notorious. We observed the presence of 14 different lineages, with occurrences (percentage of samples in which the group was detected) between 96% and 0.8% (Table 3.2). As a whole, the OTUs richness was higher for Woesearchaeota (402 OTUs) and Pacearchaeota (376 OTUs), followed by Halobacteria (105 OTUs) and Nanohaloarchaeota (62 OTUs). The highest values of abundance in the amplicon mixture were found for both Nanohaloarchaeota (24%) and Halobacteria (22%). Woesearchaeota and Pacearchaeota did not reach relative abundances >5% for any of the samples, although they were widely distributed being present in >87% of the total dataset. Halobac-

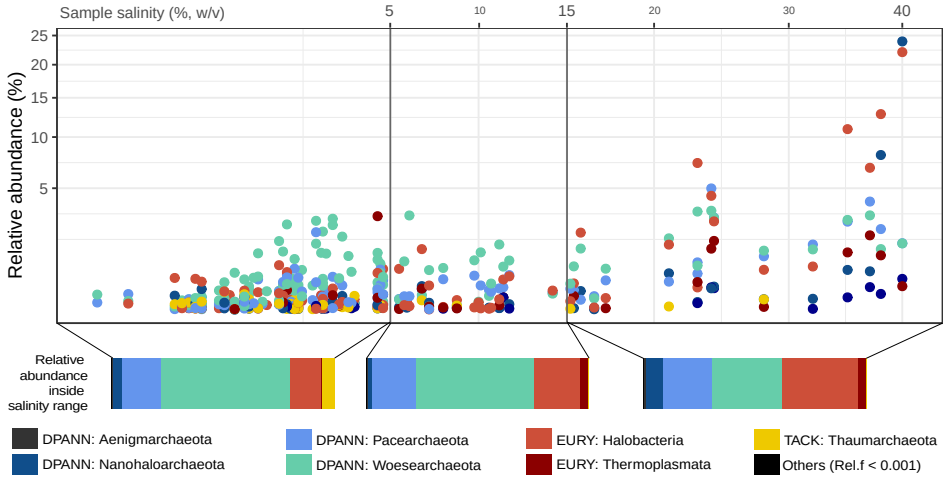


Figure 3.2: Relative abundances of the main archaeal lineages along the salinity gradient, and averaged values for the three traditional saline ranges, i.e. $>5\%$ (saline), $5\text{--}15\%$ (hypersaline), and $>15\%$ (extreme hypersaline), respectively.

teria was also found in a large percentage of samples (82%), followed by Nanohaloarchaeota (49% occurrence). Minor groups were Thermoplasmata, Thaumarchaeota and Aenigmarchaeota, among others (Table 3.2). Only six reads (1 OTU) matched methanogens (i.e., *Methanomicrobia*, genus *Methanohalophilus*).

3.3.2 Archaeal assembly along the salinity gradient

We selected the seven most abundant lineages and with occurrences $>15\%$, for a more detailed downstream analysis. These groups showed a distinctive distribution along the salinity gradient (Figure 3.2). At salinities below 15%, the most abundant groups were mainly Woesearchaeota, and in one case we detected noticeable concentrations of both Thermoplasmata (3% relative abundance in Lake Muerte on January 2013, salinity 5.8%, temperature 7.8°C; OTU 316 in Table A.2) and Pacearchaeota (2% relative abundance in Lake Salobral on July 2013, salinity 3.8%, temperature 29.6°C). Conversely, at $>15\%$ salinity the archaeal assemblages were mostly dominated by Nanohaloarchaeota (up to 24% in Lake Salineta, October 2013, salinity 40%, temperature 18.8°C), Halobacteria, and

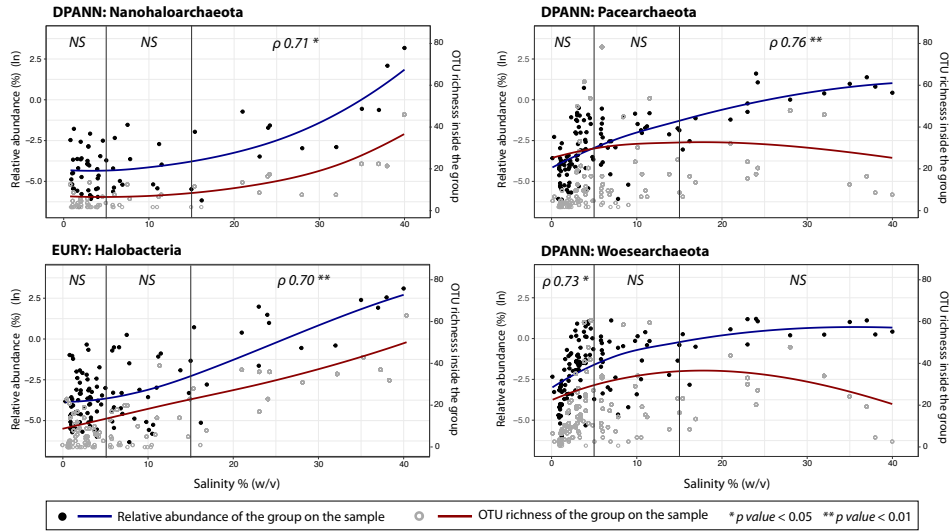


Figure 3.3: Relative abundances and OTU richness for the four most abundant groups (Woesearchaeota, Pacearchaeota, Nanohaloarchaeota, Halobacteria) along the salinity gradient. Correlations shown for the three traditional saline ranges, i.e. <5% (saline), 5%–15% (hypersaline), and >15% (extreme hypersaline), respectively. Spearman's rho correlation and p-value are shown when significant; NS = non-significant. The smooth fitted curves from these parameters were obtained by means of Local Regression (loess) approach.

Pacearchaeota. Finally, Thaumarchaeota (ammonia-oxidizing archaea, AOA) showed a maximum abundance of 0.23% at marine like salinities (close to 3% w/v) and were mostly detected at salinities up to 10% (8 OTUs, accession numbers MH028008 to MH028015) indicating a strong selection by local environmental conditions on this functional group.

Overall, salinity was the environmental variable which explained most of the variations in the relative abundances of the different archaeal populations and not major effects were observed for the additional environmental/biological parameters measured (p-value >0.05 for pH, redox potential, dissolved oxygen, Chl a concentration, and bacterial diversity). We found a strong correlation between salinity and relative abundances of Pacearchaeota (Spearman, rho 0.59, p-value <0.001) and Woesearchaeota (Spearman, rho 0.65, p-value <0.001). This correlation was also significant for Nanohaloar-

chaeta (Spearman, ρ 0.36, p -value <0.01) and Halobacteria (Spearman, ρ 0.41, p -value <0.001). The robustness of the correlations increased when the whole salinity gradient was divided into the three main salinity ranges based on both physicochemical and biological parameters (Imhoff, 2001; Kushner, 1978; Por, 1972), i.e., $<5\%$ (saline), $5\text{--}15\%$ (hypersaline), and $>15\%$ (highly hypersaline), and some unseen trends were subsequently unveiled (Figure 3.3). Nanohaloarchaeota and Halobacteria showed stronger correlations (c. ρ 0.70) at the most saline end, whereas the relationship was non-significant at $<15\%$ salinity. Woesearchaeota only showed a significant positive relationship at salinities $<5\%$. Conversely, relative abundances of Pacearchaeota showed an unexpected positive correlation with salinity in the highly hypersaline range ($>15\%$ salinity, Figure 3.3). These results unveiled that these two highly unknown taxa tend to behave differently in front of salinity. Candidatus Pacearchaeota and Woesearchaeota were the richest Archaea taxa in Monegros and showed the highest genetic variation along the salinity gradient in comparison with the remaining Archaea populations (Figure 3.4). Despite the fact that both had been grouped within the Deep-sea Hydrothermal Vent Euryarchaeota Group 6 (DHVEG-6) in former taxonomic arrangements, Pacearchaeota and Woesearchaeota appeared clearly separated from each other in the NMDS ordination according to its nucleotide sequence composition (Figure 3.4).

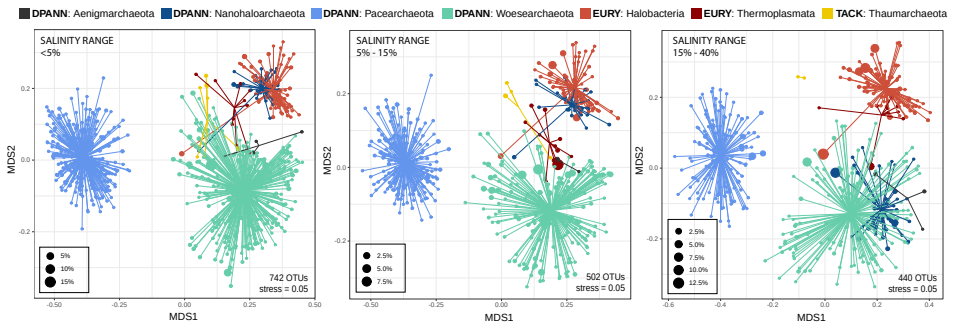


Figure 3.4: NMDS representation based on the genetic distances among OTUs in the three main salinity ranges ($>5\%$ saline, $5\text{--}15\%$ hypersaline, and $>15\%$ extreme hypersaline). Each OTU from the same lineage is connected in the two-dimensional space through the centroid. Point sizes show relative abundances.

A closer look to the most saline end of the gradient ($>35\%$ salinity) showed that Halobacteria at these salinities were spread in 12 already known genera, whereas most of the Nanohaloarchaeota were only distantly related to any previously reported counterpart (supplemental Figure A.1, with available individual accession numbers shown). Interestingly, two of the most abundant taxa in the most saline sample (Lake Salineta, October 2013, 40% salinity) were the Halobacteria OTU 320 (97% identity to *Halonotius pteroides*) and the Nanohaloarchaeota OTU 289 (100% identical to *Candidatus Haloredivivus* G17 AGNT01000024) previously found in Santa Pola solar salterns in the Spanish Mediterranean coast (Ghai et al., 2011), aprox. 450 km away from Monegros.

3.3.3 Archaeal ecology and dynamic behaviour

Community ecology analyses provided new insights and unveiled unseen ecological trends for the different Archaea groups. As a whole, group occurrence was strongly correlated with OTUs richness (Spearman, ρ 0.96, p -value <0.005), i.e., the highest occurrence was found for the richest groups. In addition, the distribution of genetic richness (i.e., nucleotidic diversity) for the seven major archaeal phyla was strongly correlated with environmental amplitude (Spearman, ρ 0.89, p -value <0.05 , Figure 3.5, panel B) but not with lake occurrence (Figure 3.5, panel A). Aenigmarchaeota had the highest environmental amplitude (i.e., larger range of habitat gradients occupied) despite the fact that only 5 OTUs were detected for this phylum, and showed the lowest occurrence (19%). Conversely, Nanohaloarchaeota, Halobacteria and Thaumarchaeota had the bulk of their abundances within a restricted environmental niche (Figure 3.5, panel B), despite having a medium-high occurrence (between 35% and 85%). The analysis of OTU richness vs. evenness also showed interesting trends (Figure 3.5, panel C); Paccarchaeota and Woesearchaeota showed the highest richness and evenness, whereas at the other end, Aenigmarchaeota, Thermoplasmata, Thaumarchaeota, and Nanohaloarchaeota were low richness-low evenness groups. Halobacteria was the only group with low richness-high evenness structure (i.e., relatively low number of OTUs and with similar population density). Finally, the estimated dynamic behaviour ('turnover' rates) for the

main archaeal lineages showed again important differences among groups (Figure 3.5, panel D). The most dynamic groups (the highest 'colonization/extinction' turnover) were Thaumarchaeota and Halobacteria. The most stable taxa (the lowest turnover) were Aenigmarchaeota and Thermoplasmata. The DPANN groups Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate turnover rates between the two ends.

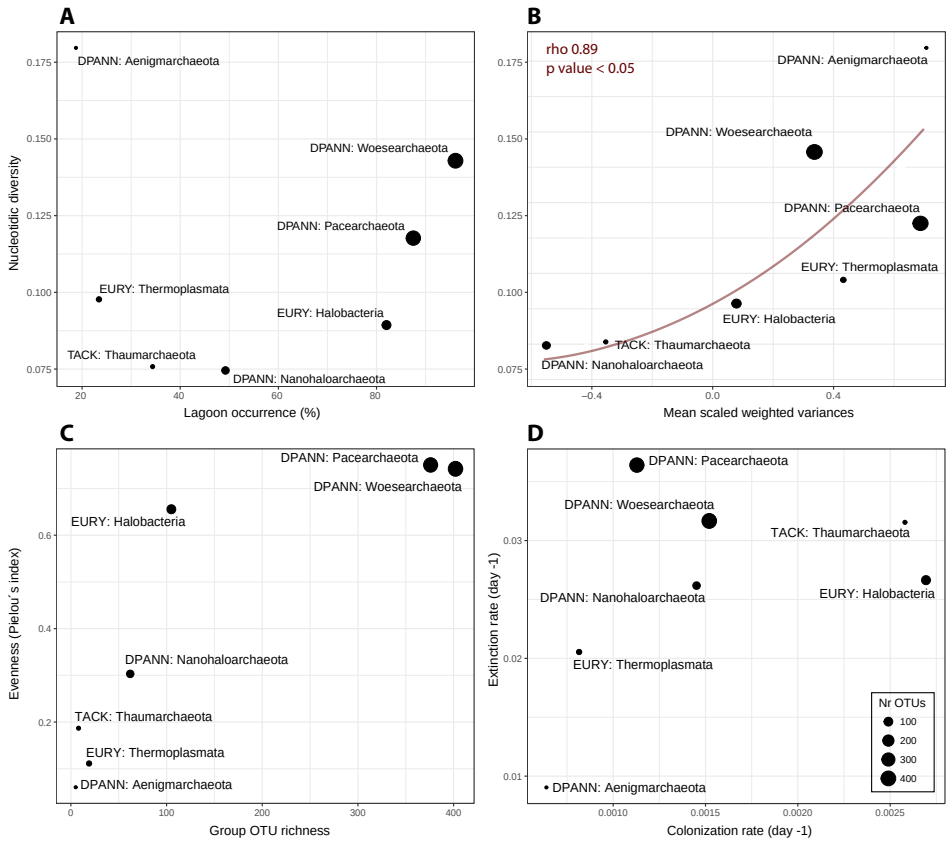


Figure 3.5: Community ecology descriptors for the main phyla of the Archaea assemblage in Monegros saline lakes. Nucleotide diversity related to both lagoon occurrence (A) and environmental amplitude (B). Spearman's rho correlation shown when significant. Evenness (Pielou's index) based on abundance distribution related to group OTUs richness (C). 'Colonization and extinction' turnover rates (day⁻¹) based on the absence/presence of the local communities in consecutive sampling times (D). Point sizes show OTU richness.

3.4 Discussion

Comprehensive studies carried out in the last years have provided a better picture on the Archaea ecology and diversity (e.g., Auguet et al. 2010; Rinke et al. 2013). There are however important gaps still in the current knowledge of the environmental Archaea microbiome. For instance, the widespread DPANN group has been recently arranged (Castelle et al., 2015; Spang et al., 2017), but little is known on the ecology, dynamics, and environmental preferences of these archaea. Our study added relevant information analysing the ecological niche, range of environmental preferences, and temporal dynamics of uncultured DPANN and other relevant archaea, using a detailed temporal survey in a high dynamic set of inland salt lakes.

We observed higher archaeal relative abundance at the most extreme salinity end, mostly of Halobacteria (superphylum Euryarchaeota) but also of the DPANN Nanohaloarchaeota. The importance of Halobacteria (haloarchaea) at extreme salinities have been recurrently reported (Oren, 2002). However, the role and distribution of Nanohaloarchaeota has been unseen for a long time (Narasimarao et al., 2012). Its recent detection in diverse hypersaline environments, has arisen a high interest for both plankton and benthic Nanohaloarchaeota populations (Andrei et al., 2015; Cowan et al., 2015; Podell et al., 2014; Ventosa et al., 2015). In our study, we observed similar relative abundances of Nanohaloarchaeota and Halobacteria at the most saline end, although with substantial differences in community structure and community ecology dynamics. Nanohaloarchaeota showed a more limited environmental breadth, being less abundant than Halobacteria at lower salinities, and lower colonization rates. More studies are needed to confirm whether or not this observation indicates a higher specialization of the DPANN Nanohaloarchaeota to the most extreme saline conditions compared to a more versatile behavior of Halobacteria. Recent discoveries on a symbiotic lifestyle for Nanohaloarchaeota (Hamm et al., 2019) indicate that it is necessary to apply an integrated view on the whole microbial community (including bacteria and eukarya) to better understand Nanohaloarchaeota ecology.

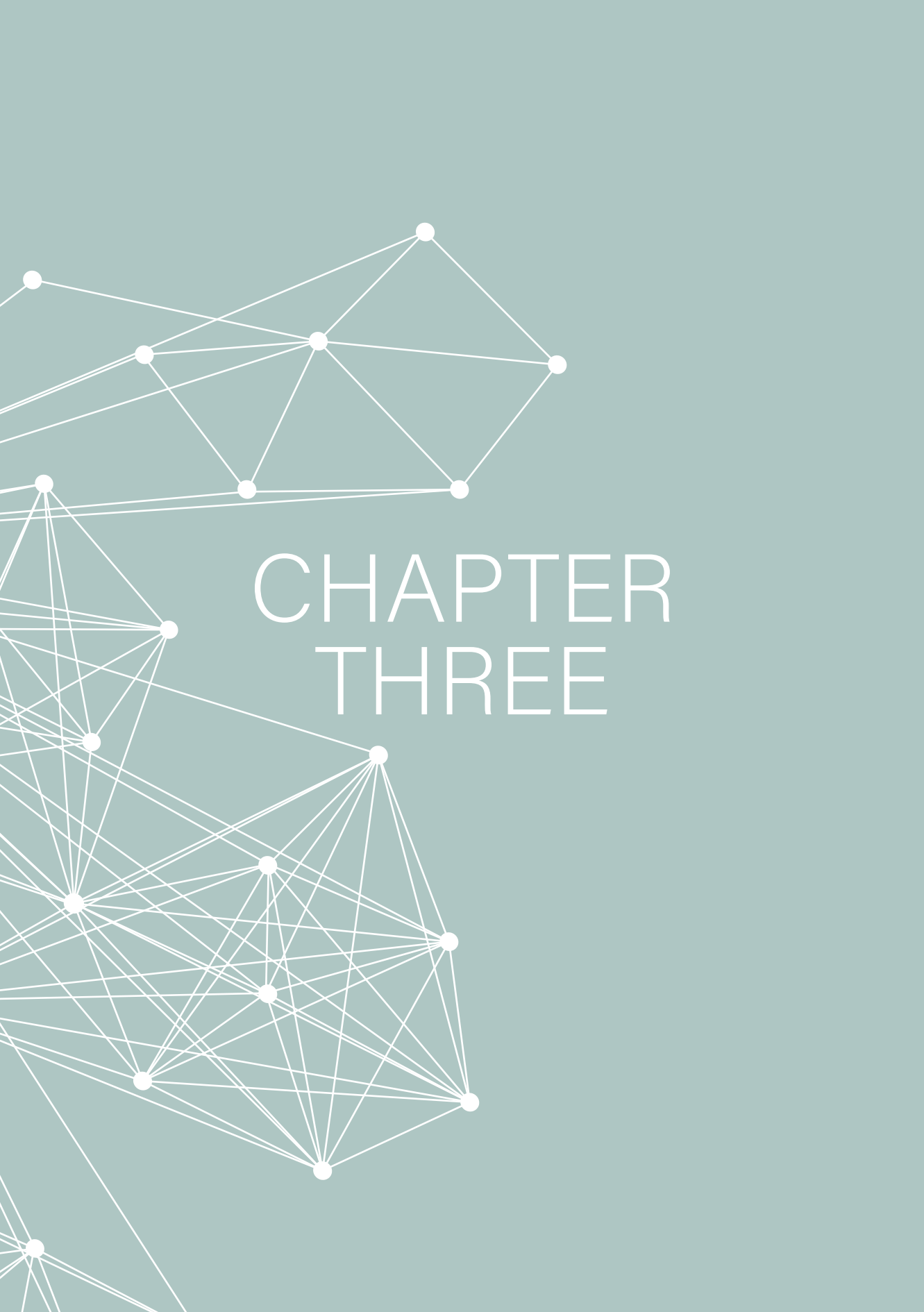
A symbiotic/parasitic lifestyle has also been proposed for the other two important DPANN groups (in terms of relative abundance and occurrence)

in the Monegros Archaea microbiome: Woesearchaeota and Pacearchaeota. Woesearchaeota and Pacearchaeota (both formerly named Deep-sea Hydrothermal Vent Euryarchaeota Group 6, DHVEG-6) have been recently proposed as candidate phyla belonging to the DPANN superphylum (Castelle et al., 2015). Since then, they have been uncovered as major components of the archaeal assemblages in a large repertoire of environments, and the need of a better understanding of their biology, ecology and distribution has stimulated research in recent years (Castelle et al., 2015; Liu et al., 2018; Ortiz-Alvarez and Casamayor, 2016). The lack of an unequivocal phylogenetic split between Woesearchaeota and Pacearchaeota has led in some cases to overlapping analysis (Lipsewiers et al., 2018; Liu et al., 2018) and blurred ecological differentiation between them. The environmental niches of these two phyla are known to be large and they are widely distributed in saline aquatic environments, mainly in sediments, microbial mats and plankton, but also in hot springs and freshwater (Kan et al., 2011; Pachiadaki et al., 2011; Robertson et al., 2009). Pacearchaeota has only been rarely detected in marine plankton (Castelle and Banfield, 2018), which matches the observed trend in Monegros where higher relative abundances of Pacearchaeota were found at higher salinities. In the studied saline lakes, these two phyla showed the highest OTU richness, evenness, sites occurrence and extinction rates among Archaea. All these traits suggest an ecological strategy based on fast colonization but low persistence along the whole salinity gradient for the different OTUs shaping these two groups. As previously mentioned, new insights in the metabolisms of DPANN archaea brought by metagenomics and single-cell genomics, suggest that small genomes, small cell sizes, and limited metabolic variability are common characteristics. Thus, the lack of core biosynthetic pathways has led to conclude that most of the DPANN archaea depend on other microorganisms –as a symbionts or episymbionts– to carry out basic metabolic functions (Castelle and Banfield, 2018). In a recent study in alpine lakes, relative abundances of Pacearchaeota and Woesearchaeota were observed to be positively correlated with phylogenetic diversity of bacterial communities, respectively (Ortiz-Alvarez and Casamayor, 2016). This correlation was however not observed in Monegros. If this correlation is expected to be consistent and somehow related to

potential bacterial symbionts, probably their high occurrence along the salinity gradient and their large environmental amplitude may have masked the detection of this potential signature in Monegros. Certainly, this point deserves further and more detailed studies considering more integrated analyses of the whole microbial community.

For the first time, our study quantifies the in situ dynamic behavior of uncultured archaea. This new perspective needs of experimental tests to be further substantiated. The more 'dynamic' taxa were Thaumarchaeota (limited to salinities below 10%, in agreement with previous studies (Fernández-Guerra and Casamayor, 2012)) and Halobacteria (with the bulk of its abundance at extreme salinities), and showed their optimal within a restricted environmental niche suggesting these are groups with a higher cell sensitivity to salinity changes. Conversely, the DPANN groups Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate turnover rates. Pacearchaeota and Woesearchaeota showed the widest environmental breadth, the highest richness, and the highest genetic variability. These results, when contextualized in a highly dynamic system with quick changes in community composition mainly driven by changes in salinity (Triadó-Margarit et al., 2019), suggest a higher resilience to environmental perturbations.

Overall, the study carried out in inland salt lakes of Monegros showed a rich and dynamic Archaea microbiome including relevant members of the still largely unknown DPANN group. We unveiled ecological properties for some of the most enigmatic DPANN members as Nanohaloarchaeota, Pacearchaeota, Woesearchaeota, and Aenigmarchaeota, and provided further evidences to support strong ecological differentiation between Pacearchaeota and Woesearchaeota. Further studies both in situ and in combination with traditional microbiological approaches, should help to improve our still very limited understanding on the biology and ecology of the monophyletic superphylum DPANN, and on the ecology of inland saline systems in general.

An abstract geometric pattern composed of white lines and dots is overlaid on a solid teal background. The pattern consists of several interconnected clusters of points. Some clusters form simple polygons, while others are more complex, with multiple lines radiating from a central point to several peripheral points, creating a star-like or web-like structure. The dots are small, solid white circles, and the lines are thin, white, and straight.

CHAPTER THREE

Mateu Menéndez-Serra, Xavier Triadó-Margarit Emilio O. Casamayor (2021). Ecological and metabolic thresholds in the bacterial, protist, and fungal microbiome of ephemeral saline akes (Monegros Desert, Spain) *Microbial Ecology*, DOI:10.1007/s00248-021-01732-9.

This chapter can be found in its publication format in the *Published Chapter* section of the present thesis. Likewise, the bibliography of this chapter can be found on the unified *Bibliography* section.

Ecological and metabolic thresholds in the bacterial, protist and fungal microbiome of ephemeral saline lakes (Monegros Desert, Spain)

Abstract: We studied the 16S and 18S rRNA genes of the bacterial, protist and fungal microbiomes of 131 samples collected in 14 ephemeral small inland lakes located in the endorheic area of the Monegros Desert (NE Spain). The sampling covered different temporal flooding/desiccation cycles that created natural salinity gradients between 0.1% (w/v) and salt saturation. We aimed to test the hypothesis of a lack of competitive advantage for microorganisms using the “salt-in” strategy in highly fluctuating hypersaline environments where temperature and salinity transitions widely vary within short time periods, as in ephemeral inland lakes. Overall, 5,653 bacterial zOTUs and 2,658 eukaryal zOTUs were detected heterogeneously distributed with significant variations on taxonomy and general energy-yielding metabolisms and trophic strategies along the gradient. We observed a more diverse bacterial assembly than initially expected at extreme salinities and lack of dominance of a few “salt-in” organisms. Microbial thresholds were unveiled for these highly fluctuating hypersaline environments with high selective pressures. We conclude that the extremely-highly dynamism observed in the ephemeral lakes of Monegros may have given a competitive advantage for more versatile (“salt-out”) organisms as compared to those better adapted to stable high salinities usually more common in solar salterns. Ephemeral inland saline lakes offered a well-suited natural framework for highly detailed evolutionary and ecological studies.

4.1 Introduction

Life in saline aquatic environments has adapted to cope with high osmotic stress, and salinity has been repeatedly described as one of the main ecological factors affecting microbial community composition at the global scale (Auguet et al., 2010; Lozupone and Knight, 2007). The progressive decrease of biodiversity as salinity increases is also usually accompanied by a substantial loss in metabolic functions (Oren, 1999), which shapes important ecological changes along the salt gradient (Gasol et al., 2004). Osmotic regulation becomes crucial to halophile and halotolerant microorganisms to avoid water deficit and plasmolysis, and a wide array of strategies are adopted to cope with changes in salt concentrations (Booth and Louis, 1999; Wood, 2011). At salts concentrations $> 0.3\%$ freshwaters turn saline (Rodríguez-Valera, 1993), and marine-like salinities can be usually found ranging between 3% and 5% salts w/v in inland waters (Imhoff, 2001; Por, 1972). At higher salts concentrations, water becomes hypersaline (Edgerton and Brimblecombe, 1981), and an additional threshold is set at 15% based on the biological differentiation between moderate and more extreme halophiles (Imhoff, 2001; Kushner, 1978, 1985).

One of the strategies used by cells to reach osmotic equilibrium between inside and outside environments is the accumulation of inorganic salts in the cytoplasm, actively “pumped in” from the saline surroundings (Galinski and Trüper, 1994). This salt-in strategy is adopted by several of the organisms living at the most extreme hypersaline conditions (e.g., *Halobacteria*, *Halanaerobiales* and *Salinibacter*), and requires of permanent adaptation in the intracellular machinery, which, in turn, limits further adaptation back to low salinities (Gunde-Cimerman et al., 2018). A second strategy is the accumulation of compatible solutes, small organic molecules with neutral charge and low toxicity at high concentrations, usually combined with active “pumping out” of inorganic salts (Lang, 2007). In general, salt-out microorganisms are better adapted to wider ranges of salinities but they have limitations to colonize extreme salinities due the high energetic cost to produce compatible solutes (Oren, 1999). Indeed, the bioenergetic requirement for maintaining these necessary osmoadaptation mechanisms is

one of the main limitations microbes have for colonizing saline environments, and therefore some energetically limited metabolic pathways cannot be present beyond a certain salinity threshold (Oren, 1999). This strategy is followed by most halophilic bacteria and most of the currently known halophilic microeukaryotes (Gunde-Cimerman et al., 2018; Hagemann, 2016; Trüper and Galinski, 1990).

Osmoadaptation mechanisms are better known in halophilic prokaryotes than in microeukaryotes. This lack of knowledge remains in the higher methodological complexity and the more recent awareness on the ecological relevance of microeukaryotes in hypersaline environments (Pedrós-Alió et al., 2000). Most of the knowledge regarding osmoadaptation in microeukaryotes had been gained from the study of the green algae *Dunaliella* and halophilic fungi, e.g. *Hortaea werneckii* (Plemenitas et al., 2008), *Wallemia ichthyophaga* (Zajc et al., 2014) and *Eurotium rubrum* (Kis-Papo et al., 2014). However, more recent works have added additional information from phagotrophic protists based on both transcriptomic analysis (*Halocafeteria seosinensis*, Harding et al. 2017) and experimental studies (*Schmidingerothrix salinarum*, Weinisch et al. 2018). Although the accumulation of compatible solutes has been described in all the cases as the main osmoadaptation strategy for microeukaryotes, the wide genetic repertory found related to specific adaptations suggests a larger variety of adaptive strategies (Kis-Papo et al., 2014). Thus, some of the main differences observed at the genome level include acidic residues in predicted proteins and genes linked to membrane cation transporters and synthesis/transport of compatible solutes. For instance, an elevated presence of genes linked to cation transporters have been associated to low intracellular concentrations of K^+ and Na^+ in *H. werneckii*. Conversely, a limited number of those genes in *W. ichthyophaga* suggests a higher resistance to high intracellular Na^+ concentrations (Weinisch et al., 2018). The expected higher adaptability of “salt-out” organisms may potentially provide competitive advantage in highly dynamic environments where frequent water level fluctuations and high evapotranspiration rates led to abrupt salinity changes, with potential dominance of organisms using the salt-out strategy even at the most extreme salinities (Filker et al., 2017; Gunde-Cimerman et al., 2018).

Ephemeral inland shallow lakes in arid and saline areas are very convenient natural model systems to test for selective microbial thresholds in highly fluctuating natural environments with high selective pressures. These ephemeral water masses are very dynamic and extremely responsive to changes in climate, with fluctuations in water volume, temperature, and salinity (Waiser and Robarts, 2009) larger than those reported in more stable hypersaline environments (e.g. solar salterns and large hypersaline lakes). The endorheic saline basin of Monegros is rich in ephemeral shallow lakes with severe spatio-temporal changes in salinity, easily ranging from freshwater to salt saturation and vice versa (Casamayor et al., 2013). We evaluated the composition and ecological changes of the planktonic microbiota (bacteria, protists and fungi) along the salt gradient of these highly dynamic environments with frequent and bidirectional salt fluctuations, and the results were contextualized according to the previous knowledge on the ecology of (hyper)saline environments.

4.2 Materials and methods

4.2.1 Study site and sampling

The Monegros Desert area (Central Ebro Basin, NE Spain, 41°42' N, 0°20' W) is one of the most arid regions in Europe with a large set of inland ephemeral saline lakes, and constitutes a unique landscape of great geological, edaphic, mineralogical and hydrological interest (Pueyo Mur and Inglés Urpinell, 1987; Samper-Calvete and García-Vera, 1998). The area is an endorheic platform hosting hypersaline brines of Cl-SO₄-Na-(Mg) type, where a rich microbial community develops (Casamayor et al., 2013; Menéndez-Serra et al., 2020, 2019). In this endorheic area, the water balance between precipitation (mean annual rainfall = 337 mm/year) and evaporation and evapotranspiration rates (mean = 1408 mm/year) is negative for most of the months and for the whole year, producing one of the highest mean annual water deficits in Europe (Herrero and Snyder, 1997). Field sampling was carried out in 14 ponds in different time periods and covered large salinity fluctuations (Supplementary Table B.1). Conductivity, temperature, pH, dissolved oxygen and redox potential were measured in situ using an HQ40d

multiparameter meter (Hach, Loveland, CO, USA). For chlorophyll *a* (Chl *a*) analysis, 1 L water was filtered through 47-mm-diameter Whatman GF/F filters (0.7 μm nominal particle retention), stored in the dark, and kept frozen. Chl *a* concentration was determined in acetone extracts by spectrophotometry (UV-2401PC, ultraviolet-visible Spectrometer; Shimadzu) as previously reported (Triadó-Margarit et al., 2019). Salinity was measured in situ by a hand salinity refractometer (Atago S-28E, Japan). Five salinity categories were established for comparative purposes as follows: Freshwater to slightly-brackish (0.1% - 1.4%, w/v), Brackish (1.5% - 2.9%), Marine (3% - 4.9%), Hypersaline (5% - 14.9%), and Extreme hypersaline (15% - 40%).

4.2.2 DNA Extraction and sequencing

For DNA analyses, water samples were prefiltered in situ through a 50 micron-pore-size net to retain large zooplankton and algae, and then between 100 and 500 mL (depending on the initial cells concentration) were successively filtered on 5-micron and 0.2-micron pore-size polycarbonate membranes (47 mm diameter). The analyzed eukaryal size range (5-50 μm) mainly covered the nanoplanktonic eukaryotes (2-20 μm) and a fraction of microplanktonic eukaryotes (20-200 μm). The 5 μm (for 18S rRNA gene analyses) and 0.2 μm (for 16S rRNA gene analyses) membranes were separately stored in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose), enzymatically digested and phenol extracted (phenol-chloroform-isoamyl alcohol; 25:24:1, v/v/v) as previously reported (Ferrera et al., 2004). The purification and concentration steps were carried out with Amicon® Ultra 4 Centrifugal Filter Units – 100 000 NMWL (Millipore). DNA extracts were quantified using Quant-it assays in a Qubit™ fluorometer (Invitrogen). PCR amplification of the V4 region of the 16S rRNA gene (primers set 515F-806R, Caporaso et al. 2010) and further sequencing with the Illumina 2x250 MiSeq platform were carried out according to the genomic core facilities protocols and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>). The V9 region of the 18S rRNA gene was PCR amplified with the primers set 1391f-EukrBr (Amaral-Zettler et al., 2009) and sequenced with the Illumina 2x150 MiSeq platform (RTSF-MSU).

4.2.3 16S rRNA and 18S rRNA genes sequence analysis

Raw rRNA gene sequences were processed with UPARSE (Edgar, 2013). The total number of sequences prior to quality filtering was 14,445,263 (mean of 79,060 per sample) for the 16S (70.6% had > Q30) and 14,374,602 (mean of 83,489 per sample) for the 18S (87.1% had > Q30) rRNA genes dataset. After merging of read pairs, each dataset was filtered by setting a minimum sequence length (250 bp for the 16S gene and 150 bp for the 18S gene) and a maximum number of expected errors of 0.25 (Edgar and Flyvbjerg, 2015). After chimera filtering using default parameters, the UNOISE algorithm (Edgar, 2016) clustered filtered sequences into Operational Taxonomic Units at 100% identity, i.e., zero-radius OTUs (zOTUs). The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012), using the SILVA 132 reference database (Quast et al., 2013). Specific BLAST searches were carried out to obtain a more detailed taxonomic classification of relevant zOTUs when needed. zOTUs were filtered by their alignment quality score (>85%) and 16S rRNA gene sequences classified as archaea, mitochondria, and chloroplasts were excluded. Archaea showed high richness and very low abundances below 20% salinities, and were treated separately in a different study covering temporal colonization-extinction turnover rates along the gradient (Menéndez-Serra et al., 2020). Only samples with > 6,800 reads for bacteria and > 8,100 reads for eukaryote were retained. The resulting zOTU tables consisted of 5,653 Bacteria (131 samples) and 2,658 Eukarya (129 samples). Data was normalized by rarefying the reads of all samples to the minimum number of reads in each dataset. For community related analysis, rarefaction at the same depth was repeated 100 times in order to avoid the loss of less abundant zOTUs, and the resulting rarefactions were unified on an average rarefied zOTU table. Phylogenetic diversity was calculated in QIIME (Caporaso et al., 2010) using *alpha_diversity.py* after building a maximum-likelihood phylogenetic tree using *make_phylogeny.py* (Price et al., 2010). General metabolic and trophic traits for bacteria and eukarya were obtained from the closest-match in SILVA database and after literature surveys (Llorens-Marès et al., 2015; Louca et al., 2016; Oren, 2011; Ortiz-Álvarez et al., 2018b). Energetic metabolism was focused on (i) photoautotrophy, (ii)

sulfur chemolithotrophs, (iii) sulfur chemoorganotrophs, (iv) methanotrophs, and (v) ammonia-oxidizing chemolithotrophs. For the trophic traits of protists and fungi, we established five groups as follows: (i) osmotrophs (which includes photosynthetic organisms), (ii) phagotrophs, (iii) mixotrophs, (iv) saprotrophic, and (v) parasitic. The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605.

4.2.4 Statistical analysis

Statistical analyses were run in R environment (<http://www.r-project.org/>). Species richness, and Shannon-Weaver diversity index were calculated using “vegan” package version 2.4-5 (Oksanen et al., 2016). Spearman’s Rank order correlation coefficient was carried out to test monotonic associations between numeric variables (environmental data, relative abundances, diversity indexes). Pearson’s correlation coefficient was used to test lineal correlation between numeric variables. Differences among group means were tested using Tukey’s range test, function *HSD.test* within “agricolae” package version 1.2-8, after ANOVA test. The Dufrene-Legendre Indicator Species Analysis, that identifies indicator taxa based on fidelity and relative abundance, was carried out at the genus level with the *indval* function on the “labdsv” package version 1.8, in order to show the genera significantly associated ($p\text{-value} < 0.05$) to each salinity category. The ten genera with the highest abundance in each salinity category were further analyzed. The *titan2* function from “TITAN2” package version 2.1 was used to identify significant variations on the taxonomic composition along the environmental gradient (Baker and King, 2013). This is done after fractioning the environmental gradient in two at each midpoint of the gradient, and calculating taxa significantly associated to each side of that midpoint, which is expressed as a Z score. Then, Z scores are based on the Indicator Species Analysis at each midpoint of the gradient. High values of negative Z scores (Z^-) indicated high association of taxa to the left side of the midpoint, the one with lower salinity, so represent points of the gradient where the environmental variable starts affecting negatively some taxa. Conversely, positive Z scores (Z^+) indicated the opposite pattern, where taxa are significantly associated to

the most saline side of the midpoint.

4.3 Results

4.3.1 Alpha diversity descriptors along the salinity gradient

Bacteria and microbial eukarya showed differences in alpha diversity community descriptors (Richness, Shannon, and Faith's Phylogenetic Diversity) along the gradient (Figure 4.1, and see additional information in Supplementary Table B.2). Bacterial richness (R) was higher (average 372 ± 116 zOTUs, range 127-793 zOTUs) than microbial eukaryotes R (average 124 ± 69 zOTUs, range 14-351 zOTUs). Shannon-Weaver diversity index (H) also reached higher values for bacteria (from 1.9 to 5.1, mean 3.47) than for eukarya (from 0.09 to 4.15, mean 2.21). The plots of richness and alpha diversity indicators

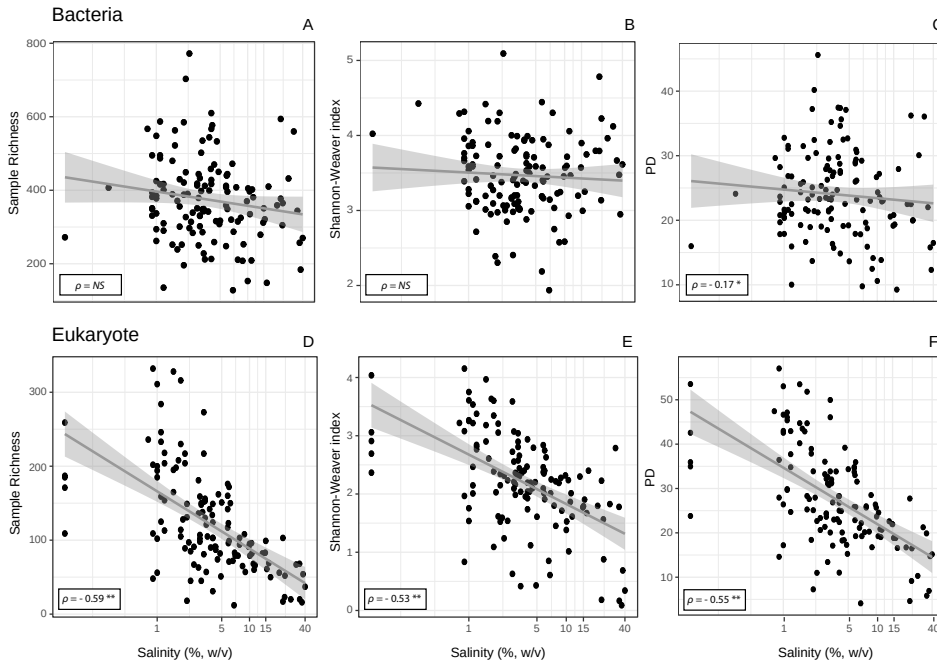


Figure 4.1: Variations on zOTU richness, Shannon-Weaver diversity index and Faith's phylogenetic diversity along the salinity gradient, both for bacteria (a, b, c) and eukaryote (d, e, f). Pearson's correlation is given when significant p-value: * p-value < 0.05; ** p-value < 0.01; *NS* = non-significant correlation.

vs. salinity showed a significant and negative trend for eukaryotes (Figure 4.1 D and E, respectively), but was not significant for bacteria (Figure 4.1 A and B). The negative trend was observed for both photosynthetic and heterotrophic protists, and also for fungi (Supplementary Table B.2). At any rate, the phylogenetic diversity of both bacteria and eukarya showed a significant reduction as salinity increased that became stronger for microbial eukaryotes (Figure 4.1 C and F).

4.3.2 Taxa replacement along the salinity gradient

We identified significant variations on the taxonomic composition along the salinity gradient based on the Z scores (Figure 4.2), with significant ecological thresholds detected. Similar values for Z+ and Z- scores were observed in bacteria whereas, interestingly, eukaryal Z+ scores were significantly lower than Z-. The Z- for both bacterial and eukarya followed a similar pattern, with its maximum and most of the contribution to the cumulative frequency observed between 2-3% salinity. The higher Z+ bacterial scores at salinities 5 to 20%, indicated a partial replacement of the bacterial community by better adapted taxa that was not observed for eukarya.

4.3.3 Indicator bacterial taxa (IndVal) along the salt gradient

As a whole, bacteria averaged 98% (relative abundances range 54-100%) of the total prokaryotes present in the PCR mixture (Supplementary Table B.1). The most abundant phyla were *Gammaproteobacteria* (25% - 44%, mean relative abundances), *Alphaproteobacteria* (10% - 35%), *Bacteroidetes* (18% - 29%), and *Actinobacteria* (c. 5%) (Supplementary Figure B.1, and see a zoom for salinities >35% in Supplementary Figure B.2). *Halanaerobiaeota* (a recently proposed phylum in the SILVA database which includes the single order *Halanaerobiales*) were only noticeable at the most extreme end. Conversely, *Firmicutes* reached higher abundances at low salinities. The phylum *Spirochaetes* showed a positive correlation between relative abundance and salinity (Spearman's rho 0.60, p-value < 0.001), with a

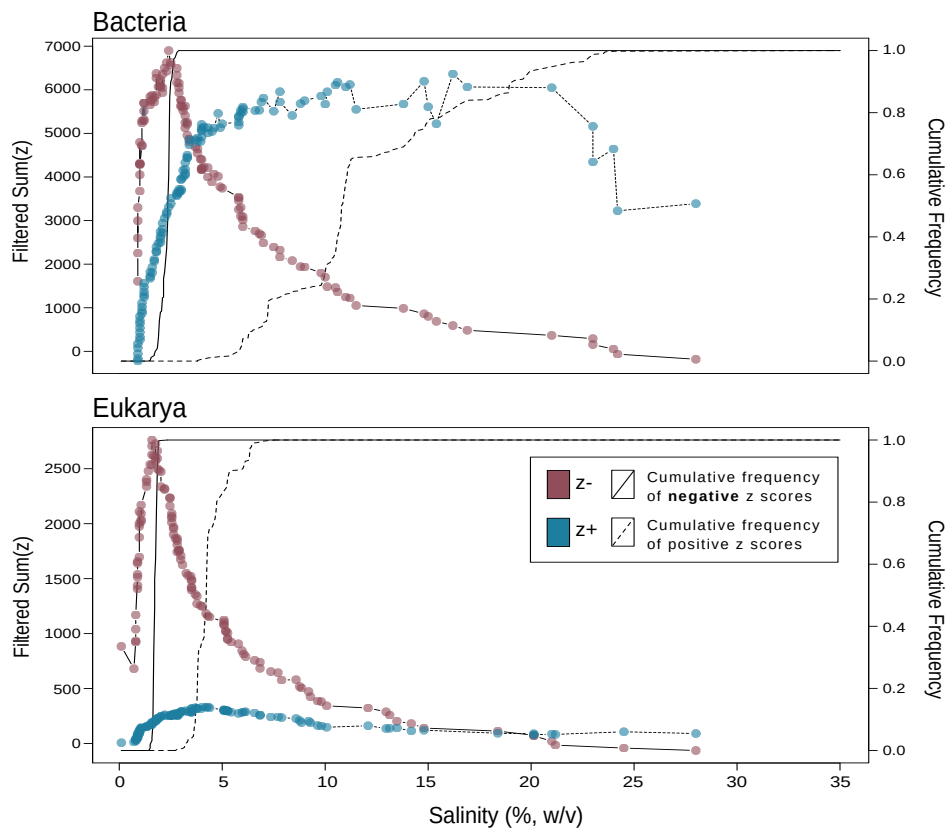


Figure 4.2: Representation of the sums of Z scores (positive and negative) calculated for each midpoint of the salinity gradient using titan2 function from “TITAN2” R package, both for bacterial and eukaryal communities. Sums of Z scores are based on the Indval analysis performed after segregating all samples along the salinity gradient as less/more saline than the tested point. Continuous and dashed lines represents the accumulated frequency of negative and positive Z scores respectively, and its pronounced increase in certain points represents potential thresholds on the community composition variation. As the Indval analysis requires of a minimum number of samples at each side of the gradient midpoint, this could not be performed at the salt gradient tails.

highly diverse composition, and reached the 3% of bacterial reads at 20-25% salinity in the ephemeral lake Salineta (Supplementary Figure B.3).

The Indval analysis unveiled the genera with the highest fidelity for each salinity category (Figure 4.3). The genera *Polaromonas*, *Polynucleobacter*, *Limnohabitans*, *Hydrogenophaga*, *Bacillus*, *Pedobacter* and *Flavobacterium*

reached significant abundances at low salinity, whereas *Salinivibrio*, unclassified *Nitrococcaceae* (94.4% of 16S rRNA gene sequence identity with *Nitrococcaceae* members included in SILVA database), *Halothiobacillus*, and genera belonging to the order *Halanaerobiales* became more important in terms of abundance at high salinities. Interestingly, the high abundance of the unclassified *Nitrococcaceae* at salinities >15% was mostly related to a single population (zOTU 7, NCBI acc. number MT129572, Supplementary Table B.3) 100% identical to the recently described *Spiribacter salinus* (*Ectothiorhodospiraceae*) (León et al., 2014). Conversely, *Salinibacter sp.* was always present at relative abundances <0.3% even at salinities >30%, and maximal abundances reached 3-4% only in two samples for the whole dataset (Salineta, samples MON 88 and 84). Finally, *Psychroflexus* and unclassified *Rhodobacteraceae*, despite being significantly associated to a given salinity category, were substantially abundant along the whole salinity gradient.

4.3.4 Indicator eukaryal taxa (IndVal) along the salt gradient

Despite the significant decrease of eukaryotic diversity observed along the salinity gradient, the whole eukaryal microbiome showed an outstanding diversity in the ephemeral saline lakes of Monegros, with up to 2658 zOTUs belonging to Amoebozoa, Archaeplastida, Centrohelida, Cryptophyceae, Excavata, Haptophyta, Opisthokonta and SAR (Supplementary Figure B.4). Unclassified *Ciliophora* and *Cryptomonas* showed the highest affinity for low salinities, whereas *Dunaliella* (Chlorophyta) prevailed at extreme salinities (Figure 4.3). Most of the reads classified by SINA aligner as unclassified Chlorophyta. belonged to zOTU 3 (100% identity to *Dunaliella salina*, NCBI acc. number MG022673), which reached 80% of the total eukaryote microbiota for most of the extreme-hypersaline samples. It was also remarkable the high abundance of diatoms of the genus *Navicula* (Ochrophyta, Bacillariophyceae) along most of the salinity gradient, reaching >25% of the relative abundance of eukaryotes in the brackish range. The dinoflagellate *Suessiacea* (unclassified) represented by a single zOTU (zOTU 2), reached high abundances and fidelity at salinities <15% and showed

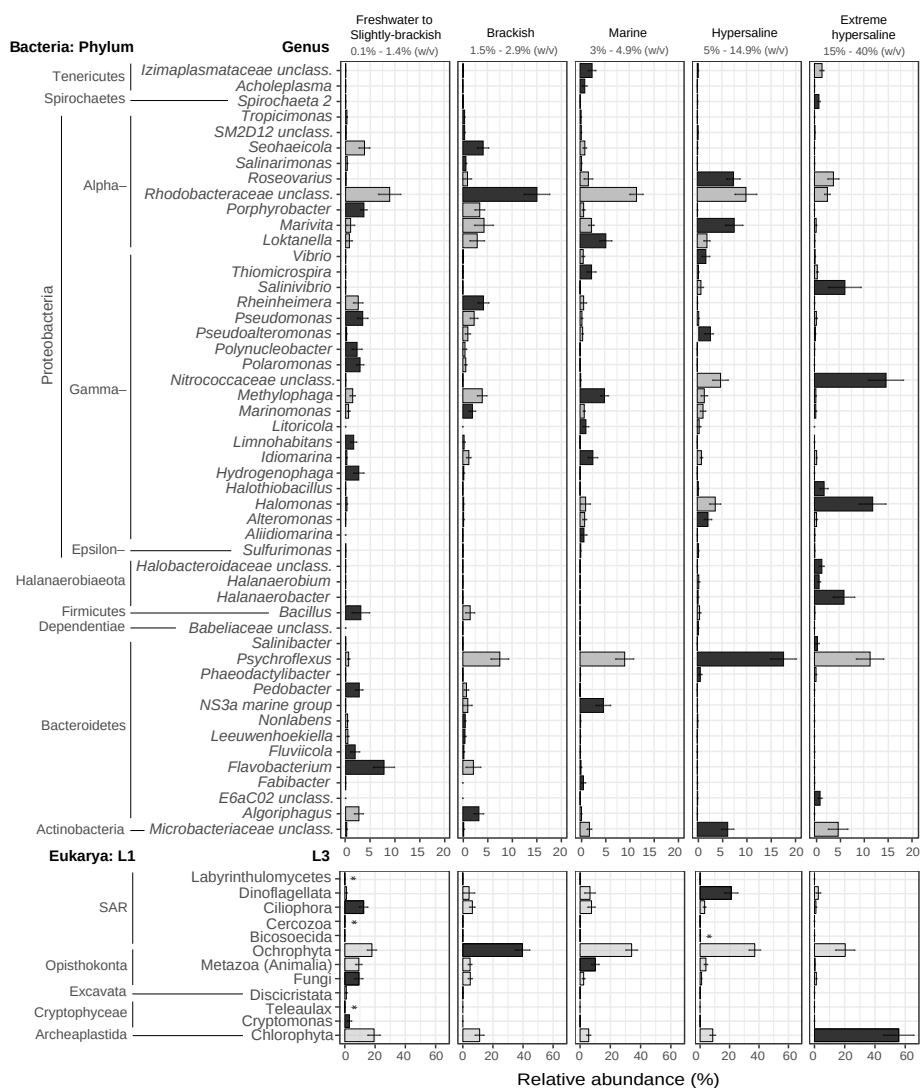


Figure 4.3: Relative abundances (%) of the genera obtained from the Indval analysis along the different salinity ranges. Highlighted abundances represent the indicator species for each salinity range.

100% identity to the halotolerant dinoflagellate *Biecheleria* sp. (NCBI acc. number KF463288 isolated from the saline lake Tirez, Spain).

Finally, fungi prevailed at the lower salinity end (Figure 4.3 and Supplementary Figure B.5) and were less abundant as salt content increased.

The most abundant taxa were Chytridiomycota, and two unknown fungi that deserve further research (zOTU 57, up to 23% of the eukaryotic reads in sample MON 2, salinity 1.2%, NCBI acc. number MW040517; and zOTU 102, up to 15.5% of the eukaryotic reads in sample MON 6, salinity 1.1%, NCBI acc. number MW040518). These unknown fungi showed <91% identity to any previously known sequences and were putatively allocated within the Basidiomycota in SINA.

4.3.5 Metabolic and trophic thresholds along the salt gradient in ephemeral lakes

Changes in the general energy metabolism processes and trophic strategies were closely linked to shifts in microbial composition. Bacterial photoautotrophy (cyanobacteria and anoxygenic photoautotrophs) was a minor component and with no significant differences along the salt gradient categories (Figure 4.4). Conversely, sulfur bacteria both chemolithotrophs

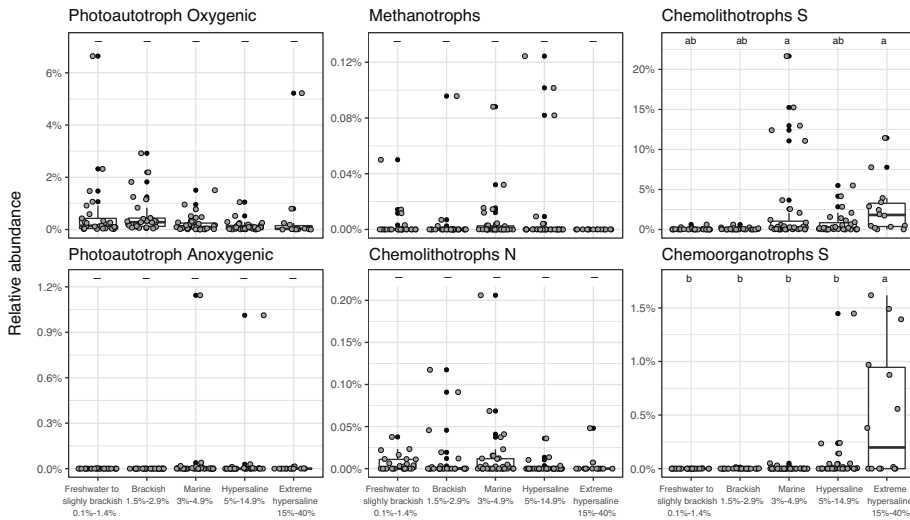


Figure 4.4: Relative abundance of the bacterial taxa involved on diverse metabolic processes. Chemoorganotrophs S: as Sulfate reducing bacteria; Chemolithotrophs S: as Sulfur oxidizing bacteria. Grey dots represent group abundance on each sample. Black dots represent boxplots outliers.

(Sulfur Oxidizing Bacteria, SOB) and chemoorganotrophs (Sulfate Reducing Bacteria, SRB) increased along the gradient. Methanotrophs showed low abundances and were undetectable at salinities $>15\%$. Ammonia oxidizers were also poorly represented in the Monegros microbiome, with AOB undetected and AOA (*Thaumarchaeota*) mostly detected at salinities up to 10% (Menéndez-Serra et al., 2020).

Osmotrophy by photosynthetic microeukaryotes was the dominant eukaryal trophic strategy along the whole salt gradient, reaching almost 100% dominance at salinities $>15\%$ (Figure 4.5) in agreement with the highest relative abundances of *Dunaliella* and *Navicula*. Conversely, phagotrophic eukaryotes (*Ciliophora* and *Cercozoa*) averaged c. 25% of the abundance at salinities $<15\%$ and became minor component at salinities $>20\%$, whereas the abundance of mixotrophic (mostly *Cryptomonas* and *Dinoflagellata*) reached its maximum at salinities $<1.5\%$ and unseen at salinities $>3\%$. Finally, saprotrophic and parasitic microeukaryotes, related to the presence of Chytridiomycota, were observed mostly at $<5\%$ salinity.

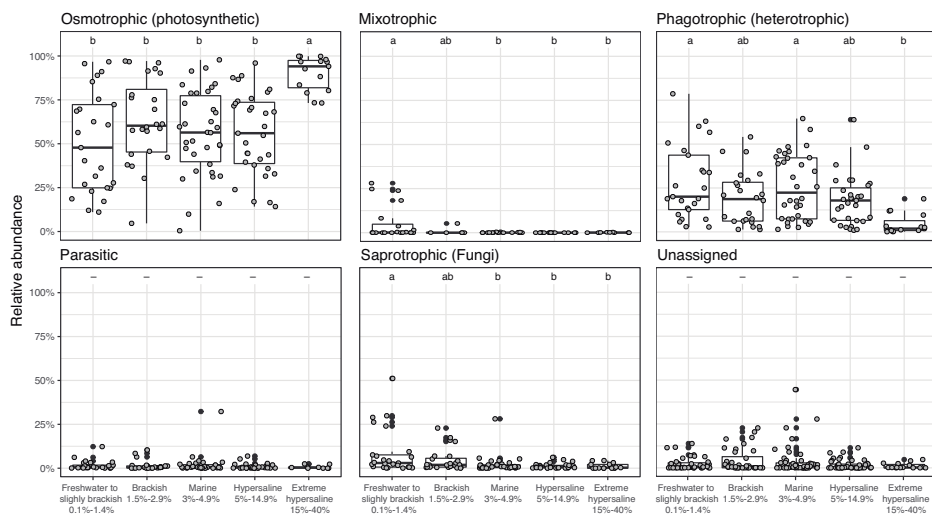


Figure 4.5: Relative abundance of the eukaryal taxa based on its trophic strategy. Grey dots represent group abundance on each sample. Black dots represent boxplots outliers.

4.4 Discussion

The dynamics and succession of planktonic microbial populations in ephemeral saline lakes have been poorly studied (Navarro et al., 2009). In hypersaline inland environments of arid regions, it is challenging to collect enough relevant information over temporal and spatial scales to properly study natural microbial processes, and successional studies are scarce (Podell et al., 2014). The Monegros desert contains one of the largest sets of inland saline lakes in Europe offering a wide repertoire of accessible heterogeneous and small ephemeral lakes within a short distance, easily covering different interannual dry and wet periods and ecological conditions (Casamayor et al., 2013; Menéndez-Serra et al., 2020). In the present study, we observed consistent changes in ecological and metabolic thresholds along the saline gradient that selectively affected different microbial populations according to their expected physiological and metabolic traits. The best known ecological thresholds with clearly segregated species distributions are usually found at the interfaces between land and water, oxic and anoxic environments, and marine waters and freshwaters. However, ecological thresholds may also exist when environmental changes are progressive (Catalan et al., 2009). As a whole, ecological thresholds are very helpful to understand ecosystems structures, to guide conservation policies, and to establish quantitative predictions for the ecological impact of environmental changes (Catalan et al., 2009; Triadó-Margarit et al., 2019). In the intermittent saline lakes of Monegros, the strong environmental selection pressures (low rain precipitation, high solar irradiation, strong and persistent winds, high evaporation rates, and salts concentration and accumulation) provide an easy to access natural model system with progressive changes in salinity where ecological and metabolic thresholds appeared finely allocated.

Microorganisms inhabiting saline and hypersaline environments have been described in detail in the literature, first using culture dependent approaches (Rodriguez-Valera et al., 1985; Ventosa et al., 1984), and more extensively after the use of molecular biology techniques. Generally, halophilic bacteria and archaea dominate in saline environments, with lower contribution of microeukaryotes to the total diversity. Most of the halophile and halotolerant bacteria are considered to be moderate rather than extreme

halophiles (Oren, 2002). The Bacteria domain contains many halophiles from diverse phylogenetic groups such as *Proteobacteria*, *Cyanobacteria*, *Spirochaetes*, *Actinomycetes* and *Firmicutes*. The genera *Bacillus* (Kim et al., 2012), *Salinibacter* (Antón et al., 2002) and members of the order *Halanaerobiales* (Antón et al., 2002) are some of the most commonly found taxa expected in these environments. Archaea become more abundant as salinity increases (Andrei et al., 2012), mostly with *Nanohaloarchaeota* and *Halobacteria* populations (Andrei et al., 2012; Menéndez-Serra et al., 2020) being *Haloquadratum* and *Halorubrum* predominant at the highest salinities (Dillon et al., 2013; Oh et al., 2010; Ventosa et al., 2014a). Microeukaryal diversity tends to be lower at high salinities (Filker et al., 2017) and are mainly represented by Chlorophyta, Dinophyceae and Bacillariophyta, with the photosynthetic algae *Dunaliella* as one of the most common taxa at extreme salinities (Heidelberg et al., 2013; Oren, 2014; Triadó-Margarit and Casamayor, 2013). Together with other reported halophilic algae (Javor, 1989), both halophilic and halotolerant fungi (e.g. *Hortaea werneckii* and *Wallemia ichthyophaga*) and halophilic heterotrophic protists (Gunde-Cimerman et al., 2018; Heidelberg et al., 2013) are also usually reported in saline environments.

Therefore, in principle, a similar assembly should be present in the saline lakes of Monegros. However, the taxonomic composition found, specially at the most saline end, significantly differed from the expectations following previous reports. Although specialist groups as *Salinibacter* and *Halanaerobiales* were detected at extreme salinities, they only represented a very low fraction of the total bacterial assembly. Conversely, a very rich assembly of *Gammaproteobacteria*, *Bacteroidetes* and *Alphaproteobacteria* using “salt-out” strategies dominated the bacterial community at extreme salinities. Most of the populations found matched *Halomonas* (El Hidri et al., 2013; Vreeland et al., 1980), *Salinivibrio* (Ventosa et al., 2014a), *Spiribacter salinus* (León et al., 2014) and members of the order *Halanaerobiales* (Oren, 2002), among others, with high relative abundances at salinities close to NaCl saturation (Supplementary Figure B.2). This fact led to a higher bacterial diversity at extreme salinities than expected from studies carried out in more stable environments, where usually “salt-in” archaea, *Salinibacter* and

Halanaerobiales dominate (Fernández et al., 2014; Ghai et al., 2011). As previously reported, this strategy is more effective at the most extreme salinities and reduces adaptation to lower salinities (Galinski and Trüper, 1994; Oren, 2008) and, therefore, should be less competitive for the most dynamic natural environments (Gunde-Cimerman et al., 2018; Sleator and Hill, 2002). Conversely, “salt-out” mechanisms provide higher plasticity at the expense of poorer adaptation at the most saline end (Kempf and Bremer, 1998). Microeukaryotic assembly was mostly dominated by *Dunaliella* and *Navicula*, both accumulating compatible solutes as the main osmoadaptation strategy (Oren, 2014; Scholz and Liebezeit, 2012), with progressively lower contribution of halophilic fungi and heterotrophic protists as salinity increased. Overall, the extremely high dynamism observed in the ephemeral lakes of Monegros selectively enriched more versatile ("salt-out") organisms rather than those better adapted to stable high salinities, usually more common in solar salterns.

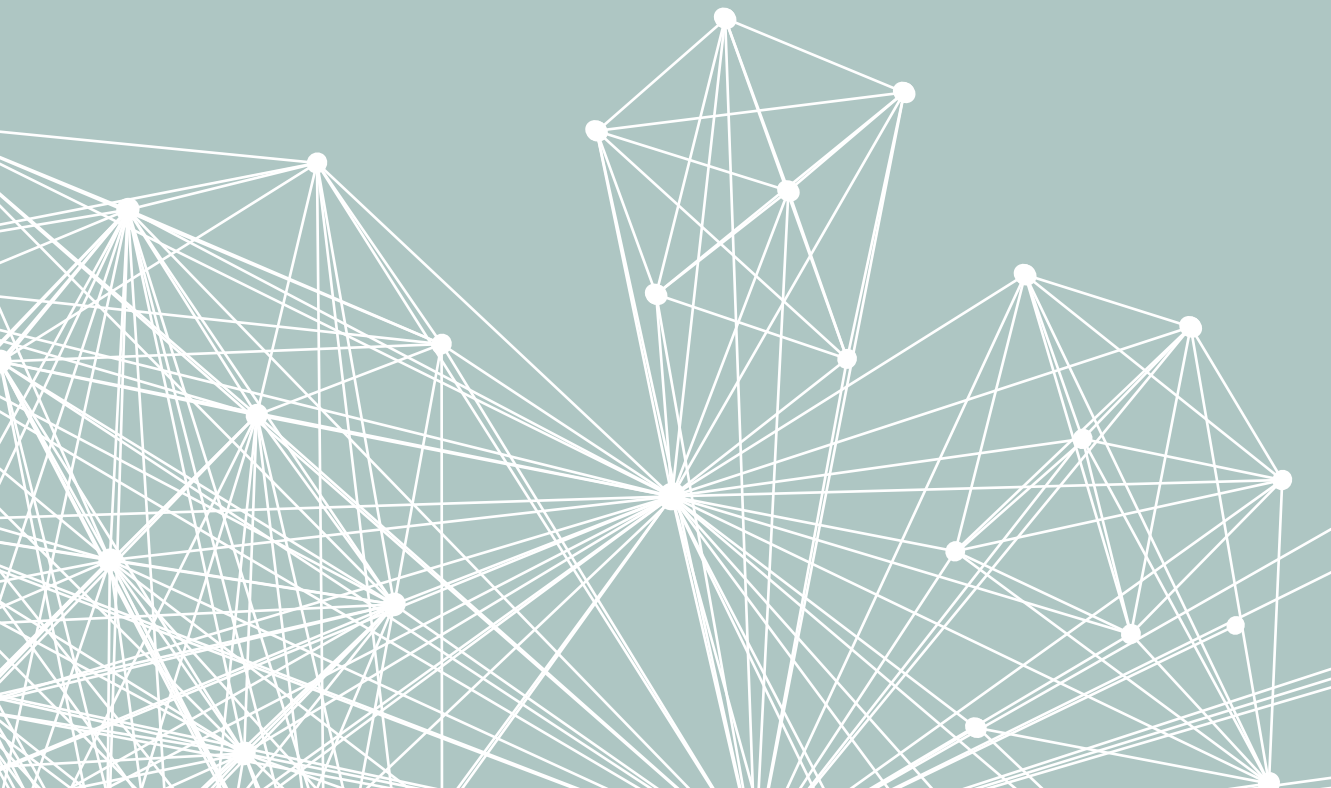
Changes in metabolic and physiological performance were also consistent along the salinity gradient. Autotrophic ammonia oxidation, autotrophic nitrite oxidation, and acetate-based methanogenesis, among others, are usually restricted to low salinity environments (Fernández-Guerra and Casamayor, 2012; Oren, 1999, 2011). The low abundance of methanotrophs observed in the ephemeral lakes of Monegros may be explained by constraints for aerobic methane oxidation in high-salt environments (Oren, 2011) but also for limitations in methanogenesis. Methanogenic activity based on H_2+CO_2 (group 1) and acetate (group 3) have never been detected above c. 12.5% salinity (Oren, 2011; Pattnaik et al., 2000). Methanogenesis based on methylated amines (group 2) may potentially occur up to 25% of salts (w/v), under anaerobic conditions (Andrei et al., 2012; Oren, 2011). However, Monegros is a gypsum-rich area with a rich presence of sulfur bacteria and a potentially active sulfur cycle in most of the sediments (Menéndez-Serra et al., 2019) that do not favor methanogenesis.

We also observed consistent changes of the trophic strategy for microbial eukaryotes in the ephemeral lakes. The dominance of osmotrophy (which includes photosynthetic organisms) was linked to the high abundance of *Dunaliella* and *Navicula* covering all the gradient. The low abundance

of *Cyanobacteria* at high salinities agrees with previous reports (Oren, 2015) being *Dunaliella* the main primary producer at the most extreme salinities. The progressive decrease of phagotrophic organisms agrees with previous studies in coastal solar salterns that showed lack of bacterivory above 25% salinity (Pedrós-Alió et al., 2000). Groups potentially including halophile and halotolerant melanized fungi ('black yeasts') usually found in hypersaline environments (Gunde-Cimerman et al., 2018), were detected at low abundances along the gradient in Monegros. Saprotrophic fungi were mostly found at salinities <5%. However, more detailed studies and refined taxonomic assignments are needed for the aquatic fungi found in Monegros. Finally, parasitic microeukaryotes were present all over the gradient suggesting potential key missing pathways in ephemeral lakes dynamics. Taken all together, these unseen groups offered new perspectives for improving the understanding of planktonic food webs (Ortiz-Álvarez et al., 2018b) and should be more carefully studied in the future.

Overall, we have reported a very dynamic and rich microbial assemblages in the Monegros wetlands system, even under the most extreme conditions. The particular characteristics and high environmental variability of this area, allow us to observe natural transitions in detail both at the taxonomic and at the functional level. Most of the high saline waters traditionally studied are quite stable environments with rather constant salinities (Heidelberg et al., 2013; Jacob et al., 2017; Podell et al., 2014) or where salt production practices guides salinity increases unidirectionally (Ventosa et al., 2014a). Ephemeral saline lakes with high turnover and frequent salinity transitions follow closely environmental perturbations and flooding and desiccation dynamics, offering a well-suited natural framework for the study of evolutionary and ecological microbial responses (Triadó-Margarit et al., 2019) under highly dynamic fluctuations and strong selective pressures.

CHAPTER FOUR



Xavier Triadó-Margarit, José A. Capitán, **Mateu Menéndez-Serra**, Rudiger Ortíz-Álvarez, Vicente J. Ontiveros, Emilio O Casamayor and David Alonso (2020). A Randomized Trait Community Clustering approach to unveil consistent environmental threshold in community assembly. *The ISME Journal*, 13, 2681-2689, DOI:10.1038/s41396-019-0454-4.

This chapter can be found in its publication format in the *Published Chapter* section of the present thesis. Likewise, the bibliography of this chapter can be found on the unified *Bibliography* section.

A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly)

Abstract: Similarities and differences of phenotypes within local co-occurring species hold the key to inferring the contribution of stochastic or deterministic processes in community assembly. Developing both phylogenetic-based and trait-based quantitative methods to unravel these processes is a major aim in community ecology. We developed a trait-based approach that: (i) assesses if a community trait clustering pattern is related to increasing environmental constraints along a gradient; and (ii) determines quantitative thresholds for an environmental variable along a gradient to interpret changes in prevailing community assembly drivers. We used a regional set of natural shallow saline ponds covering a wide salinity gradient (0.1–40% w/v). We identify a consistent discrete salinity threshold (ca. 5%) for microbial community assembly drivers. Above 5% salinity a strong environmental filtering prevailed as an assembly force, whereas a combination of biotic and abiotic factors dominated at lower salinities. This method provides a conceptual approach to identify consistent environmental thresholds in community assembly and enables quantitative predictions for the ecological impact of environmental changes.

5.1 Introduction

The main goal of community ecology is to understand the underlying forces that determine species numbers and identities, and their relative abundances, across spatio-temporal scales. Many recent efforts are aimed at developing robust methods to estimate and understand the relative importance of different community assembly processes (van der Plas et al., 2015). Ecologists generally accept that community assembly must necessarily be driven by both dispersal-assembly and niche-assembly mechanisms (Gravel et al., 2006; Vellend, 2010, 2016). Dispersal-assembly theories emphasize (1) stochastic colonization (driven largely by species pools) and (2) local extinctions (driven by random events affecting small populations sizes) (Arthur, 1969; Hubbell, 2001), while niche-assembly highlights (1) the importance of species differences and (2) their interactions in predicting community composition (Law and Morton, 1996; Law et al., 2000; Tilman, 1982). Niche-assembly mechanisms can affect either extinction or colonization. From this perspective, local extinctions result in general from negative interactions. They are often caused by competitive exclusion of sub-dominant species disappearing from the system due to a worse overall performance (Chesson, 2000; Mayfield and Levine, 2010), but strong environmental filtering (cf. species sorting (Leibold and Chase, 2018) can also prevent colonization and establishment of poorly adapted species.

In contrast, positive interactions (facilitation) decrease the probability of local extinctions, enhancing persistence and biodiversity (Bastolla et al., 2009; Bertness and Callaway, 1994; Maestre et al., 2009). Assessing the link between trait variation, the environment, and biotic interactions (either positive or negative) is not trivial (Helmus et al., 2007; Kraft et al., 2015; Peres-Neto et al., 2017; van der Plas et al., 2015). Further, there is an increasing recognition of the need for methods to make robust predictions that help understand responses of species assemblages to environmental shifts (Peterson et al., 2011; Wisz et al., 2013). In studying community assembly of macroorganisms, spatio-temporal scales are commonly huge, making long-term ecological research challenging. Microbes, in contrast, thrive in highly diverse communities and are characterized by short time scales. This makes microbial communities ideal study systems, providing robust insights

into assembly processes (Barberán et al., 2014a; Dini-Andreote et al., 2015). Species and individuals within assemblages respond to the environment according to their functional characteristics. In particular, microbes with stress tolerance traits are selected in specific environmental conditions (Goberna et al., 2014). Genomic traits allow the linkage of bacterial functions to their environmental preferences at different scales, from local habitats to biogeographical regions (Barberán et al., 2014b). This approach could be especially helpful across habitats with strong environmental constraints, where a variety of functional adaptations can emerge along a gradient. Moreover, there is an increasing recognition that variations in the distribution of functional traits along environmental gradients can be quantified (Edwards et al., 2013; Peres-Neto et al., 2017; ter Braak et al., 2012). Whenever niche-assembly mechanisms are overridden by strong random processes, the local community composition is expected to be akin to a random sample from the regional pool (Etienne and Alonso, 2005). When this happens, local trait distributions should not differ significantly from the regional trait distribution. Conversely, differences between local and regional trait distributions are a fingerprint of the existence of specific drivers of community assembly. The classical framework by Webb (Webb et al., 2002) considers clustering in trait values as a signature of environmental filtering, while Mayfield and Levine (Mayfield and Levine, 2010) remarked that such clustering can result also from competitive exclusion.

Here, we use shifts in the distribution of functional traits along an environmental gradient to assess the ecological impact of a changing environment on community trait structure. Our methods are inspired by the data randomization/resampling null model approach (Gotelli and Graves, 1996). Our trait-based approach consistently finds thresholds in an environmental variable along a gradient with the aim of interpreting meaningful changes in the type of drivers responsible for community assembly. We call this index-based method RTCC—Randomized Trait Community Clustering.

5.2 Materials and methods

Details related to sampling sites, molecular methods, bioinformatics processing, and additional statistics are provided in SI Materials and Methods, while null models of community assembly are fully described in this section. See also supplementary text and Figure C.1 for a summarized work-flow of the methods carried out in this study. In addition, source code for the RTCC Method and sample data files are available in GitHub. Briefly, samples were collected from the Monegros Desert area (NE Spain, 41°42'N, 0°20'W) which harbors among the largest number of inland saline lakes in Europe (Casamayor et al., 2013). A total of 148 samples were taken from 14 different shallow lagoons (Figure C.4). We explored the spatio-temporal variation of a regional set of these local aquatic communities driven by a wide salinity gradient (0.1–40% of dissolved salts) (Figure C.5 A). Temporal data of wind speeds were obtained from the Meteorological state agency reporting data of an automatic station in the area (Bujaraloz) (Figure C.5 B). For DNA analyses, water samples were pre-filtered in situ through a 50 μm -pore-size net, to retain large zooplankton and algae, and 100–500 mL were subsequently filtered on 5 μm and then 0.2 μm pore-size polycarbonate membranes (47 mm diameter).

The membranes were enzymatically digested and phenol–chloroform extracted (Casamayor et al., 2013). To obtain community profiles, we performed a NGS sequencing step targeting the V4 region of the bacterial 16S rRNA gene by means of a high-speed multiplexed Illumina MiSeq platform. Raw sequences were processed using the UPARSE pipeline (Edgar, 2013). After the processing steps, we only kept those samples that presented library sizes larger than 10,000 reads/sample. The final pool consisted of 136 samples with 9,993 operational taxonomic units (OTUs). In order to associate OTUs to genome content, we used a 16S rRNA database (Wattam et al., 2014) linked to the IMG genomic database (Markowitz et al., 2014). OTUs were matched to 16S rRNA gene records available in the PATRIC genomic database (as of January 2016) (Wattam et al., 2014). The average percentage of abundance matched per sample was ca. 62%. We tested the representativeness of the subset matching genomes in our approach (see Figure C.1 and C.2), i.e., we explored if the ordination pattern of

the whole dataset was consistent with the ordination of the subset. In addition, a Mantel test (based on Spearman correlation) was also used to conclude on this decision-making step. See also the RTCC Method protocol in SI document. Functional predictions based on representative genomes allow for an inferring of genomic and metabolic potential of 16S rRNA sequences (Langille et al., 2013; Nemergut et al., 2016). For this we downloaded available genomic traits of potential interest from IMG (Markowitz et al., 2014): genome size, gene count, % GC, % coding base, % CDS, % RNA, rRNA count, % transporter proteins, % signal peptide, and % transmembrane proteins. We selected 10 traits that summarize genomic structural/functional variability, some of them known to respond to ecological adaptations.

5.2.1 Null models of community assembly

In order to quantify the importance of environmental filtering in the assembly of ecological communities, we developed a measure of the degree of clustering of a set of samples (or sites). The pool of species was formed by all the species observed in the complete set of samples taken from the lagoons. First we calculated a measure of average dissimilarity in the trait distribution of a sample as the pairwise mean trait difference between every species pair in the sample, the functional mean pairwise distance (MPD) (Webb et al., 2002)

$$\Delta x = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n |x_i - x_j| \quad (5.1)$$

where x_i is the trait value of species i in the sample (with n reported species). Next we tested if the observed measure Δx significantly differed from the expectation of the same measure across realizations of consistent null models with the same number of species as appearing in the empirical sample. We assumed that any species from the pool can take part in a null-model synthetic community. We build the null-model dissimilarity distribution over replicates to obtain a p-value, i.e., the quantile defined by the empirical difference within the distribution of simulated differences across realizations.

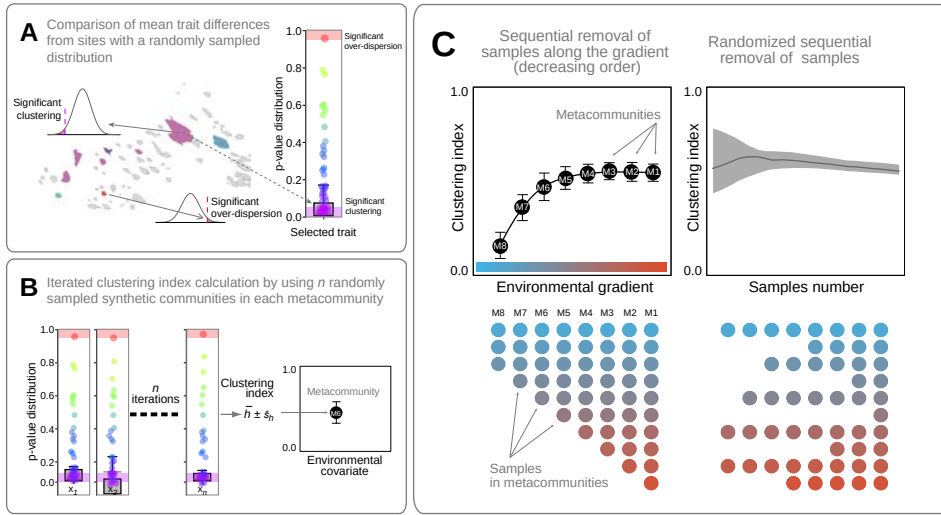


Figure 5.1: Methodological framework for the Randomized Trait Community Clustering (RTCC) method. a For each sample from different water bodies over a given region we test the hypothesis that observed average dissimilarity (see Eq. (1)) is compatible with a given null model. These non-parametric randomizations assign a p-value for each sample. For all samples in the dataset we represent the p-value distributions obtained by the null model for a given functional trait. Significantly low p-values are related to clustering ($p < 0.05$), and significantly high p-values to over-dispersion. b For the whole sample set, which we define here as a metacommunity, we calculate the fraction of samples that yield significant clustering, i.e., its clustering index (see Eq. (2)). We average several realizations of the null model in order to get average clustering index and error bars. c Samples are sequentially removed along decreasing values of the environmental variable to define nested metacommunities; for each of them we calculate the clustering index (left). The significance of this curve is assessed by building an ensemble of curves corresponding to removing samples in random orderings, which are then used to define the shaded area (right).

We tested whether the empirical sample dissimilarity value was significantly large or small, indicating over-dispersion or clustering, respectively. If the observed difference of our empirical sample was under the lower 5% of the simulated distribution trait differences, we considered that sample as showing significant clustering (Figure 5.1 a, left). We repeated this test for every sample in our data set, which yielded the p-value distributions for each target trait (Figure 5.1 a, right). Let N be the number of samples in the data set, and N_c the number of samples that showed significant clustering

when tested against the null model. We defined a clustering index,

$$h = \frac{N_c}{N} \quad (5.2)$$

as the proportion of samples showing significant clustering over all the samples considered. Note that the clustering index is a global measure for the whole set of samples. Error bars for the clustering index were calculated by averaging this quantity over 100 repetitions of the distribution of p-values according to the null model (Figure 5.1 b). The null models developed differ in the definition of the species pools potentially available for each of the samples, that is, in the way species were drawn from the species pool:

- i *Random assembly*: For all samples, all species are potentially available. They were chosen at random from the pool with equal probability.
- ii *Abundance-based assembly*: For all samples, all species are potentially available, but had a probability of being chosen proportional to their abundance in the entire species pool.
- iii *Environmental range-based assembly*: For a given sample, only those species whose ranges for the environmental factor encompass the value observed in that sample are eligible to be drawn and form the corresponding randomly generated sample. Potentially eligible species are drawn with uniform probability. Then, in this constrained way, we build a set of random synthetic samples, each characterized by a given value of the environmental factor and the same number of species as observed in each corresponding observed sample. For each species, the range of the environmental factor is defined by the minimum and maximum values of the variable across the samples in which the species is observed.
- iv *Environmental range and abundance-based assembly*: As (iii), but, additionally, species whose ranges for the environmental factor contain the value of the empirical sample are randomly drawn proportionally to their abundances in the species pool.

For the purpose of discerning if significant clustering (or the lack of it) was influenced by the environmental gradient, we iteratively excluded the sample from the set with the highest value of the environmental variable of interest, also excluding from the pool, the corresponding species that appeared only in that sample, and recalculated the clustering index with the remaining samples (Figure 5.1 c, left). Additionally, we tested the significance of clustering patterns according to the four null models by applying the same sequential removal of samples, but first randomizing their salinity order (see Supplementary Materials for further details). These randomizations provided a 95% confidence interval for the clustering index (Figure 5.1 c, right) that allows you to decide if the observed clustering pattern as the environmental variable decreases is related to the environmental factor or not.

5.2.2 Statistics

Significant breakpoints of observed patterns along environmental gradients, either from ordination analyses or clustering indices, were estimated by means of the maximum F statistic derived from sequential Chow-tests (*sctest* function from *strucchange* package, Zeileis et al. 2003) (see Figure C.3).

5.3 Results

We developed a detailed theoretical and methodological framework for the RTCC method (see Figure 5.1, the methods section for more details, and Figure C.1 for a summarized workflow of the methodology). To test the potential of this trait-based approach for gaining new insights about different community assembly mechanisms, we explored microbial communities along an environmental (salinity) gradient in a set of shallow saline ponds. We first evaluated 10 genomic traits that were available in public databases and that summarized genomic structural/functional variability, some of them known to respond to ecological adaptations (Figure 5.2 and see Supplementary Material, section *Matching genomes*). Among them, the percentage of DNA dedicated to signal peptide synthesis exhibited the strongest clustering signal. Therefore, this trait, characterizing each species within a sample,

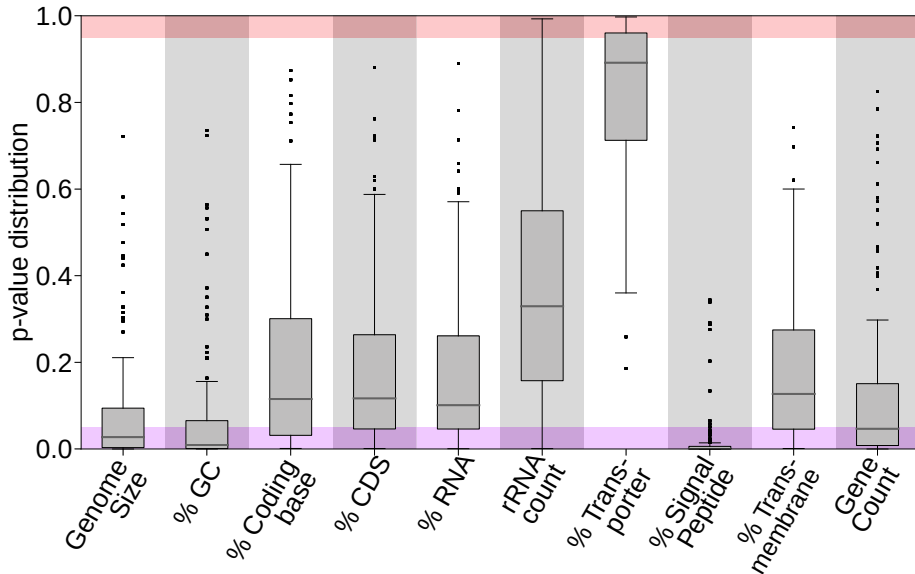


Figure 5.2: Clustering distribution patterns for different functional traits. The p-value distributions obtained by the random assembly null model for 10 functional traits are shown. Colored stripes show the ranges of significant over-dispersion ($p > 0.95$) and significant clustering ($p < 0.05$). Signal peptide exhibit a strong signal of clustering.

was primarily used to further conduct the RTCC analyses. The pairwise differences for this trait were significantly smaller than those expected in synthetic communities randomly assembled from the pool according to the first null model (see *Methods* section below).

To assess the role of environmental factors in sorting species from empirical communities along the gradient, we analyzed the relationship between clustering patterns and two environmental variables with a battery of assembly null models (see the *Methods* section). Salinity and wind velocity should have different impacts on community assembly and, therefore, result in distinct patterns of clustering index values. We use wind as a control variable because it should represent a stochastic continuous perturbation, while salinity would provide a deterministic influence. We represented clustering index values by sequentially removing samples in decreasing order of the environmental variable (see Figure 5.3 a, b). We then tested the

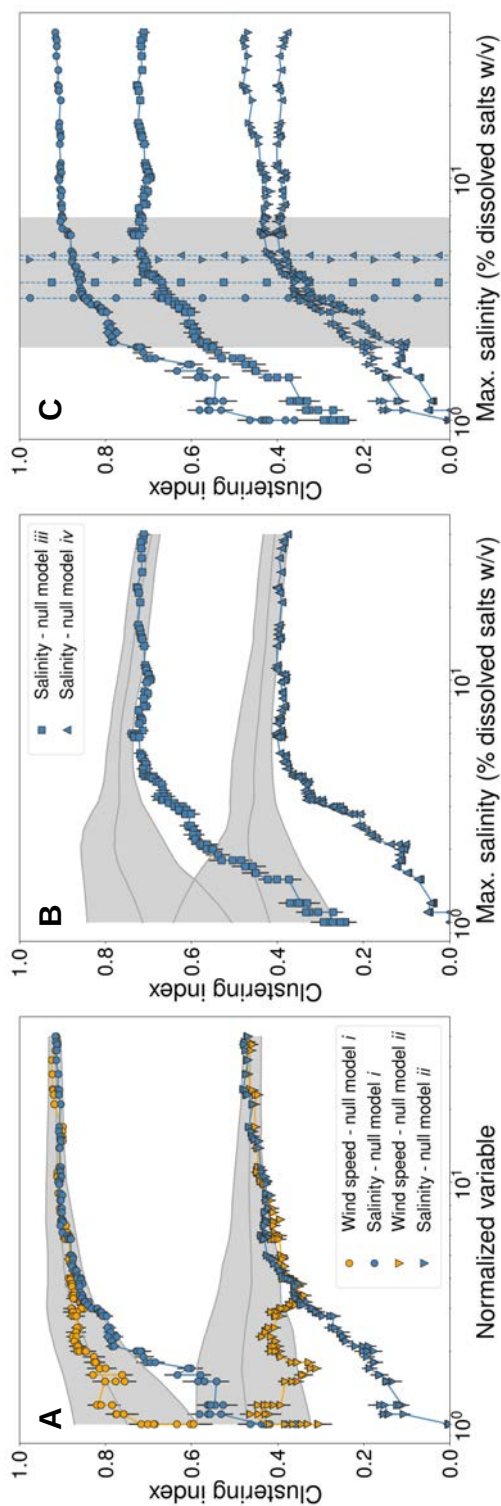


Figure 5.3: Null models for community assembly. (A) Randomized signal peptide community clustering in salinity and maximum wind speed according to null models (i) and (ii), random assembly and abundance-based assembly, respectively. The clustering index for the signal peptide trait is plotted versus the maximum value of the environmental variable among all samples still present after every removal step. Samples are sequentially removed one-by-one in decreasing order of the variable. Note that wind velocities have been normalized so that their maximum values correspond to the maximum salinity observed. The shaded gray areas correspond to confidence intervals for the different null models when the random removal of samples is repeated multiple times. (B) Same as A for null models (iii) and (iv), environmental range-based assembly, and environmental range and abundance-based assembly, respectively, which are meaningful only for salinity. (C) Clustering index curves are plotted again along with their thresholds, all of them lying within the rectangle (shaded in gray) yielded by the OTU-based approach, whose limits corresponds to the significant breakpoints of the observed pattern along the salinity gradient (Figure C.6).

significance of the environmental variable by comparing the declining patterns with replicates of the same methodology based on random orderings (95% confidence intervals are shown as shaded areas). The rightmost point of the curves corresponded to the clustering index for all samples in the data set. The random assembly null model yielded a clustering index value $h \sim 0.9$, which means that around a 90% of samples were not compatible with random assemblages, that is, showed an average trait dissimilarity significantly different from null model expectations. In other words, the vast majority of empirical samples exhibited a high degree of clustering in signal peptide compared to purely random assemblages. Such level of clustering declined when synthetic communities were built by selecting species with a probability proportional to relative abundances, and when ranges of salinity were taken into account in the assembly of simulated communities. As null model complexity increased, the fraction of samples compatible with the null hypothesis increased (see Figure 5.3 b). For the most constrained model, which takes into account both species abundances and salinity ranges, only at most $\sim 30\%$ of samples were not compatible with the null hypothesis underlying this model (see Figure 5.3, lower curve). All null models showed a plateau in clustering index as the environmental variable decreased, but, at some point, a sharp decline occurred. The curves for salinity lay outside the confidence interval expected for randomized orderings, hence the null model was not able to explain the variation of clustering as salinity decreased. This result is consistent with a deterministic role of the salinity environmental gradient in community assembly. We repeated the same methodology using wind speed as a control environmental variable. Wind velocity showed a different pattern from salinity, that is, white curves laying within the shaded area in both tests (Figure 5.3 a). Since clustering values for ordered sample removal lie within null model confidence intervals for randomized orderings, wind speed is not a deterministic driver of community assembly, but, as expected, plays a stochastic role (Figure 5.3 a). Although the OTU-based ordination approach revealed a weak positive correlation signal, this marginal correlation is well below the strong correlation observed for salinity (compare D and B panels in Figure C.6).

Chow tests were used to determine a threshold value in salinity over

which environmental constraints emerge as a leading driver. We found percentages of salinity ranging from 3.2% for the random assembly null model and 4.8% for the ranges and abundances null model. These values lay within the range of salinity that presented significant breakpoints on the changing rate of communities in the ordination analysis (Figure 5.3 c, see also Figure C.6). We observed that traits other than signal peptide could be used to determine thresholds along the gradient (for example, the percentage of GC pairs in sequences; see Figure C.7). However, this was not necessarily true for any trait (for example, the percentage of DNA associated to transporter proteins, see also Figure C.7). Finally, signal peptide average values across genomes per sample (relative to the metacommunity average) were plotted along the salinity gradient. We observed that average trait values decreased as salinity increased (Figure C.8).

5.4 Discussion

Competitive exclusion increasing local extinctions of subadapted species and strict environmental filtering preventing effective colonization of non-tolerant species along an environmental gradient, lie at two extremes of a continuum. Accurately disentangling these two main drivers is challenging. We developed a method to evaluate the effect of an environmental gradient on community trait structure, and provide a conceptual framework (Figure 5.4) to interpret consistent discrete environmental thresholds that helps separate an assembly regime driven by a rich combination of several biotic and abiotic factors from a regime characterized by strong environmental constraints. The observed threshold unveils the switch point above which a structuring force starts dominating along the gradient (see the highest values of the clustering index along the plateau). Conversely, the sharp decline of the clustering index below the threshold value can be regarded as a signature of the rapid decay of the predominant role of a given environmental variable in driving local community assembly (Figure 5.4). Thus, the lower the environmental pressure the lower the proportion of samples exhibiting significant clustering (i.e., the lower clustering index) due to the increased dispersion of trait values within local samples.

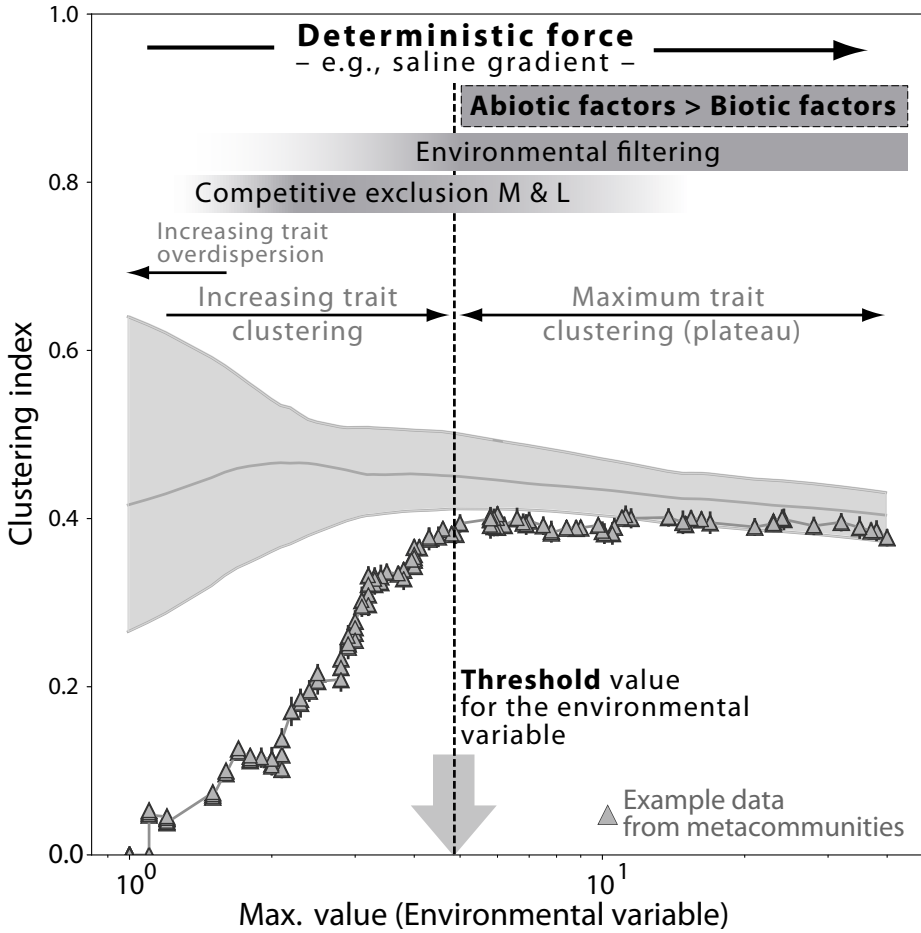


Figure 5.4: Conceptual framework. A conceptual model for the interpretation of results based on the Randomized Trait Community Clustering (RTCC) method. We indicate the type of dominant deterministic forces acting along a gradient of a meaningful environmental variable. The picture shows the pattern observed along the salinity gradient by using null model (iv). The shaded gray area corresponds to confidence intervals for the same null model but randomizing sequential sample removal. A significant correlation of the trait clustering pattern with an environmental gradient exists if clustering index values for a given trait (under ordered sequential removal) lie outside the shaded area, this is, the null model fails to explain the variation of clustering along the gradient. The vertical line is the threshold salinity value separating the two community assembly regimes.

Trait-based approaches have been widely used by ecologists to identify general community-level patterns of macro-organisms based on phenotypic characters (Adler et al., 2013; McGill et al., 2006; ter Braak et al., 2012; Thakur and Wright, 2017). Some authors have made considerable efforts to introduce this framework in the study of microbial communities (Barberán et al., 2012b; Dini-Andreote et al., 2015; Krause et al., 2014; Ortiz-Álvarez et al., 2018a), among others. Burke et al. (2011) suggests that gene function, rather than the RNA taxonomy approach, holds the key to properly studying bacterial community assembly. Community composition may vary through several processes, while community functionality may remain stable due to functional redundancy of taxa (Grossmann et al., 2016; Louca et al., 2018). If stochastic forces are not too strong, selected species traits are expected to be differently favored and cluster around certain optimal values along a structuring environmental gradient. The distribution of trait values across species along a given gradient will be the result of a balance between environmental filtering and biotic interactions. Some traits can be unambiguously related to either stress-tolerance or competitive abilities. In this case, they can be used to distinguish between environmental filtering vs. competitive exclusion (Goberna et al., 2014; Kraft et al., 2015) in spite of these two process producing the same trait clustering patterns (Mayfield and Levine, 2010).

Our findings indicate that the relative amount of genomic DNA coding for signal peptide shows a significant clustering pattern along the salinity gradient. Signal peptides are short peptides (5–30 aminoacids) present at the N-terminus of newly synthesized proteins that control the final destiny of these proteins, this is, whether they end up anchored to the cell membrane or are extracellularly secreted (von Heijne, 1985). The diversity of these signaling systems decreases along salinity gradients in favor of the signal peptide-dependent Tat system, which dominates under high salinity conditions (Rose et al., 2002; van der Ploeg et al., 2011, 2012). The observed decrease in average signal peptide values per sample along the salinity gradient (see Figure C.8) suggests that this trait could be involved in salinity stress tolerance, although confirming this requires further experimental research. Likewise, as salinity increases, we also observed higher similarity

in GC content across locally co-occurring species (see Figure C.7). Several studies have suggested that hypersaline inhabitants are characterized by high GC content although there are some exceptions to these general patterns (Paul et al., 2008). Moreover, it has been suggested that the stress effect of extreme environments is reflected in the nucleotide composition (Barberán et al., 2012b). Other functional traits are not critically influenced by this gradient. In other words, some traits with relevant ecological roles may be related to other functional strategies not linked to the main environmental gradient, such as, for instance, the rRNA operon copy number, known to reflect bacterial growth rates (Klappenbach et al., 2000), which are related to competitive ability (Krause et al., 2014).

Salinity has been described as one of the major environmental drivers of microbial community composition at a global scale (Auguet et al., 2010; Barberán and Casamayor, 2010; Lozupone and Knight, 2007). Important ecological changes have been described along salinity gradients from multipond solar salterns (Gasol et al., 2004), with decreasing biodiversity as salinity increases (Casamayor et al., 2013), with an accompanying loss of metabolic processes (Oren, 1990). Here, the specific trait pattern reveals the ecological functional adaptation of bacterial communities to the salinity gradient. Structural changes shown under a taxonomic view followed a similar pattern along the gradient (Figure C.6 B). An increase in salinity beyond a threshold close to 5% triggers both a rapid change in community composition, and a decrease in mean relative signal peptide values (Figure C.8, upper panel). Maximum trait clustering was also reached at the same threshold. Above 5% salinity, community trait structure tends to stabilize, which means that there is functional redundancy in taxa while community composition still changes, but showing a slower turnover.

Between ca. 10% and 40% salinity range, community composition remained more stable. This shows that the RTCC method provided information not only about functional adaptation of communities—complementing results from their taxonomic turnover—but also about the accurate position of a salinity threshold (5%) over which environmental constraints strongly shape community assembly. Since traits are selected based on the degree of clustering signal rather than on previous trait knowledge, our method

opens new possibilities to objectively determine both whether or not and where a functional adaptation occurs along an environmental gradient.

It is reasonable to think that, in principle, if an environmental variable changes very smoothly, community composition, and, consequently, community trait structure, should also change smoothly. However, the opposite seems to be true in microbial communities along a salinity gradient. The RTCC accurately finds the position of a salinity threshold.

Ecological theory predicts that the end points of community assembly can be multiple (Bunin, 2017). The possibility that several stable species assemblages can exist under the same environmental conditions opens the door to observing abrupt transitions in community composition as a single environmental variable slowly changes. The study of the origins and underlying causes of these break points in community composition along environmental gradients deserves further empirical and theoretical work.

To conclude, we provide a new conceptual approach with the RTCC method that could be useful in guiding hypothesis-driven studies coping with the high complexity of biological systems. Although we used a data set from microbial communities, we highlight the potential of this method to be applied more generally to quantitative phenotypic or genotypic trait data measured for macroorganismal communities as well. A combination of theoretical modeling through trait-based analyses allows us to go beyond description and deepen our mechanistic understanding of community assembly. Quantifying how community assembly is shaped by the environment is critical to predicting the ecological impact of environmental changes, and our approach provides a powerful tool to advance this analysis.

CHAPTER FIVE



Mateu Menéndez-Serra, Vicente J. Ontiveros, Joan Cáliz, David Alonso and Emilio O. Casamayor. Understanding stochastic and deterministic assembly processes in microbial communities along temporal and environmental gradients. *Molecular Ecology* (submitted).

The bibliography of this chapter can be found on the unified *Bibliography* section.

Understanding stochastic and deterministic assembly processes in microbial communities along temporal and environmental gradients

Abstract: Identifying the main drivers of community assembly remains an open fundamental question in microbial ecology. Dispersal processes introduce randomness in community composition while selection for particular environments creates predictable assemblages. However, the interaction of selection and dispersal processes is still poorly understood. Here, we address this question in bacterial and microeukaryal communities inhabiting a highly dynamic system of ephemeral (hyper)-saline ponds. We showed that the combination of beta-diversity decomposition methods and a novel temporal approach yield new insights into the relative effect of selection and dispersal along environmental gradients. Selective pressure and dispersal-related processes simultaneously shape each local community with variable strength and net effect. The dominance of selection vs dispersal changed along the gradient, shifting from stochastic to deterministic assembly as salinity increased. Likewise, temporal community turnover decreased at high salinities because of both colonization and extinction rates slowed down. Only eukaryal richness decreased along the gradient due to lower effective colonization at higher salinities, suggesting that the net effect of selection and dispersal is determined by environmental conditions and the idiosyncratic microbial ecologies. Our results emphasized the use of novel temporal approaches in combination with standard statistical methods for a better understanding of dynamic processes underlying community assembly.

6.1 Introduction

Mechanisms driving community assembly have been recurrently studied as one of the major questions in ecology (Legendre et al., 2005; Leibold et al., 2004; Shmida and Wilson, 1985). According to the framework proposed by Vellend (Vellend, 2010), all factors affecting community assembly can be condensed in four main processes: selection, drift, dispersal, and speciation. In this context, selection is defined as a process by which individual abundances change due to the different deterministic fitness among organisms, which will create predictable communities based on particular environmental features. Conversely, other processes are neutral and introduce randomness in the assembly mechanism. Drift is defined as variations in species abundances due to stochastic fluctuations in birth/death events, while speciation creates new species by random mutations. Dispersal, i.e. the movement of organisms across space, also introduces randomness in the community assembly. Several theories on the organization of biodiversity over space and time have been postulated according to the relative importance of each process (Etienne and Alonso, 2007; Wennekes et al., 2012). The balance between determinism and stochasticity has historically led to two contrasting theories, niche-based and neutral theories, and several authors have tried to unveil which process dominates across diverse communities and environments (Hubbell, 2001; Leibold et al., 2004; van der Plas et al., 2015). Empirical support alone for the prevalence of any of the two theories remains controversial (Fierer and Jackson, 2006; Logue and Lindström, 2010; Van der Gucht et al., 2007), suggesting that there is a balance between neutral and niche processes depending on the idiosyncrasy of the studied system (Gravel et al., 2006).

In nature, ecological processes very often operate in a metacommunity context, and expanding the analyses from single communities to metacommunities permits to evaluate community assembly at a multiscale level (Leibold et al., 2004). This framework proposes a multidirectional network of local communities linked by dispersal processes (Holyoak, 2005). Dispersal has particularities in microbial organisms as it is mostly driven by external factors (Barberán and Casamayor, 2014). Recent descriptions of geographical patterns in microbial distribution have challenged the accepted statement

“everything is everywhere, but the environment selects” (Baas Becking, 1934), conferring higher importance to dispersal processes on microbial community assembly. The main signal of the importance of dispersal is a characteristic pattern of distance-decay of similarity (Nekola and White, 1999) that has been observed repeatedly across many microbial communities, on both continuous (Habtom et al., 2019) and discrete (*“islands”*) scenarios (Darcy et al., 2018). The interaction between selection and dispersal is crucial to better understanding community assembly as strong selective pressures may limit the colonization success of particular organisms, and very few studies have simultaneously assessed their relative weight in microbial communities (Logares et al., 2018; Ofițeru et al., 2010; Van Der Gast et al., 2008; Wu et al., 2018). However, it remains unclear still what is the effect of both microbial and environmental idiosyncrasies shaping the relative importance of selection and dispersal. In addition, the lack of intensive temporal surveys makes it very difficult to study how these processes influence community temporal dynamics. In this sense, the development and application of novel dynamic approaches would be very welcome.

In this investigation, we studied a set of shallow lakes characterized by wide salinity variations throughout the hydrological cycle, which allows to evaluate bacteria and microeukaryotes assemblies under increasing selective pressure along an stress gradient. We hypothesized that the idiosyncratic microbial ecologies might led to a differential response to the dominant shaping processes. We applied a simple metacommunity framework (Leibold and Chase, 2018; Leibold et al., 2004) rooted in the Island Biogeography Theory (MacArthur and Wilson, 1963, 1967), combining standard statistical methods with a novel temporal approach (Ontiveros et al., 2019) that estimated colonization and extinction rates for each salt lake and microbial group (i.e., the characteristic time (Ontiveros et al., 2021)). This may help to improve our understanding on the mechanisms underlying the observed dynamics, and on the differential effect of these processes in the environmentally-driven temporal variability of microbial communities.

6.2 Materials and methods

6.2.1 Study site, sampling and genetic analyses

We studied the effect of selection and dispersal processes on bacteria and microeukaryotes inhabiting a group of ephemeral salt ponds in the Monegros Desert. The Monegros Desert area (Central Ebro Basin, NE Spain, 41°42'N, 0°20'W) contains one of the largest sets of inland salt lakes in Europe and includes saline depressions of variable sizes (<2 to 239 ha) located at the same elevation without physical dispersal barriers (Figure 6.1). Distance between ponds ranged between <1 km to 24 km². This system is characterized by wide oscillations of water volume which leads to severe variations of salinity both among and within ponds. Salinity is one of the main ecological factors shaping microbial communities at a global scale (Auguet et al., 2010; Lozupone and Knight, 2007) and has been proved to be a key factor determining the distribution of archaeal (Menéndez-Serra et al., 2020), bacterial and eukaryal (Menéndez-Serra et al., 2021) populations in these salt ponds. An increase of salinity exercises a tight control on the microbial abundance and growth rate (Gasol et al., 2004) and on alpha diversity indicators (Casamayor et al., 2013), which makes this system a perfect scenario for analysing the responses to differential selective pressures. A total of 144 water samples were analysed from 14 different shallow ephemeral ponds sampled monthly over three consecutive calendar years, covering different dry–wet hydroperiods as recently described (Menéndez-Serra et al., 2021). Conductivity, temperature, pH, dissolved oxygen, and redox potential were measured in situ using an HQ40d multiparameter meter (Hach, Loveland, CO, USA). Salinity was estimated in situ by a hand refractometer (Atago S-28E, Japan). Sampling information and the whole physico–chemical dataset have been previously reported in detail (see Table 3.1 and supplementary Table A.1 in Menéndez-Serra et al. 2020).

Bacteria (size fraction 0.2–5 μm) and microeukaryotes (size fraction 5–50 μm) were independently analysed after PCR–amplified 16S (V4 region, 515f–806r) and 18S (V9 region, 1391f–1510r) rRNA genes and high–throughput sequencing using the Illumina MiSeq platform (Michigan State University, <https://rtsf.natsci.msu.edu/>), respectively. Raw sequences

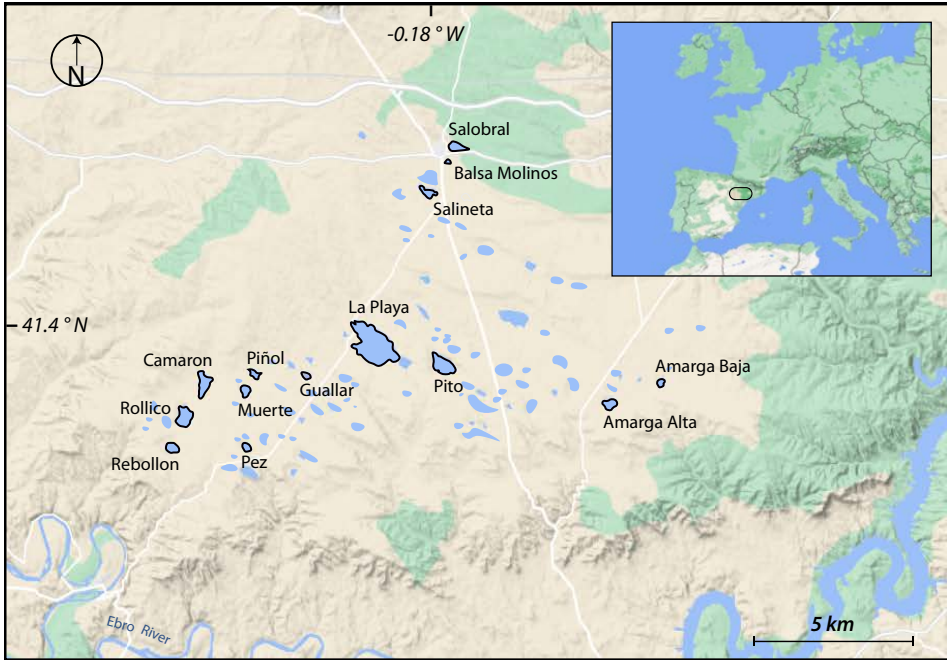


Figure 6.1: Monegros desert plateau in NE Spain covering a endorheic area of 400 km². The 14 sampled sites are highlighted. Map data: Google, DigitalGlobe.

were processed using the UPARSE pipeline (Edgar, 2013). Sequences were merged, quality filtered, and chimera removed. The UNOISE algorithm (Edgar, 2016) defined the Operational Taxonomic Units at 100% identity, i.e., zero-radius OTUs (zOTUs). The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012), using the SILVA 132 reference database (Gurevich et al., 2013). The resulting zOTU tables contained 5653 bacteria and 2658 eukaryotes. See more details in Menéndez-Serra et al. (2021). The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject ID PRJNA429605.

6.2.2 Environmental and spatial analyses

A connectivity index was calculated for each site (pond) considering distance to the surrounding ponds and the respective areas. This index is based on the connectivity index proposed by Hanski (1994) further simplified

by Reche et al. (2005). The influence of environmental variables in the microbial composition and structure (alfa- and beta-diversity) was preliminarily assessed by correlation approaches, Mantel tests and Non-Metric Multidimensional scaling ordinations based on Bray-Curtis dissimilarity index.

In order to test the combined effect of salinity and spatial distance, the Bray-Curtis dissimilarity index between samples was linked to the additive effect of both scaled salinity differences and spatial distances between each sample and pond, respectively. This relationship followed a nonlinear trend which best fitted to a sigmoid function:

$$BC = BC_0 + \frac{1 - BC_0}{1 + e^{-(\alpha S + \beta G - I)}} \quad (6.1)$$

where, for each pair of samples, S_{ij} G_{ij} stands for salinity difference and spatial distance between samples i and j , respectively, BC_0 represents a reference value related to the Bray-Curtis dissimilarity index when both salinity difference and spatial distance are equal to zero, α and β are parameter weights estimated independently for each microbial domain, and I is the function inflection point. All parameters were estimated using the *nls* function within the ‘stats’ package version 3.5.2. The relative contribution of each factor was assessed by defining two new variables that represent the weighted scaled salinity difference and spatial distance across pairs, αS and βG respectively, and comparing their distribution average values through a t-test. Model fit was estimated by Spearman’s correlation between predicted and observed values.

The influence of the environmental and spatial variables in the structure of microbial communities (beta-diversity) was also evaluated by Variation Partitioning Analysis (VPA) (*varpart* function in “vegan” package). VPA assigned a percentage of community composition variability among samples to environmental effect [E|S], spatial effect [S|E] and the interaction component of them [E $\cap$ S]. A pairwise distance matrix based on Bray-Curtis dissimilarity index between samples was used as response matrix. For spatial data, Principal Coordinates of Neighbour Matrices (PCNMs) were

calculated (*pcnm* function in “vegan” package) from geographic coordinates in order to generate numerical variables which represented multiple spatial scales. Both environmental variables and PCNMs were forward selected by permutation tests using the *ordistep* function also in “vegan”. In order to test the influence of both processes at different salinity levels, VPA was also independently performed over low (0.1–4.9% w/v) and high salinity samples (5–40% w/v). These salinity ranges were previously established based on an unveiled environmental threshold after a trait-based approach (Triadó-Margarit et al., 2019). To contrast the relative weight of both selection and dispersal processes among microbial groups and samples configurations, we compared the environmental to spatial effect ratio ($[E|S]/[S|E]$) (Wu et al., 2018) rather than absolute values which might be influenced by species pool size (Kraft et al., 2011) and noise background (Stegen et al., 2013).

6.2.3 Dynamic analysis

A characteristic time for each salt pond was calculated using the “island” package version 0.2.2 on R (Ontiveros et al., 2019). This characteristic time is based on “colonization” and (local) “extinction” rates estimated from presence-absence time series transitions on consecutive samplings for all the zOTUs present in the pond along the whole temporal survey. It is expressed in days and is inversely related to the turnover rate of the community (i.e., more dynamic communities show lower characteristic times (Ontiveros et al., 2021)). The relationship between colonization and extinction rates also provides an effective expected richness at equilibrium as a fraction of the global pool richness, which is defined as:

$$S = \frac{1}{1 + \frac{e}{c}} \quad (6.2)$$

When rates are assumed constant in time, this expected richness remains also stable. However, environmental conditions may influence colonization and extinction processes. Therefore, we added salinity as a factor shaping colonization and extinction rates which, in turn, resulted in environmentally-driven temporal variability of microbial richness. Here, salinity was introduced on the original model as two independent components influencing colonization (c) and extinction (e) rates at each time:

$$c_t = C_0 + C_L S_{Lt} + C_H S_{Ht} \quad (6.3)$$

$$e_t = E_0 + E_L S_{Lt} + E_H S_{Ht} \quad (6.4)$$

where, for each time, S_{Lt} captures salinities at the low range (from 0 to 4.9%, w/v) and S_{Ht} captures salinities at the high range (from 5 to 40%, w/v). This implies that a sample with salinity equal to 3‰ will be characterized by the variable S_{Lt} equal to 3 and S_{Ht} equal to 0, while for a sample with salinity equal to 10‰, S_{Lt} will be 0 and S_{Ht} will be 10. C_0 and E_0 represent basal colonization and extinction, respectively, while C_L , C_H , E_L and E_H are model parameters estimated for each pond and microbial community (considering all sampling times) which capture the effect of salinity over colonization and extinction. In order to make colonization and extinction parameters comparable across ponds and groups, we standardized each model parameter dividing it by their corresponding basal colonization (C_0) or extinction (E_0). Next, we evaluated the relative contribution of each salinity component to the total c and e as the product of the standardized rate and the mean salinity of the pond for each salt range. Model fit was evaluated by Akaike information criterion (AIC). Model predictability improvement after adding salinity was quantified by differences in mean square deviations of model predictions from actual observations (*simulated-model* function from “island” package).

6.3 Results

6.3.1 Selection by environmental filtering

We found significant correlations between zOTUs distance matrices based on Bray–Curtis dissimilarity index and between-sample pairwise Euclidean salinity distance for both bacteria and microeukaryotes (Mantel, R: 0.450 and 0.374, respectively). The same test showed weaker correlations with pH (Mantel, R: 0.155 and 0.198), redox conditions (Mantel, R: 0.154 and 0.103) and dissolved oxygen (Mantel, R: 0.199 and 0.163), respectively.

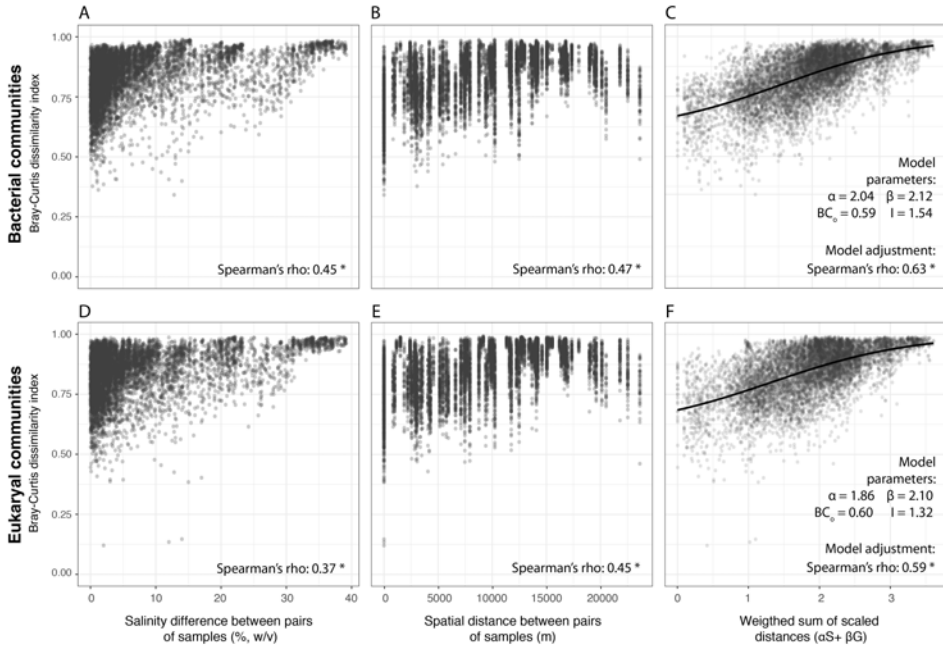


Figure 6.2: Representation of Bray-Curtis dissimilarity index against pairwise salinity difference (A, D), spatial distance (B, E) and the weighted additive effect of both factors (C, F). Model prediction is plotted over the observed data in panels C and F.

The microbial community composition similarity strongly decreased as salinity difference increased (Figure 6.2, panels A for bacteria, and D for microeukaryotes). However, for narrow salinity differences a high variability in the Bray-Curtis index was observed, reporting highly differentiated communities between similar salinity samples, mostly at low-medium salt concentrations (Supplementary Figure D.2 and D.3). At high salinities, we consistently observed high similarities in the microbial communities for all the sampling period (Supplementary Figure D.3).

6.3.2 Dispersal

Community composition showed a consistent decay of similarity (Bray-Curtis index) with geographic Euclidean distance between samples (Figure 6.2, panels B and E) with similar correlation values for bacteria and microeukaryotes (Spearman's rho 0.47 and 0.45 respectively, p-value <

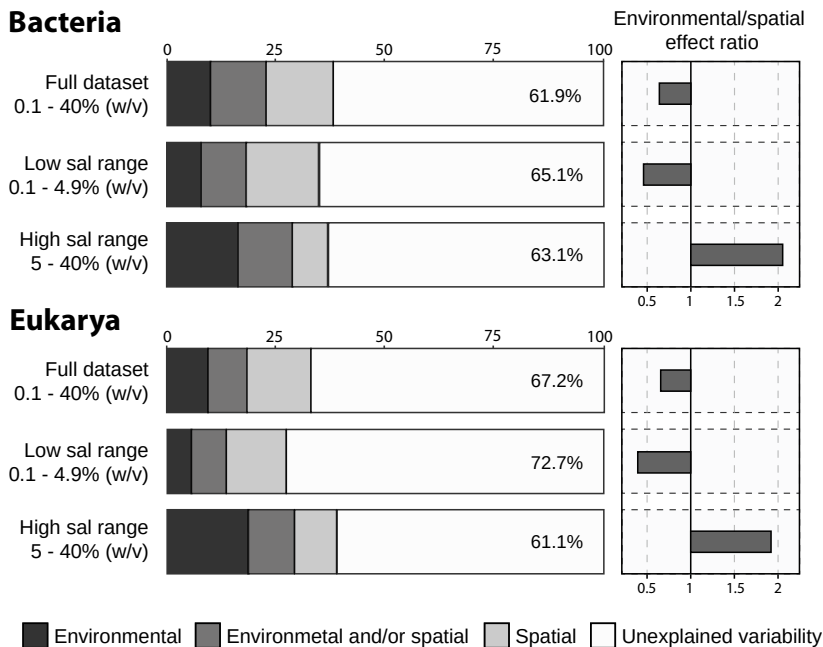


Figure 6.3: Community composition variability explained by the Variation Partitioning Analysis for bacteria and microeukaryotes communities, including the environmental to spatial ratio ($[E/S]/[S/E]$) for each metacommunity and salt range (right panels).

0.001). The connectivity index calculated for each pond showed a positive correlation with the mean bacterial richness of the pond considering all sampling times (Pearson's correlation = 0.50, p-value = 0.07). Interestingly, for microeukaryotes this correlation was not found.

6.3.3 Relative weight and combined effect of selection and dispersal

Decomposition of beta-diversity of the whole metacommunity by VPA showed similar results for both bacteria and microeukaryotes. We found that spatial factors explained more variability than environmental factors for both communities (see percentages in Figure 6.3 and Supplementary Table D.1). The similar environmental to spatial ratios observed for bacteria and

microeukaryotes indicated a similar effect of spatial and environmental factors on the two communities. This fact was also evidenced after representing the Bray-Curtis pairwise dissimilarities for both communities against each other (Figure 6.4), showing similar values in most cases (Mantel, R : 0.74). However, if the metacommunity was splitted into the two salt ranges, the environmental to spatial ratios showed a change in the explanatory factors. Thus, spatial factors explained higher variability at low salinities whereas environmental conditions explained better at high salinities. Most probably, the low environmental/spatial ratio found for the whole metacommunity might be influenced by the unbalanced distribution of samples along the salinity gradient (i.e., up to 67% of samples had salinities $< 5\%$, w/v).

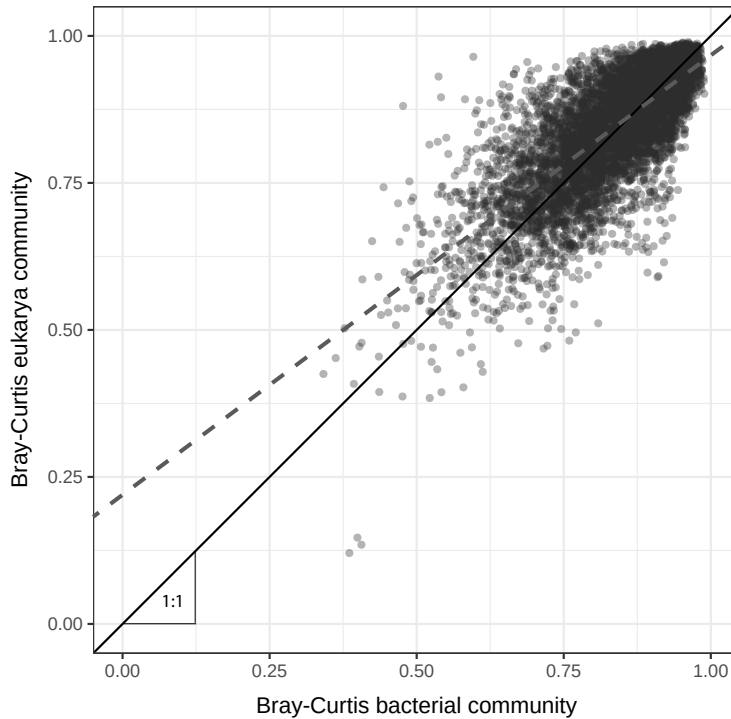


Figure 6.4: Bray-Curtis dissimilarity index of bacterial and microeukaryotic assemblages for each pair of samples. Black line highlights the 1:1 diagonal and the best linear fit is showed in dashed line (slope = 0.75; $R^2 = 0.56$).

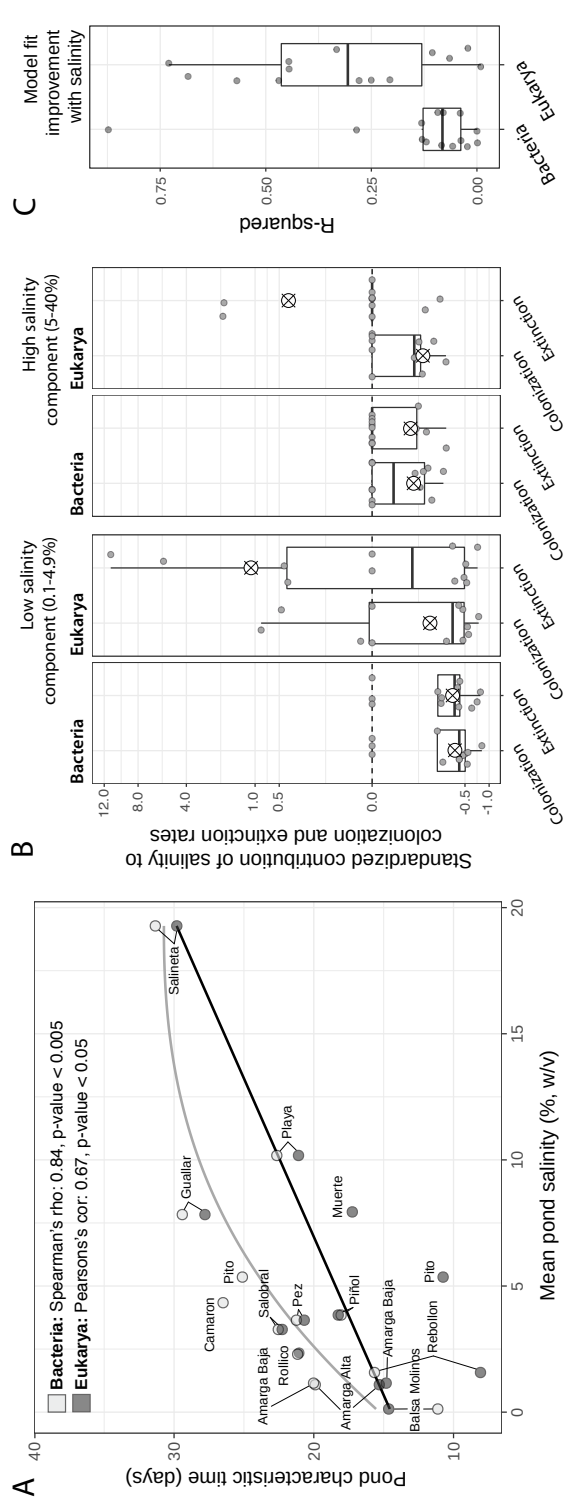


Figure 6.5: Characteristic time for all ponds and both microbial communities against the mean pond salinity, considering all sampling times (panel A). Relative effect of salinity over colonization and extinction rates (panel B). The standardized contribution, compared basal colonization/extinction, is represented for each salt range and microbial group. Crossed dots represent rates mean values (panel B). Differences of the dynamic model fit predicting sample richness along consecutive sampling times when considering salinity variations. Since R-squared can range between -1 and 1, values above 0 indicate an improvement of the fit compared the model without salinity in all cases (panel C).

The representation of the additive effect of salinity difference and spatial distance against Bray–Curtis dissimilarity index showed a strong decay in community composition similarity as the sum of both factors increased (Figure 6.2, panels C and E). By comparing the weighted scaled variables (pairwise salinity differences and spatial distances, αS and βG , respectively, see Eq 1), we found that the weight of the spatial variable was higher than the salinity variable for predicting Bray–Curtis dissimilarity in both communities (Student’s t-test, $t=68.76$ for bacteria and $t=80.34$ for microeukaryotes, p -values < 0.001). According to the model, the expected reference value of Bray–Curtis index when salinity differences and spatial distance were zero (BC_0) was close to 0.6 for both communities. The incorporation of other environmental variables, as dissolved oxygen and pH, into the sigmoidal model was tested but model fit did not significantly improve (data not shown).

6.3.4 Dynamic behavior

We noticed that characteristic times consistently increased with salinity for both bacteria and microeukaryotes communities (Figure 6.5 A), being bacterial communities generally more stable (higher characteristic time). For most cases, Bacteria and microeukaryotes of the same pond showed similar community dynamics.

While colonization and extinction rates captures pond dynamism, the extinction to colonization ratio e/c ultimately determines average community richness at equilibrium. We analyzed how salt variations affected the expected richness within each pond and each type of microbes by comparing the product of the standardized model parameters capturing the effect of salinity over colonization and extinction rates, and the pond mean salinity for each range (see Eq. 3 and 4). Interestingly, salinity variations differentially affected bacteria and microeukaryotes community dynamics (Figure 6.5 B). For bacteria, a similar effect of salinity into colonization and extinction rates resulted in a constant e/c ratio along the gradient, i. e., there was no significant variation of the overall metacommunity richness along the salt gradient. The contribution of salinity to bacterial colonization and extinction rates was negative (or zero) for all ponds and salt ranges, which

means a decrease of colonization and extinction rates as salinity increases, this is, longer characteristic times and, therefore, lower temporal community turnover.

Conversely, for microeukaryotes the effect of salinity was more heterogeneous among ponds. The contribution of the low salinity component to colonization and extinction showed both positive and negative values. This fact implies that some ponds turnover did not decrease with salinity, which explains the lower observed correlation compared to bacteria (Figure 6.5 A). For the high salinity range, only two ponds (Muerte and Piñol) presented a positive contribution of salinity to extinction, while the contribution to their colonization was negative or zero. For most ponds, the differential contribution of salinity to colonization and extinction rates translated into e/c ratio increasing along the gradient, which resulted on a significant decrease of the overall metacommunity richness as salinity increased (see Supplementary Fig S1, and Figure 1 in Menéndez-Serra et al. 2021).

Finally, we compared the observed richness along the temporal series with the expected richness for each pond and time based on the dynamic models, with and without salinity as a factor shaping colonization and extinction rates. Model predictability increased if salinity variations were considered (Figure 6.5 C and Supplementary Figure D.4). However, predictability was significantly higher for microeukaryotes than for bacteria, showing again a higher relationship between salinity and microeukaryotic richness. In addition, the highest improvement in model predictability was observed for the ponds submitted to the widest salinity oscillation (Supplementary Table D.2).

6.4 Discussion

Understanding the dominant processes shaping metacommunity assembly is extremely challenging due to the high complexity of ecological systems. Historically, the idiosyncrasy of microbial systems has limited the description of temporal community variations to a sequence of changes in terms of diversity and composition, with limited knowledge of the mechanisms driving assembly changes. The lack of a consistent methodological framework for

evaluating the effect of the different shaping processes makes it very hard to compare results among studies and environments, in addition to limitations in the sampling designs (number and distribution of samples, spatial and environmental scales, etc.) (Viana and Chase, 2019) and the use of a particular quantification method that, altogether, strongly influence the study conclusions. Furthermore, it has been recently shown that co-existing microbial assemblages may have differential responses against the same assembly mechanisms (Logares et al., 2018; Wu et al., 2018).

The possibility to assign a single dominant process to explain compositional differences (beta-diversity) between pair of samples has been shown as an effective way to summarise the main factors driving metacommunity assembly (Logares et al., 2018; Stegen et al., 2013), although it might imply a substantial oversimplification. It is well accepted that selection, dispersal and drift simultaneously determine community composition with variable strength and direction in space and time, and with complex interactions among them (Vellend, 2010). However, the mechanistic way driving the interaction of these processes remains still unknown. In the present study, we combined statistical methods for decomposing beta-diversity with a temporal analysis based on extinction-colonization dynamics. The novel approach provided useful insights into the mechanistic pathway where selection and dispersal interact to shape community assembly.

We noticed that bacterial and microeukaryotic communities followed similar trends along the salt gradient, space and time. This convergent behavior suggests equivalent effect of the main factors driving the assemblies. A previous study in a set of Eastern Antarctica lakes reported bacterial assembly dominated by selection while the microeukaryal assembly was dominated by drift (Logares et al., 2018). In our study we surveyed highly dynamic systems, with large salinity changes within the same pond in short time periods that limited the effects of differential ecologies and potential local adaptations of specific groups (Menéndez-Serra et al., 2021). However, differences were still observed. Thus, dissimilarity in bacterial communities was more predictable based on salinity differences (Figure 6.2, panels C and F), suggesting a higher effect of unmeasured factors for the eukaryotes. This result agreed with the previously reported higher

sensibility of microeukaryotes to drift (Lankiewicz et al., 2016; Logares et al., 2018), probably due to the lower population densities. Beta-diversity indicators showed higher convergence when selective pressure increased, because the high salt concentration requires of physiological adaptations restricted to specific groups (Oren, 1999, 2011; Ventosa et al., 2014a). The metabolic limitations for maintaining the energetic-demanding adaptations also results in salinity thresholds for certain metabolisms (Menéndez-Serra et al., 2021; Oren, 1999, 2011) and decreased richness as salinity increased. This pattern is especially significant in microeukaryotes, while in bacteria was non-significant, probably due to higher taxa replacement along the gradient (Menéndez-Serra et al., 2021) and higher metabolic plasticity of smaller organisms (Finlay, 2002; Langenheder et al., 2005). Conversely, the pattern of distance decay of similarity seemed to be quantitatively comparable for the two assemblages, most probably because of the size ranges considered in the present study ($< 50 \mu\text{m}$).

Along the stress gradient, we consistently found a change in the balance between community assembly processes. Dispersal dominated over selective pressures at low salinities, shifting to a selection-dominated assembly as the environmental filter increased. Dispersal processes, and potentially also other unmeasured factors as drift, were the main factors explaining the high dissimilarity between samples of similar low salinity. In this sense, the high variability among samples of the same (or close) ponds sharing similar environmental characteristics suggests a prominent effect of unmeasured environmental variables and drift on community assembly. However, the effect of drift potentially decreased as the salinity-induced selective pressure increased, reducing beta-diversity among samples with similar salinity. Although salinity was the environmental variable that better explained variations in community composition, the amount of beta-diversity explained was lower than expected. The unbalanced sample distribution in the Monegros dataset dominated by low salinity samples highly influenced by dispersal processes and the lack of larger number of high salinity samples because of the quick ponds desiccation, may partially explain the higher contribution of dispersal to the overall metacommunity assembly. That is a clear example of how beta-diversity decomposition is strongly dependent on

spatial scale, sample distribution, and spatial autocorrelation (Smith and Lundholm, 2010) and should be carefully considered.

Interestingly the novel dynamic approach applied (Ontiveros et al., 2019, 2021) allowed us to dig into the underlying mechanisms behind the observed patterns. The observed pattern of decreasing colonization and extinction rates as pond salinity increased indicated a higher stability, i.e. a decrease in community composition turnover, which had been previously observed in artificial setups associated to increase in selective pressures (Van Der Gast et al., 2008). The increase on temporal stability together with the homogenization in community composition at high salinities suggests a gradual shift from stochastic to niche-based community assembly, similar to the observed during primary succession under increasing selective conditions (Dini-Andreote et al., 2015). The dynamic model predicted the temporal variation of community richness for each pond as a response to salinity changes. The quality of the fit in the colonization-extinction model was changing among ponds and communities, but highly remarkable considering that only one environmental variable was taken into account. Altogether, these results support the major role of salinity as a major factor for metacommunity assembly, but also the potential of the dynamic model for capturing realistic temporal site dynamics. For bacteria, the comparable effect of salinity on both colonization and extinction rates resulted in overall richness stability along the gradient. By contrast, in microbial eukaryotes, colonization rates decreased more than extinction rates, resulting in lower species richness at higher salinities. This observation can be probably explained by salinity strongly constraining colonization success of microeukaryotes. This hypothesis could be further tested with a temporarily intensive sampling scheme, which may show sporadic appearance of groups with low residence time, evidencing the arrival of potential colonizers with low colonization success. However, this factor did not seem to be limiting for bacteria, and even ponds with higher connectivity presented higher richness mean values regardless their mean salinity.

Overall, we have shown that the metacommunity of a set of ephemeral shallow saline lakes is the result of a complex balance between selection and dispersal with different dominance along the salt gradient. Although these

processes dominated the assembly of both bacteria and microeukaryotes, their different ecologies resulted in a different net effect of selection/dispersal processes and its interactions. While dispersal seemed to affect similarly bacterial and microeukaryotic communities, selection acted differently at the high-end of the salinity range for the microeukaryotic communities, as the environmental filter constrained successful colonization of microeukaryotes showcasing the environmentally-driven temporal variability of the microbial assembly.



CHAPTER SIX

Mateu Menéndez-Serra, Vicente J. Ontiveros, Albert Barberán, Emilio O. Casamayor. Absence of stress-promoted facilitation coupled with a competition decrease in the microbiome of ephemeral saline lakes. *Ecology (in press)*.

The bibliography of this chapter can be found on the unified *Bibliography* section.

Absence of stress-promoted facilitation coupled with a competition decrease in the microbiome of ephemeral saline lakes

Abstract: Salinity fluctuations constitute a well-known high stress factor strongly shaping global biological distributions and abundances. However, there is a knowledge gap regarding how increasing saline environmental stress affects microbial biological interactions. We applied the combination of a probabilistic method for estimating significant co-occurrences/exclusions and a conceptual framework for filtering out associations potentially linked to environmental and/or spatial factors, in a series of connected ephemeral (hyper) saline lakes. We carried out a network analysis over the full aquatic microbiome –bacteria, eukarya and archaea– under severe salinity fluctuations. Most of the observed co-occurrences/exclusions were potentially explained by environmental niche and/or dispersal limitation. Co-occurrences assigned to potential biological interactions remained stable, suggesting that the salt gradient was not promoting interspecific facilitation processes. Conversely, co-exclusions assigned to potential biological interactions decreased along the gradient both in number and network complexity, pointing to a decrease of interspecies competition as salinity increased. Overall, higher saline stress reduced microbial co-exclusions while co-occurrences remained stable suggesting decreasing competition coupled with lack of stress-gradient promoted facilitation in the microbiome of ephemeral saline lakes.

7.1 Introduction

Interactions among organisms are fundamental for the establishment and functioning of ecological communities. However, assessing the existence, magnitude and ecological role of biological interactions is one of the most challenging goals in complex microbial communities (Faust and Raes, 2012). Microbial interactions are diverse and cover competition, facilitation, and cooperation. At the organism and population levels, competition among microbes is prevalent across different ecosystems (Ghoul and Mitri, 2016). The most generalized mechanism relies on the increase of growth rates to better compete for nutrients, space or oxygen (Ghoul and Mitri, 2016). Positive interactions such as cooperation or facilitation have been underexplored in the microbial world, and have been hypothesized to be infrequent (Foster and Bell, 2012). Likewise, stable positive interactions are only expected for organisms with limited metabolic capabilities (Machado et al., 2021) or when highly restrictive conditions prevail (Foster and Bell, 2012).

The Stress Gradient Hypothesis (SGH) proposes that under increasing stress conditions positive interactions are more prevalent, while negative interactions are reduced as stress increases (Bertness and Callaway, 1994). This hypothesis has been supported by a wide repertory of studies since its formulation, mostly in plants (Bertness and Ewanchuk, 2002; Callaway et al., 2002; Maestre et al., 2003). However, there is a current debate regarding the generality of the SGH and the role of different factors for determining expected interactions along the gradient, fueled by some potentially contradictory results (Granda et al., 2012; Maestre and Cortina, 2004). In this context, Maestre and collaborators (2009; 2005) proposed the nature of the stressor and species life history as key factors determining the resulting interaction scheme. The SGH has been mostly tested in plants, but a few studies have also found consistent results for microbes (Lawrence and Barracough, 2016; Piccardi et al., 2019). Though, most of them only consider a few species confined in artificial setups. Most recently, results coherent with the SGH have been found along an increasing elevation gradient for bacteria and fungi (Hernandez et al., 2021). One of the main reasons for the limited number of available microbial studies at the community level

relies on the difficulty for assessing potential microbial interactions along wide environmental gradients.

Community microbial interactions cannot be directly observed, and statistical indirect approaches are needed to infer biological interactions (Faust and Raes, 2012). In this context, network analysis has recently gained importance to approach interactions and their effects on microbial community structure and functioning (Barberán et al., 2012a; Faust and Raes, 2012; Goberna et al., 2019; Hernandez et al., 2021). However, potential artifacts and false positive interactions have historically limited the application of correlation approaches when analyzing large environmental gradients (Connor et al., 2017) although more complex statistical methods have recently tried to circumvent this limitation (Niku et al., 2019).

In the present study, we applied a combination of both a probabilistic method for determining significant co-occurrences and co-exclusions between pairs of species (Veech, 2013), and a conceptual framework for classifying the nature of the observed co-occurrences and co-exclusions (Blois et al., 2014), using a microbial metacommunity studied along a wide salinity gradient. We focused on how both positive and negative potential biological interactions changed along a gradient of physiological stress linked to increasing salinity. This allowed us to test whether or not the SGH applies for complex highly dynamic microbiomes in natural environments, reproducing the previously observed responses in plant communities and simplified microbial consortia.

7.2 Methods

7.2.1 Study site and molecular dataset

We analyzed a recently reported spatio-temporal survey on planktonic bacterial, eukaryal and archaeal communities inhabiting the Monegros Desert lacustrine system (NE Spain, 41° 42'N, 0° 20'W) (Menéndez-Serra et al., 2020, 2021). Sampling was carried out monthly in 14 different ephemeral lakes along three hydroperiods, showing large salinity variations ranging from freshwater to salt saturation. The prokaryotic (bacteria and archaea) and microeukaryotic communities were analyzed by PCR amplification and sequencing of 16S and 18S rRNA genes, respectively. See Supplementary

Material for a detailed description of samples processing, sequencing protocols, raw sequences processing for zero-radius operational taxonomic units (zOTUs) definition, taxonomic assignment, and environmental data.

7.2.2 Network construction

Ecological networks were constructed based on pairwise zOTUs co-occurrences and co-exclusions for 116 samples. Only zOTUs with occurrences $> 5\%$ (i.e., detected on at least 6 samples over 116) and relative abundances $> 0.1\%$ in at least one sample, were considered. We applied a probabilistic model (Veech, 2013) in order to state whether two zOTUs were observed together significantly more or less frequently than expected by chance (Figure E.1). Pairwise zOTUs associations were classified as positive (co-occurrence), negative (co-exclusion) or random. Link strength was defined as $s = \log(p_0/p)$, where p_0 is the hypothesis testing confidence level (set here at 0.01) and p is the observed co-occurrence/exclusion probability, being s proportional to the order-of-magnitude difference between p and p_0 (Martinez-Romo et al., 2011). Significant co-occurrences/exclusions by environmental preferences and spatial factors were tested according to the framework proposed by Blois and collaborators (2014). One-way ANOVA was applied for testing significant (p -value < 0.05) environmental differences among sites where pairs of zOTUS were observed simultaneously based on salinity, temperature, pH, dissolved oxygen and redox conditions. Environmental variables were selected according to the previously reported effect over the microbial community assembly and activity in the same area (Casamayor et al., 2013; Menéndez-Serra et al., 2020, 2021; Triadó-Margarit et al., 2019). One-way MANOVA was applied to test for significant (p -value < 0.05) spatial segregation or aggregation of sampling sites where pairs of zOTUs were and were not observed, respectively. Associations not explained by environment or space were considered potential biological interactions (Figure E.1). Subnetworks were built following the same workflow after the whole salinity gradient was split in three salinity ranges, defined after a consensus between ecological significance and a balanced number of selected samples. Thus, the dataset was fractionated in low (0.1–2.4% salinity, 36 samples), intermediate (2.5–4.9% salinity, 35 samples) and high

salinity samples (5–40%, 44 samples). For the subnetworks, a minimum occurrence threshold was not established, while the abundance threshold was retained. The source code in R for networks analysis is available at <https://github.com/mateumenendez/networksConstruction>.

7.2.3 Network analysis and null network model simulation

The visualization platform Gephi (Bastian et al., 2009) was used for network representation and computation of basic network parameters such as modularity, average path length, network diameter, node degree and heterogeneity on degree distribution (Newman, 2003). The relative weight of each taxonomic group on the co-occurrence networks was defined as the proportion of the total weighted degree accumulated by the nodes corresponding to its members. This measure was compared with its representativeness in the dataset used for making the co-occurrence network in each case, which was defined as the proportion of the observations on the presence-absence data. Random graphs were computed by two complementary approaches in order to test the statistical significance of network descriptive parameters: (i) Erdős-Rényi (1960) random graph model and (ii) the randomization of the presence-absence observation matrix prior to networks construction, respectively. The Erdős-Rényi random graphs were computed based on the number of nodes and the probability of interaction between nodes of the observed networks (Connor et al., 2017), preserving therefore the average degree of the derived networks (Erdős and Rényi, 1960). Conversely, the randomization of the presence-absence matrices with the “curveball” algorithm (Strona et al., 2014) maintains species occurrences and samples richness, preserving biologic features as the distribution of generalists and specialists groups. For each network and subnetwork, 1000 random graphs were generated based on each null model approach and its descriptive parameters were calculated to create a parameter distribution to further calculate the statistical significance ($p\text{-value} < 0.05$) of the observed parameters. See Supplementary Material for a detailed description of random networks construction workflow.

7.3 Results

7.3.1 Pairwise co-occurrences and co-exclusions

Filtering based on minimum abundance and occurrence resulted in 1140 zOTUs (762 prokaryotic zOTUs and 378 eukaryal zOTUs) and 116 samples. Significant co-occurrences (57,795 cases detected) prevailed over co-exclusions (26,618). Salinity explained the highest number of the co-occurrences and co-exclusions based on niche coincidence or segregation, followed by temperature and redox conditions (Figure 7.1 a and e). In most cases, it was impossible to split the independent effect of the environment due to the high spatial autocorrelation found. The accumulated effect of all measured environmental variables explained 92.4% of the co-occurrences and 98.3% of the co-exclusions. The independent effect of space was limited (3.23% and 0.98% in co-occurrences and co-exclusions, respectively). Overall, the number of potential biological interactions was reduced to 2543 for co-occurrences (4.4%) and 181 for co-exclusions (0.68%). Subnetworks based on the three salinity ranges showed higher percentages of interaction assigned to potential biological interaction than full networks in all cases, with a decreasing trend along the salt gradient (Figure 7.1 b-d and f-h). Salt concentration remained as the most important parameter for the high and low salinity ranges, while pH was the environmental variable explaining most interactions at intermediate salinities.

7.3.2 Networks and subnetworks composition and topology

Network topology showed remarkable differences between positive (co-occurrence) and negative (co-exclusion) networks (Figure 7.2 a and 7.2 e). Co-occurrence network showed a higher number of nodes (550 against the 221 nodes of the negative network) and links (2543 against 181), being in all cases significantly lower than the observed for the randomized null networks. Nonetheless, mean degree showed a higher connectivity between nodes than the expected by chance for the positive network, being significantly lower for the negative counterpart (Figure E.4). Network modularity was higher than expected for both networks, while network diameter and average path length were lower for the co-exclusion network

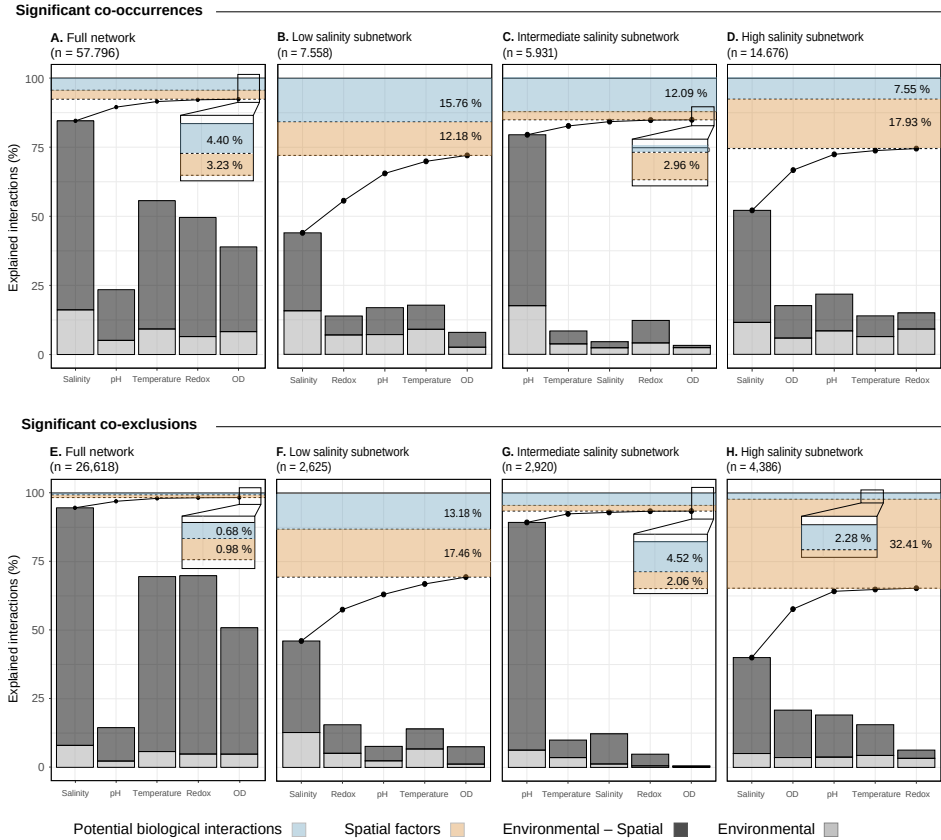


Figure 7.1: Pairwise co-occurrences and co-exclusions explanations according the framework proposed by Blois and collaborators. For each network and subnetwork, the potential explanation of all significant co-occurrences/exclusions by differences in environmental niche or samples spatial configuration was independently and simultaneously tested. Then, environmental variables were ordered, starting from the one which explained a higher number of associations, and followed by those which mostly contributed to the cumulative proportion of interactions not previously explained by other environmental variables (black line and points). The effect of samples spatial configuration was represented above the accumulated percentage of interactions explained by the environment. Interaction not explained by any environmental variable or the spatial configuration were assigned to potential biological interactions.

and showed inconsistent results for both null models and co-occurrences (Figure E.2). A few microbial groups were strongly overrepresented in the co-occurrence network (Figure E.2). Members of the Candidate Phyla Radiation Patescibacteria such as Parcubacteria, ABY1, and Gracilibacteria

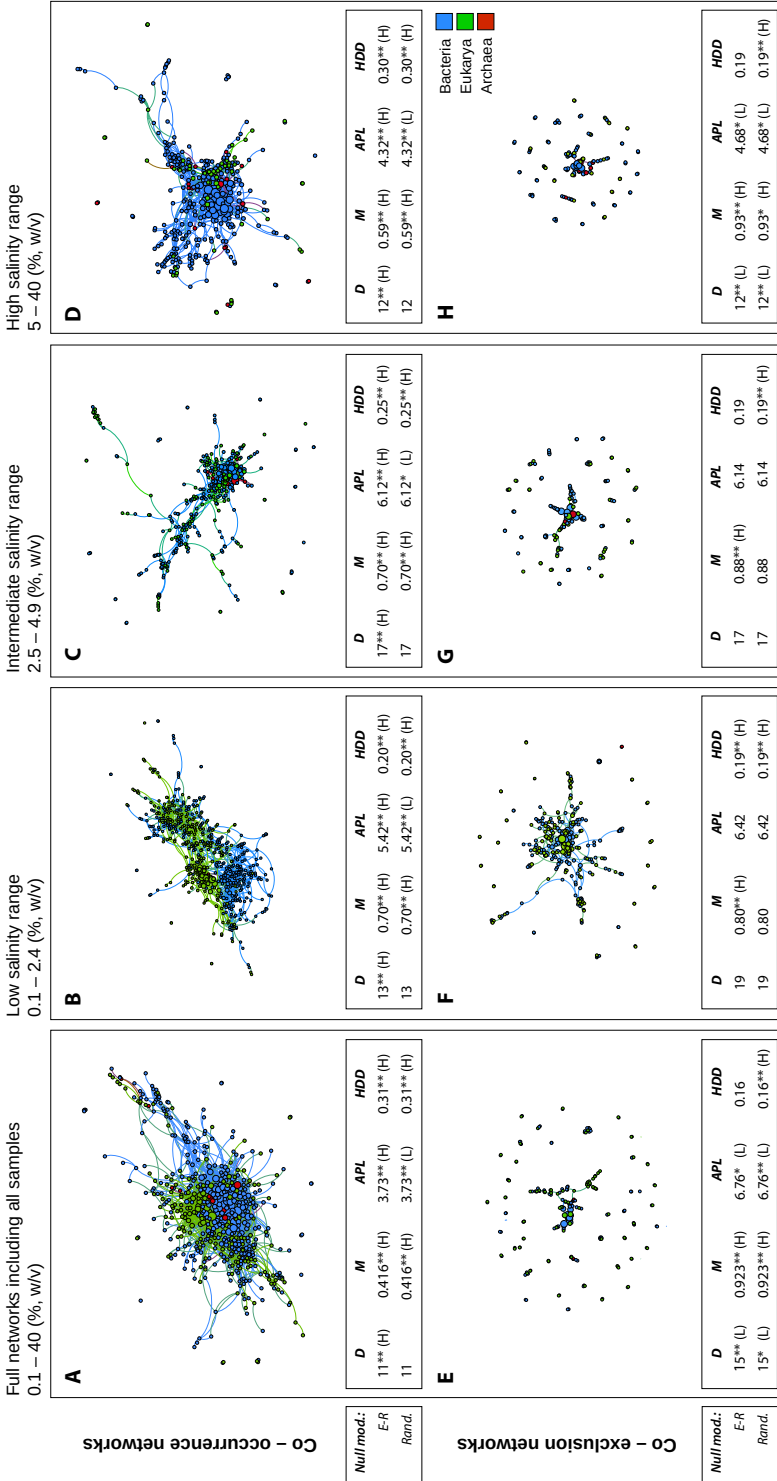


Figure 7.2: Networks graphical representation including positive and negative networks, and their salinity ranges subnetworks. Subnetworks were constructed independently for each salt range. For each network and subnetwork, descriptive parameters and its statistical significance according to both Erdős-Rényi random graphs and the randomized null models are given. D = diameter, M = modularity, APL = average path length, HDD = heterogeneity on degree distribution. Significance of descriptive parameters was established by determining a p-value of the observation compared to a parameter distribution from 1000 randomly constructed networks. L = significantly lower than expected on a random network. H = significantly higher than expected on a random network. Significance codes: * p-value < 0.05; ** p-value < 0.001.

showed the highest overrepresentation ratios vs their general presence in the dataset (up to 17:1). For the salinity ranges subnetworks (Figure 7.2 b-d and f-h) significant differences were observed. Co-exclusions networks appeared simplified as salinity increased, losing around 65% of its nodes and links while its mean degree dropped from higher to lower than the observed in the randomized null models (Figure E.4). The heterogeneity sized-down from higher than expected to random-like values when compared to the Erdős–Rényi model, being in all cases higher than the randomized null models values. Network diameter and average path length progressively shifted from random-like values to lower than expected by chance (Figure E.5). Conversely, co-occurrence networks maintained its structure and complexity along the gradient. The observed significant differences found in the full network topology as compared with the null model were retained in all subnetworks. Several of the specific microbial groups which were overrepresented in the full co-occurrence network were also overrepresented in the subnetworks, while others groups such as *Spirochaetes*, *Lenthisphaeria* or *Kathablepharidae* only showed a remarkable overrepresentation on specific salinity ranges (Figure E.3).

7.4 Discussion

The SGH has been very popular among plant ecologists though whether or not it can be extended to the microbial world is poorly known. Network analysis of the microbiome present along wide environmental stress-gradients can unveil valuable information on how stress affects potential biological interactions and community structure. Our results showed a decrease of co-exclusions along the increasing stress-gradient, suggesting a reduction of competition. Simultaneously, co-occurrences remained stable suggesting a potential lack of facilitation processes.

7.4.1 Inference of potential microbial interactions and ecological interpretation

Network analysis has gained importance as a tool for understanding (microbial) community assembly and functioning, complementing more classical

approaches based on diversity and compositional descriptions (Barberán et al., 2012a). However, current methodologies for estimating biological interactions among microbes at the community level suffer of several limitations. The robustness of correlation approaches over compositional data has been debated, and specific methodology has been developed with the aim of reducing potential artifacts (Friedman and Alm, 2012). In parallel, co-occurrences analyses have been also developed and applied (Araújo et al., 2011; Faust and Raes, 2012). Presence-absence based methods have been proven useful (Veech, 2013) and its application over strong gradients could add more robust information than correlations. Besides the potential methodological limitations to explore organisms distributions and abundances, any statistical method by definition has limitations if lacks a framework for reducing the potential effect of environmental and spatial factors. Here, we found that environmental and spatial variability explained a high proportion of the observed co-occurrences and co-exclusions (85% and 95%, respectively). These results highlight that, even for subnetworks covering lower environmental variations, the network construction workflow strongly benefits from the application of the framework proposed by Blois and collaborators (2014).

However, to establish the potential effect of the reported trends over the community structure and stability is not straightforward. Although an increase on the positive to negative interactions ratio might imply a decrease of community stability (Hernandez et al., 2021), the maintenance of interspecies dependencies by positive-feedback loops in the co-occurrence networks suggests that its potential effect over the community stability would remain stable along the gradient. Moreover, the elevated modularity observed in both co-occurrences and co-exclusions networks would reduce the effect of potential perturbations on specific groups over the whole community.

7.4.2 The stress gradient hypothesis in the microbial world

The stress gradient hypothesis has fueled numerous studies in the last decades as a conceptual framework to better understand biological interactions, mostly in plant communities (He et al. 2013, and references therein) and, more recently, also in microbes (Hernandez et al., 2021; Lawrence

and Barraclough, 2016; Piccardi et al., 2019). Nonetheless, there is still a lack of studies assessing this question in microbial communities inhabiting natural environments where variations in stress levels occur within a limited area and can be directly quantified. In this context, saline ephemeral lakes represent a perfect scenario for assessing the role of increasing stress since salinity fluctuations occur in-site and within short times periods and distances. Incorporating the stress nature and organism’s life history into the SGH framework (Maestre et al., 2009, 2005) improves the prediction of potential outcomes and helps interpretation. Thus, the nature of the stress factor and the specific biology of the organisms cannot be considered independently from each other. For instance, an increase of facilitation will be only observed when, for a particular stressor, at least one specie of the community has the ability to ameliorate the stress (Piccardi et al., 2019). However, the mechanistic explanation on how one aquatic species could benefit from another under saline stress is unclear. Potentially, under salt stress a microbial facilitation mechanism could be related to the production and excretion of compatible solutes which are used by halophilic and halotolerant organisms to maintain intracellular osmotic pressure under saline conditions (Roberts, 2005). These molecules can be either synthesized by the organism after a high energy expenditure or, alternatively, be directly incorporated from the environment (Kempf and Bremer, 1998), which substantially reduces the high bioenergetic requirements for maintaining the osmoadaptative cell machinery. However, although halophile groups relying on this mechanism dominate the assembly at intermediate salinities (Menéndez-Serra et al., 2021) and take part of the retrieved co-occurrence networks (Figure E.2 and E.3), this strategy has not been fully described as a stable facilitation mechanism and deserves further research to establish its potential scope.

Another potential facilitation mechanism relies on intracellular parasitism, where the parasite escapes from the cellular stress exerted by salinity by developing its vital cycle within the host intracellular medium. Interestingly, specific microbes potentially holding this ecological strategy showed a remarkable role in the observed co-occurrence network. Mostly unknown microbes of the recently proposed Candidate Phyla Radiation

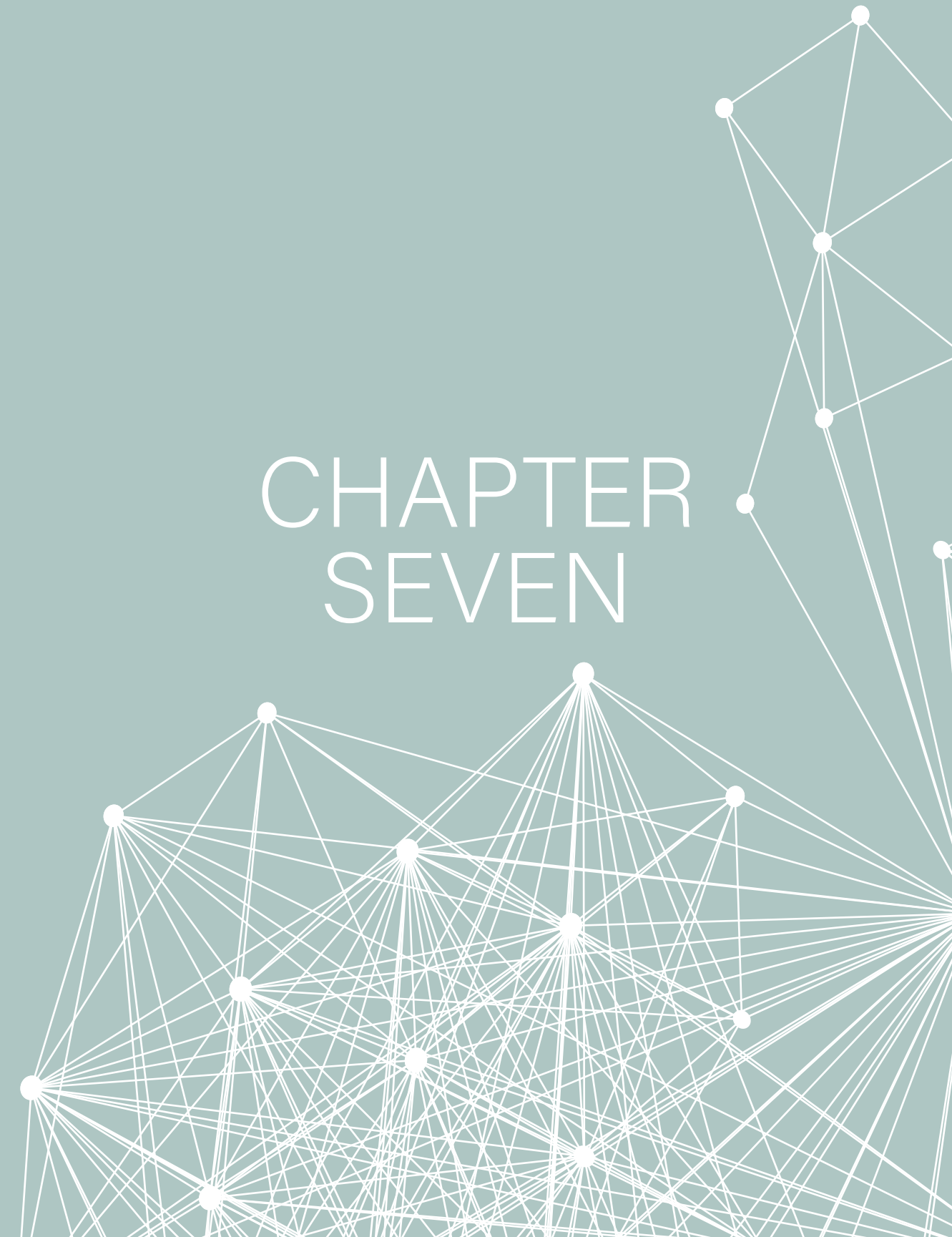
(CPR), e.g. the phylum *Dependentiae* (Yeoh et al., 2016), and DPANN (Castelle and Banfield, 2018) members accounted for a higher proportion of interactions than expected according to their mostly modest presence in the original Monegros dataset. CPR and DPANN have been hypothesized to present a symbiotic/parasitic lifestyle, according to both the lack of essential metabolic functions in their simple genomes (Castelle and Banfield, 2018; Yeoh et al., 2016) and the direct microscopic observations in agreement with intracellular parasitism (Gong et al., 2014). Although a direct link between co-occurrence analyses and ecology is extremely controversial, it could be still considered as a useful complementary approach. The biological nature of these potential associations, together with its perseverance along the salinity gradient, suggests a behavior unaffected by the stress factor. Conversely, co-exclusions experienced a remarkable effect along the gradient, suggesting a mass shift from negative interactions to neutrality as salinity increased. An equivalent transition has been already described under increasing stress in a simplified microbial consortia (Piccardi et al., 2019) and is most likely linked to a progressive and generalized change of the dominant strategy within the communities, moving from competition to stress tolerance (Ho et al., 2017).

7.4.3 Concluding remarks

The application of indirect approaches to infer biological interactions has several methodological limitations. The main limitation is the impossibility to discern the true nature of the observed association, which limits an accurate results interpretation. Although recent studies have questioned the validity of indirect methods for inferring specific trophic associations between pairs of species (Freilich et al., 2018), when these methods are applied at larger scales consistent patterns may arise to unveil high order interactions. Endowing observed patterns with ecological significance is as controversial as necessary for moving from describing statistical trends to generate hypothesis on the ecological processes shaping communities. Our study suggests that, in agreement with the predictions of the SGH, competition decreased along the gradient potentially linked to a dominance of osmoadaptation mechanisms over growth in terms of resources allocation.

The potential lack of stress-specific facilitation reinforces the importance of the interaction between the stress factor and the specific biology of the organisms to accurately predict potential interactions outcomes.

CHAPTER SEVEN



Mateu Menéndez-Serra, Joan Cáliz, Xavier Triadó-Margarit, Emilio O. Casamayor. Increasing saline stress reduces allochthonous microbial colonization in ephemeral ponds, driving the transition from habitat generalists to specialists. *Manuscript in preparation*

The bibliography of this chapter can be found on the unified *Bibliography* section.

Increasing saline stress reduces allochthonous microbial colonization in ephemeral ponds, driving the transition from habitat generalists to specialists.

Abstract: Dispersal of organisms across space is known to be one of the key factors in determining communities composition. Dispersal is usually only considered among local communities within the same type of biome. However, the idiosyncrasy of microbial organisms allow them to successfully colonize and develop on several biomes (e.g. terrestrial, aquatic, aerial), making therefore possible its dispersal among biomes. The strong influence of terrestrial-derived colonizers into the aquatic microbiome has been previously reported on freshwater environments. Here, we tested the potential influence of terrestrial derived organisms into the aquatic microbiome by the joint analysis of the 16S rRNA dataset corresponding to the water samples and an analogous dataset corresponding to 19 sediment samples from five saline ponds within the Monegros lacustrine system. We reported a decreasing contribution of sediment derived organisms into the aquatic communities as water salinity increased. The divergence of both groups of samples reached a maximum plateau at salinities above 5‰. The source-tracking approach showed a transition from habitat generalists to specialists along the salinity gradient, both for organisms present on both biomes as for those only present on water samples. Lastly, our results showed the remarkable role of sediments within the studied aquatic system, acting as a “seed bank” during desiccation stages and driving interannual recolonizations by habitat generalist but also halophilic groups.

8.1 Introduction

Dispersal of organisms across space is, together with selection, a main factor determining community assembly within a metacommunity context (Leibold et al., 2004; Vellend, 2010). Its potential effect has been repeatedly reported for all types of organisms and environments, with different processes (e.g. dispersal limitation, mass effect, homogenizing dispersal) acting at different spatial scales (Mony et al. 2020 and references therein). Potential dispersal of organisms within a given metacommunity is usually only considered for the same type of biome (e.g. terrestrial biome, aquatic biome). This approach is conceptually valid for most macroorganisms that hold physiological specialization that limits their ability to colonize heterogeneous environments. However, the high physiological and metabolic plasticity in microorganisms (Allison and Martiny, 2008) allow them to survive and develop within a wide range of biomes. Nonetheless, little is known about the processes mediating dispersal of microorganisms among biomes, and the potential colonization of these allochthonous microorganisms in new sink environments. Understanding dispersal processes between biomes may help to elucidate the assembly of communities that is not mediated by local processes (environmental selection, biological interactions, local dispersal).

The importance of microbial dispersal among biomes has been already reported in aquatic environments. Indeed, soil-derived colonizers have a strong influence on diversity and composition of aquatic (microbial) communities in freshwater systems (Crump et al., 2012; Monard et al., 2016; Ortiz-Álvarez et al., 2020; Ruiz-González et al., 2015). In complex hydrographic networks, the reported findings suggest a massive inoculum of soil-derived microorganisms on segments of the network where both compartments (i.e., soil and water) are highly interconnected (small streams) followed by a progressive selection of the best adapted microorganisms (Ruiz-González et al., 2015). Thus, water communities progressively differ from soil communities along the hydrographic network, which has led to hypothesize that water communities in freshwater environments are mainly constituted by a selected subset of richer soil communities (Crump et al., 2012). The high contribution of soils to freshwater communities has been reported for bacteria and archaea, while the role of soil colonizers seems to be lower for microeukaryotes (Crump

et al., 2012). At the interface of the aquatic and terrestrial biomes, both specialist and generalist taxa have been reported according to their potential to cross the biome barrier and colonize multiple environments (Monard et al., 2016). Previous studies have also reported that habitat generalist and specialist are primarily assembled by different mechanisms as selection or dispersal, increasing metacommunities dynamics complexity (Pandit et al., 2009; Székely and Langenheder, 2014). Additionally, the potential role of mass effect processes is also crucial on the interface between aquatic and terrestrial biomes, retaining populations of less adapted microorganisms due to dispersal from source habitats and adding complexity to the evaluation of the potential role of each microorganism within the community.

To date, studies that report the exchange of microorganisms between biomes has mainly focused on unidirectional systems where the flux of colonization is easily traceable, occurring in most cases exclusively from soils to water masses (Crump et al., 2012; Monard et al., 2016; Ruiz-González et al., 2015). However, the directionality of the colonization is less clear on highly interconnected systems such as water and sediment. Although many studies have analyzed the microbial composition of sediments and water masses within the same system (Hongchen et al., 2006; Jiang et al., 2006; Sommers et al., 2019; Zhang et al., 2019), most studies focus on describing differences in communities composition between both compartments linked to its physicochemical differentiation (e.g. anoxic conditions, light and nutrients availability). The link between biogeochemical cycles driven by microbial communities inhabiting both compartments has been also well described, e.g. methane production by anaerobic archaea on the sediment which is consumed by aerobic methanotrophic bacteria on the water column (Holmer and Storkholm, 2001) or sulphate reduction in anoxic sediments coupled with sulphur compounds oxidation within the water column (Holmer and Storkholm, 2001). However, the exchange of microorganisms between both compartments, as well as the factors modulating this exchange and its potential effect over both communities, remains poorly explored.

In ephemeral ponds, the relationship between water and sediment is extremely complex. During desiccation, dry sediments act as a seed bank hosting microorganisms which will later recolonize the aquatic environment

after ponds re-inundation (Sisson et al., 2018). While ponds are inundated, the exchange of microorganisms between water and sediments is presumably bidirectional and mainly modulated by the environmental characteristics of both compartments. Furthermore, mass effect processes between water and sediment might increase assembly complexity. In this context, we aim to test the potential contribution of external sources, with special focus on sediments, to ephemeral saline ponds water masses assembly combining two complementary approaches: (1) the joint analysis of water and sediment communities of the same area; and (2) the use of a database source tracking approach based on gene sequences to determine potential environmental sources. Using this combination we tested (i) the direct contribution of sediments communities into water masses along the hydrological cycle and (ii) the role of habitat generalists and specialist taxa on the community assembly. We hypothesize that this contribution will be variable and strongly dependent on the differential selective pressure reported along the salinity gradient.

8.2 Materials and Methods

8.2.1 Water and sediments molecular datasets

The available 16S rRNA gene datasets corresponding to bacteria and archaea in water and sediment samples of the Monegros desert lacustrine system was selected for the present study. The sediments dataset contained 19 samples from five saline ponds in the Monegros desert (Menéndez-Serra et al. 2019, which can be found on the *Related works* section of this thesis). Sediment dry samples were sampled within the inner margin of the ponds during a desiccation stage, being therefore influenced by water masses during ponds inundation periods. As previously reported, while water samples composition was strongly linked to changes in salinity (Menéndez-Serra et al., 2020, 2021), sediment samples did not show a significant response to any of the tested environmental variables (Menéndez-Serra et al., 2019). The use of the same primers set (515F – 806R) and sequencing platform (Illumina 2x250 MiSeq) allowed us a joint analysis of both molecular datasets. For this, raw sequences of water and sediment samples were processed together

following the UPARSE pipeline (Edgar and Flyvbjerg, 2015). Sequences were merged and quality filtered (expected error of 0.25). After de-noising and chimera filtering, the UNOISE algorithm (Edgar, 2013) defined Operational Taxonomic Units at 100% identity, i.e., zero-radius OTUs (zOTUs). The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012) and the SILVA 132 reference database (Quast et al., 2013). zOTUs were filtered by their alignment quality score ($>80\%$), and those classified as mitochondria and chloroplasts were removed. Only samples with $> 10,000$ reads were retained, and the resulting table consisted of 13,677 zOTUs distributed in 155 samples. Both gene sequence datasets have been separately deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605 (water samples) and PRJNA420947 (sediment samples).

8.2.2 Environmental Ontology sequences annotation

We use seqenv (Sinclair et al., 2016) to link gene sequences to potential sources based on the reported habitat for similar sequences on public repositories. For a given zOTU sequence, seqenv performs a BLAST search against the NCBI nucleotide database to detect sequences with high identity and applies a text-mining protocol to extract environmental ontology terms from the data associated to those sequences. Environmental ontology (EnvO) terms are based on a concise, structured and controlled vocabulary which helps to define environmental entities of all kinds (Buttigieg et al., 2013). The seqenv pipeline (version 1.3.0) was run over the predicted zOTUs sequences with a minimum identity threshold of 99%. As each zOTU might be linked to several EnvO terms, the relative contribution of each term was determined by the proportion of matched sequences reported from each source. The richness and diversity of sources reported in databases provide information on the different environmental conditions where a given zOTU was previously detected, and can be considered a proxy for habitat specificity or generalism at the global scale.

8.2.3 Specificity index

We made use of the specificity index applied by Mariadassou and collaborators (2015) to determine the specificity of each zOTU to each salinity range, considering only water samples. This index is, at the same time, based on the IndVal index (Dufrene and Legendre, 1997) but keep the specificity value of each zOTU within all groups of samples, not only its maximum. Briefly, the index is the product of two independent parameters defined by Dufrene and Legendre as specificity and fidelity to the group. For a single zOTU and group of samples, specificity is the ratio between mean abundances within the group and the sum of mean abundances in all groups, and fidelity is the proportion of samples within the group where the zOTU has been detected. The specificity index was then represented against the mean abundance of each zOTU within each salt range in order to assess the potential link between specificity and dominance (Mariadassou et al., 2015). The significance of the observed trends was tested by permutations in order to discard potential artifacts. For this, the same index was calculated for each zOTU and salt range after randomly assigning each sample to a salinity range while maintaining the number of samples of each group. This process was repeated 999 times in order to compute a mean value and a confidence interval for each zOTU and range, to be represented together with the observed data. For approximating potential differences between trends, both the observed data and the mean values of the permutation test were fitted to a linear model and the models slopes were statistically compared (function *lstrends* from *lsmeans* package).

8.2.4 Additional statistical analysis

Statistical analyses were run in R environment (<http://www.r-project.org/>). Environmental sources richness and the Shannon–Weaver index were calculated using “vegan” package version 2.4-5 (Oksanen et al., 2016). Spearman’s Rank order correlation coefficient was carried out to test monotonic associations between numeric variables. Differences among group means were tested using Tukey’s range test, function *HSD.test* within “agricolae” package version 1.2-8, after significant ANOVA. Differences in the microbial

composition between water and sediment samples were assessed by Non-Metric Multidimensional scaling (NMDS) ordinations and permutational multivariate analysis of variance (PERMANOVA) using distance matrices based on Bray-Curtis dissimilarity index.

8.3 Results

8.3.1 Water–sediment communities relationship

Sediment and waters samples showed significant differences in terms of community composition based on the Bray-Curtis dissimilarity index (ANOSIM statistic R: 0.90, significance: 0.001) (Figure 8.1). Distribution of Bray-Curtis dissimilarity indexes between water-sediment pairs of samples did not show significant differences between samples from the same or a different pond (Figure F.1, right panel). This lack of water-sediment pond specificity allowed us to analyze water-sediment pairs of samples with independence of its pond, providing analysis of higher statistical robustness when necessary. A significant and positive correlation (Spearman’s rho = 0.46, p-value < 0.001) was found for intercompartment (water-sediment) dissimilarity along

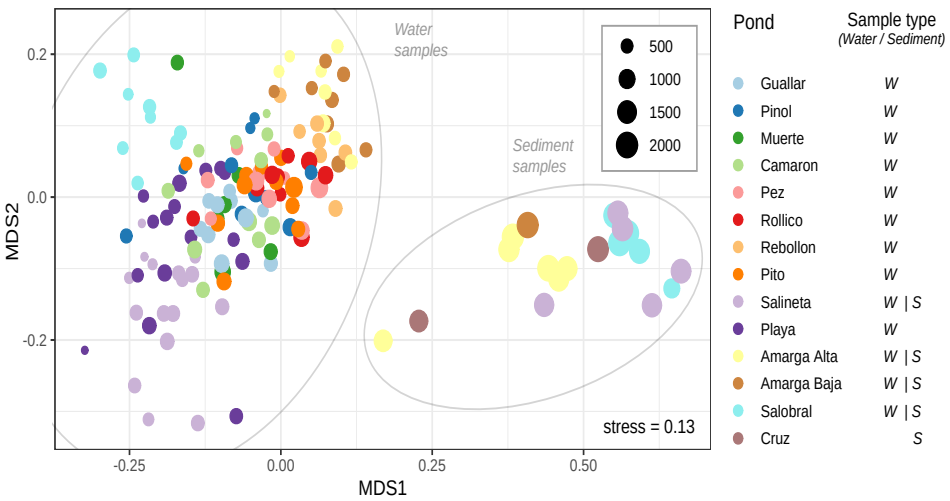


Figure 8.1: Non-metric multi-dimensional scaling (NMDS) representation of water and sediment samples based on Bray-Curtis dissimilarity index.

the salinity gradient (Figure F.1, left panel), i.e. sediment and waters samples progressively differed in terms of community composition as water salinity increased. No significant trend was found when the same indexes were represented against sediments conductivity, supporting the lack of consistent trends along the conductivity gradient in sediment samples.

Of the 13,671 zOTUs distributed among all samples, 3,060 (22.4%) were common to water and sediment, while 4,209 (30.8%) and 6,402 (46.8%) were detected exclusively in sediment and water, respectively. The weight of water-exclusive zOTUs within each water sample significantly increased along the salinity gradient, both in terms of zOTU proportion (Spearman's $\rho = 0.78$, $p\text{-value} < 0.001$) and relative abundance (Spearman's $\rho = 0.60$, $p\text{-value} < 0.001$). Most of the variation was recorded at low salinities ($< 5\%$ w/v) where zOTU proportion and relative abundance increased from 50% to 75-80%, and reached stabilization (Figure 8.2). In sediments, the proportion (mean = 0.65, sd = 0.13) and relative abundance (mean = 0.67, sd = 0.16) of soil-specific zOTUs remained stable along the electrical conductivity (salinity) gradient.

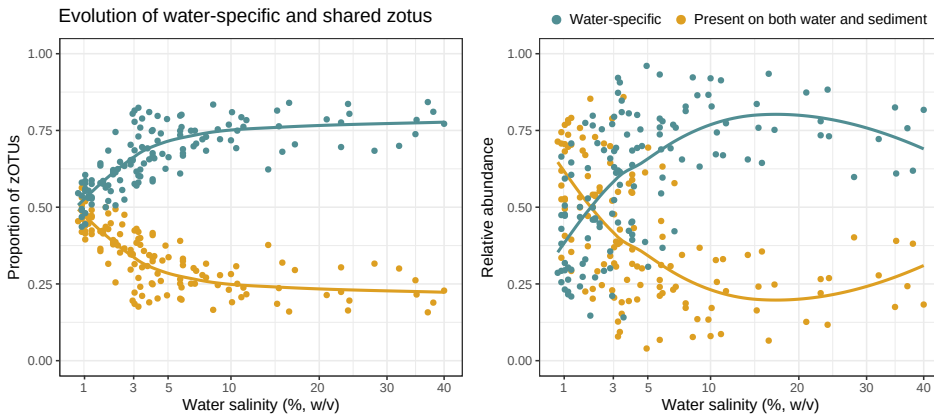


Figure 8.2: Proportion (left panel) and relative abundance (right panel) of zOTUs found only in water samples and zOTUs found both in water and in sediment samples, along the salinity gradient.

8.3.2 Predicted sources for pond microorganisms

The relative contribution of the most abundant (mean relative contribution $>5\%$) predicted environmental sources for microorganisms in the water samples changed along the salinity gradient (Figure SF.2, left). As expected, the abundance of microorganisms with EnvO source “sea water” was higher ($\sim 10\%$) in ponds with marine-like salinities (close to 3%). Conversely, microorganisms related to “saline lake” and “saline evaporation pond” sources remained low at salinities $<5\%$, but significantly increased along the gradient reaching relative abundances $>15\%$. The other abundant sources such as “lake” (10-20%), or “soil” and “sediment” (5-10%), remained relatively constant along the salinity gradient.

The number of sources (richness) and their abundance and evenness (Shannon index) of all predicted environmental sources was calculated for each zOTU. The diversity of sources for zOTUs shared in water and sediment samples was generally higher (Figure 8.3, left), as expected, in

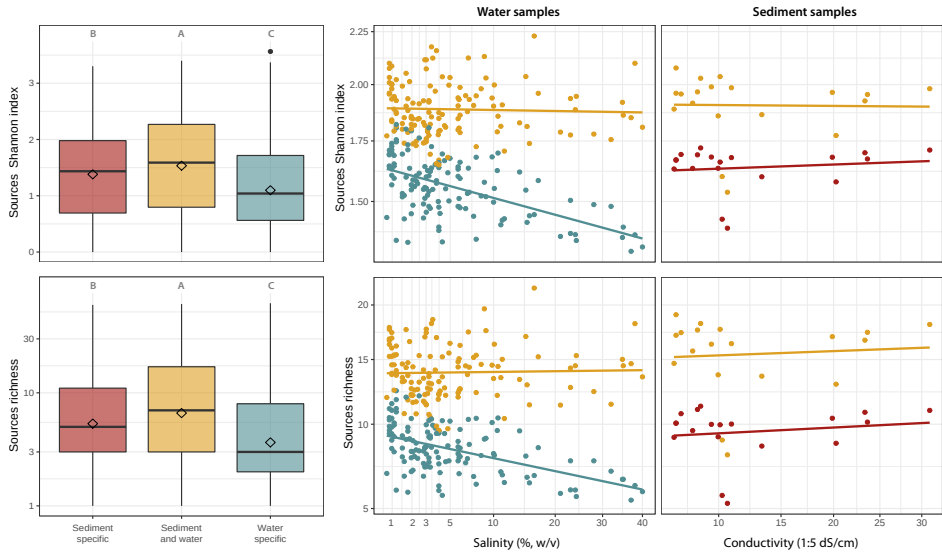


Figure 8.3: Boxplot of the diversity (richness and Shannon index) of sources (EnvO terms) for each zOTU, according to their assigned environmental compartment (water, sediment or both) (left panel). Mean diversity of the sources for each sample and zOTU group, represented along the salinity (water samples) and conductivity (sediment samples) gradients (center and right panels).

comparison to the compartment-specific zOTUs (Significant HSD Tukey test after significant ANOVA: F-statistic = 162.2 and 142.7 for richness and Shannon index, respectively; p-value < 0.0001). Moreover, the diversity of sources in each sample, taken only the zOTUs defined in each sample group (water- and sediment-specific, or present in both compartments), decrease along the salinity gradient for plankton (Spearman's rho -0.51 and -0.46 for Shannon index and richness respectively, p-value < 0.001) (Figure 8.3, center and right). The diversity of sources for zOTUs in sediments, or in shared compartments, showed no significant trends along the conductivity/salinity gradients.

For zOTUs found on both biomes, we calculated the ratio between its mean relative abundance within sediment and water samples. The distribution of these ratios presented a bimodal distribution (Figure 8.4, left). A main peak centered at zero evidenced a high proportion of zOTUs detected at similar abundances on both compartments, while a second peak centered at two evidenced a remarkable proportion of zOTUs with relative abundances two order of magnitude higher on sediments compared to water. Accordingly, the predicted primary source (i.e. the source accounting for the higher proportion for each zOTU) (Figure 8.4, right), showed that zOTUs overrepresented on sediment samples were mainly related to “soil”

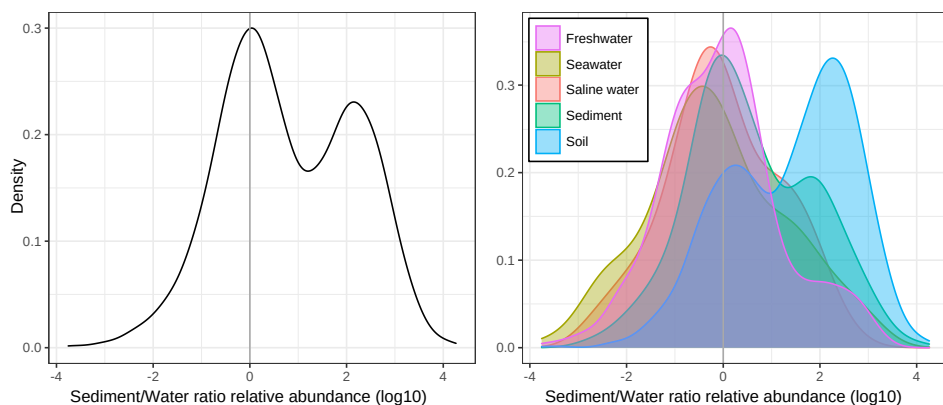


Figure 8.4: Distribution of sediment to water relative abundance ratio for zOTUs found in both biomes (left panel), and the same distribution with information linked to its primary predicted source (i.e. source accounting for the higher proportion for each zOTU) (right panel).

and “sediment” sources, with low contribution of the aquatic sources. Conversely, those presenting similar abundances on both biomes presented a high variability of primary sources, including both aquatic and terrestrial sources.

8.3.3 Local and global specificity and generalism

The specificity index calculated for each zOTU and salinity range showed a positive correlation with zOTUs local abundance in all five salinity ranges (Figure 8.5, top panels). This result indicates that, the more abundant a zOTU was on a given salinity range, the more specific was to that range. Although the specificity index calculated after the permutation tests (red dots and trend) also showed a positive correlation to local abundance, this was significantly much lower than the obtained for the observed communities. Slopes comparison after the adjustment to linear models showed significant differences in all ranges, with maximum differences between observations and permutation tests at both ends of the gradient. These differences suggests higher dominance of local specialists at the lowest and at the highest salinities.

Next, we plotted the specificity index against the richness and diversity of predicted sources (Shannon index) (Figure 8.5, bottom and middle panels, respectively). At the lowest salinity range (slightly brackish), local specificity showed a positive and significant correlation with the diversity of sources, i.e. low-salinity specialist were potentially global generalists. Conversely, local specificity of microorganisms at the higher end of the salinity gradient (highly hypersaline) showed a negative correlation, and these local specialist could be potential specialist at a global scale. No significant correlations were found for the other salt ranges. Taken together, a consisting decrease of global generalist was observed along the gradient.

In order to support these findings, we explore the genomic traits information reported in Chapter 5 of the present thesis. As reported by Barberán and collaborators (Barberán and Casamayor, 2014) and supported by recent studies (Chen et al., 2021), generalist taxa with a broader geographic and environmental distributions show larger genomes and higher

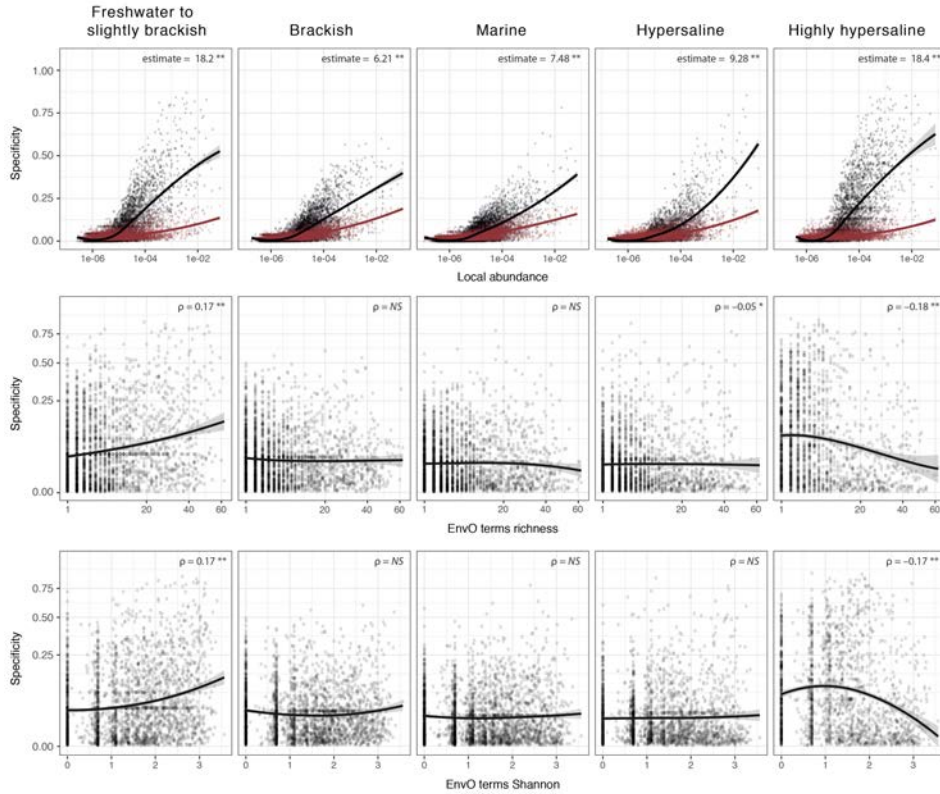


Figure 8.5: Positive relationship between specificity index and local abundance for each zOTU and salinity range. Red dots and trends represent the expected relationship under the null distribution when samples are randomly assigned to each salinity range. Differences between observations and permutations test were approximated by slopes comparisons after fitting to a linear model (top panel). Specificity index and salinity range represented against the richness and diversity (richness in the middle, and Shannon index in the bottom panel) of sources (EnvO terms) for each zOTUs. Significance codes: * p-value < 0.05; ** p-value < 0.001.

metabolic repertory to better adapt to multiple habitats. Based on this premise, we studied how genome sizes and gene counts changed along the salinity gradient. Genomic traits were inferred for a representative subset of the community by sequence matching against the IMG genomic database (see *Materials and Methods* and *Supporting information* of Chapter 5 for a detailed description). As expected, a significant and negative correlation was observed for both traits mean values and sample salinity (Spearman's correlation -0.53 and -0.52 for genome size and gene count, respectively,

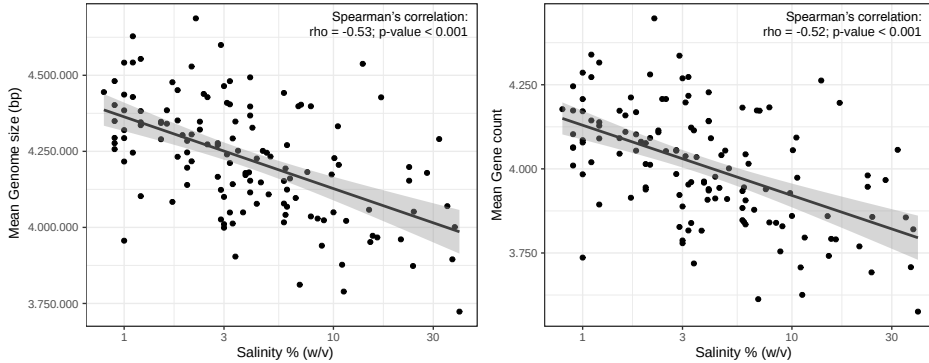


Figure 8.6: Mean genome size and gene count for each sample along the salinity gradient. Genome size and gene count were inferred for a representative subset of the aquatic community by sequence matching against the IMG genomic database.

p-value < 0.001) (Figure 8.6) supporting a significant decrease of genome sizes and gene counts along the salinity gradient.

8.4 Discussion

Dispersal is a key factor determining communities assembly for many organisms, including microbes. However, dispersal processes are highly controlled by environmental constraints which limits the success of potential colonization events. Environmental control is determinant for dispersal exchanges between communities which differ in environmental features, such as different biomes but also along environmental gradients. Experimental and computational approaches using broader perspectives incorporating multiple biomes as source of potential colonization may help to better understand the role of dispersal for microbial metacommunities assembly. This integrative approach provided better understanding for the microbial communities assembly in freshwater environments (Crump et al., 2012; Ruiz-González et al., 2015). However, little is known about inter-biome exchanges and its influence on the assembly on more complex systems such as ephemeral saline ponds, where the combination of periodical recolonizations and highly selective pressures add higher complexity to these systems.

We have shown in our study that at low salinities, water and sediment communities were highly interconnected. At the less saline end, c.a. 50% of

the bacterioplankton components were also detected in sediment samples. This was a similar proportion to the ones reported in highly interconnected segments of freshwater hydrographic networks (Crump et al., 2012; Ruiz-González et al., 2015) suggesting that those previously reported assembly mechanisms based on a massive inoculum and a progressive selection, might be also the origin of the observed Monegros local communities at low salinities. Therefore, water communities would be mainly constituted by a selected subset of sediment communities. However, unlike in unidirectional systems, sediment communities might by at the same time influenced by present or past water communities. With this, sediments would not only act as a source of sediment specific –or generalists– organisms but as a seed bank of water specific organisms which will recolonize the aquatic biome after ponds re-inundation, as previously reported for vernal pools (Sisson et al., 2018).

Although water and sediment communities shared a high proportion of taxa we also observed that, even under low selective conditions, not all organisms had the capability to cross the environmental barrier and establish on both compartments, as previously reported (Monard et al., 2016). Furthermore, our results also suggest that this ability is intrinsic to the microorganisms and it can be traced back in terms of habitat generalism or specialism behaviour at a global scale following the information provided by public gene databases. Thus, we observed that microorganisms found in both water and sediment showed a high richness and diversity of habitats where they had been previously detected compared to those only found in one compartment. Differences in habitat generalism or specialism were also detected on zOTUs only retrieved from water samples. In this case, zOTUs dominating the assembly at low salinities showed higher diversity and richness of associated habitats than those present at high salinities. This habitat specialization along the gradient is most likely given due the increasing specific requirements for life performance at high salinities (i.e. to cope salt stress), which progressively limits organisms competitiveness outside its specific habitat breadths. Conversely, groups dominating the assembly at the lowest salinities presented features of habitat generalists according to the source-tracking approach. These groups showed a high specificity

for the lowest salinity range and they were not able to successfully develop under increasing saline conditions.

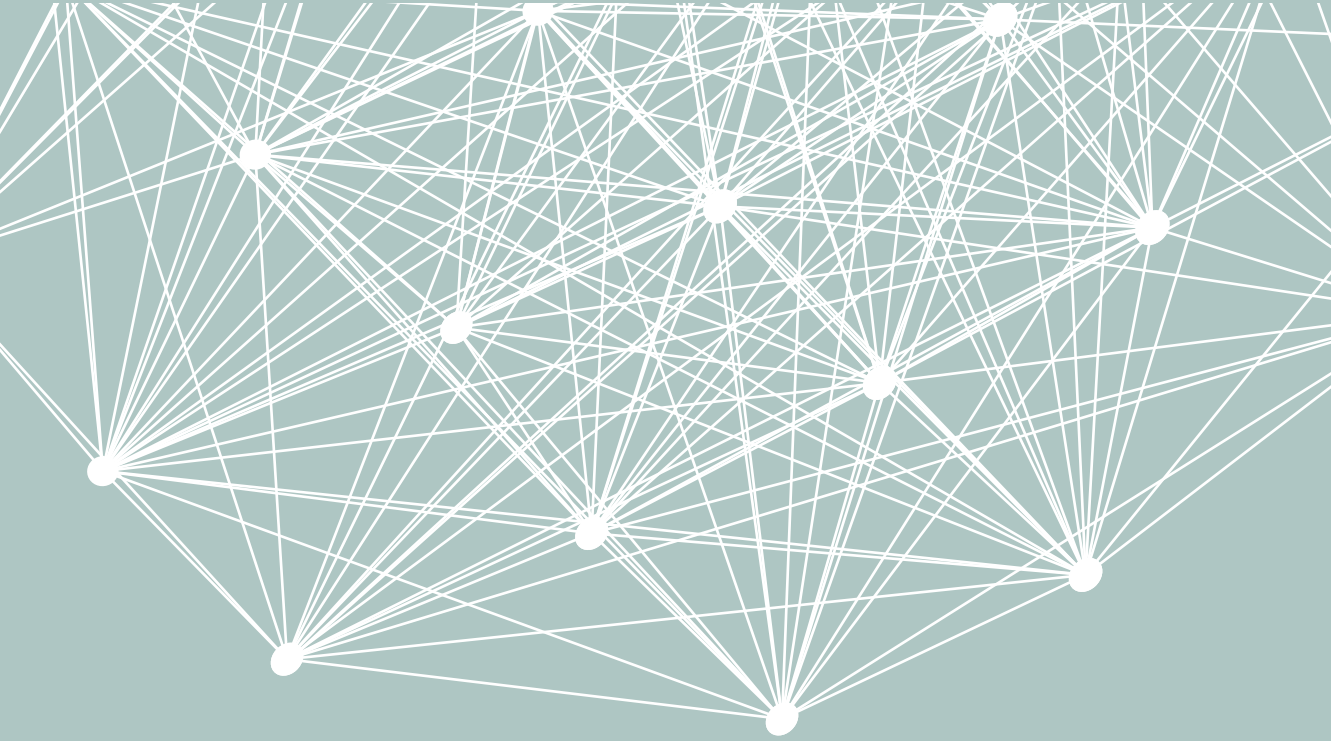
Mass effect may play an important role on shaping communities diversity and composition under low selective pressures. Here, microbial groups without the required traits needed to successfully develop in both biomes can be maintained due the combination of high immigration rates and low osmotic pressure. Although determining the role of specific groups on the assembly is extremely complex, approximations based on shared relative abundances in both compartments (Ruiz-González et al., 2015) and primarily sources (Ortiz-Álvarez et al., 2020) may provide insights into a differential role. Based on this, we identified a group of zOTUs which were overrepresented on sediment samples and primarily associated to soils and sediment sources, which led us to hypothesize that its detection on water samples might be the result of mass effect processes. Conversely, any significant group of zOTUs was found to reproduce the opposite pattern (enriched on water compared sediments and associated to water sources), suggesting a lower effect of mass effect from water to sediment.

As salinity increased up to 10-15% of salts, a remarkable change in the microorganisms dominating the aquatic assembly was observed. Beyond this point, dominant zOTUs behaved as specialist at local and global scales, showing lower diversity and richness of reported sources. This transition is potentially linked to the previously described threshold which separates moderate and extreme halophilic, historically set at 15% of salts (Imhoff, 2001; Kushner, 1978, 1985). Above this salinity threshold, salt-in groups including halophilic archaea and bacteria increased its abundance and become a relevant proportion in the assembly (Menéndez-Serra et al., 2020, 2021). The structural osmoadaptation in these groups limits its adaptation to lower salinity conditions (Oren, 2002), showing limited habitat breadths at any studied scale. The high specialization of these organisms challenges the hypothesis, potentially valid for freshwater environments, that water microbial communities are mainly constituted by a selected subset of soils or sediments assemblies. Then, where did these highly specified organisms come from? After environmental fluctuations, changes in community compositions might be the result of two different processes: recruitment over time and

over space. On the one hand, recruitment over time (Caporaso et al., 2012; Comte et al., 2017, 2014; Johanna et al., 2012) is produced when rare or inactive taxa already present in the community increases their abundance due to a fitness increase following the environmental perturbation. In this scenario, the configuration of the local seed bank is driven by past dispersal events. On the other hand, recruitment over space (Comte et al., 2017; Crump et al., 2012; Ruiz-González et al., 2015) is produced by the selection of specific groups, better adapted to the novel environmental conditions, among the pool of potential colonizers from external sources. Both processes simultaneously act following salinity increases in ponds systems, with sediment and air deposition as proven sources of potential colonizers (Comte et al., 2017, 2014). Our results suggest that the two processes may act simultaneously in Monegros salt lakes. The detection of extreme halophiles present along the whole salinity gradient support the recruitment over time of these groups already present on the water column. In this scenario, local seed bank within the water column is configured by past dispersal events. Additionally, taxa recruitment over space can be still produced at high salinities by the environmental selection from the pool of potential colonizers from sediments. Finally, previously described spatial patterns on the area (Chapter 6) as well as previously reported long-range dispersal and colonization of marine representatives (Comte et al., 2017), supports air deposition as a complementary source of potential recruitment over space. Nonetheless, our hypothesis apparently showed a notable flaw: many extreme halophiles, although being present on water samples along the whole gradient, were not detected in the sediments which are supposed to act as seed bank. This might be explained, we believe, because of the extremely low abundance these groups may have within the sediment. In order to overcome this potential limitation, Ruiz-Gonzalez and collaborators (Ruiz-González et al., 2017) carried out a deeply-sequencing of samples, reaching c.a. 2.2-3 million reads per sample, to assess the presence of rare taxa in the seed bank. These authors showed a remarkable increasing detection of the potential “reactive” pool, i.e. those groups showing transitions from rare to abundant between communities, up to a sequencing depth of 500.000 reads per sample. In our work, our sediment dataset had mean raw reads

per sample of c.a. 51.000, clearly under the detection limit at which rare taxa recovery reach the plateau. This fact, we hypothesize, could at least partially explain the non-detection of highly specific halophilic groups in the sediment.

In summary, we have shown that aquatic microbial communities inhabiting ephemeral shallow saline ponds are highly influenced by sediments with variable strength along the salinity gradient. The selective pressure exerted by increasing salinity progressively reduces the role of sediment-derived taxa, as well as other allochthonous groups, in the planktonic assembly. However, sediments also played a major role in the microbial assembly at high salinities, acting as seed banks and driving interannual recolonizations by habitat generalist but also highly specialized halophilic groups. Overall, increasing salinity drive a transition from habitat generalist to specialist by the establishment of severe physiological and metabolic requirements. Our work also reinforces the need to broadening the definition of microbial metacommunities by considering the interaction among multiples biomes in order to improve the understanding of the processes driving microbial community assembly.



GENERAL DISCUSSION

General Discussion

9.1 The hydrological cycle through the microbial metacommunity perspective

Ephemeral saline ponds of the Monegros desert lacustrine system show a high dynamism, which is one of the main differentiating attributes as compared to the traditionally better studied solar salterns systems (Fernández et al., 2014; Gasol et al., 2004; Ventosa et al., 2014b). Because of the endorheic origin and the climatic characteristic of the Monegros area (Herrero and Snyder, 1997), saline ponds temporally show severe variations in water volume along the rain-feeding hydrological cycle which are translated into wide salinity fluctuations. Increasing saline stress establish a series of metabolic, physiological and functional requirements which determine the abundance and identity of life-forms along the gradient. The stress gradient, furthermore, modulates the effect of other factors which potentially affect communities assembly as dispersal or biological interactions, determining complex equilibriums along the hydrological cycle.

Right after ponds inundation during the early rains of the wet season, water masses hold low salinities which a priori do not represent insurmountable stress conditions for most microbial organisms, and a wide variety of microorganism colonize and develop there, resulting in rich and diverse bacterial and eukaryal assemblies. Under the initial low selective pressure, habitat generalists dominate the assembly. These groups usually present high growth rates (Ramin and Allison, 2019) and elevated metabolic diversity (Barberán and Casamayor, 2014) which allow them to efficiently exploit resources under non-stressful conditions. These generalist taxa may proceed from multiple environments and biomes, not being exclusively water specific microorganisms. Soils and sediment communities contribute largely to the aquatic prokaryotic assembly (Ruiz-González et al., 2015) and the

low salinity ponds are partially constituted by a selected subset of soil and sediment communities. In the ephemeral ponds, sediments act as a “seed bank” during desiccation periods, hosting water specific organisms and saline and hypersaline specific groups in dormant states (Sisson et al., 2018) further driving aquatic recolonization. The combination of low stress levels and dominance of generalists groups, yields high interspecies competition which is detected as co-exclusions. Also, local dispersal of organisms between ponds operates originating a pattern of distance-decay of similarity among ponds as a potential result of two complementary processes: dispersal limitation and homogenizing dispersal. Homogenizing dispersal between close ponds potentially plays a major role in the assembly at low salinities as the low restrictive conditions most likely permit higher colonization success.

As evaporation dominates over precipitation, water volume progressively decreases and salinity increases. Slight salinity increases within the low salt range promote remarkable changes in the community. At this point, massive extinction and colonization events are registered for both bacteria and eukarya linked to minor salinity fluctuations. The observed appearance and disappearance of local taxa are the result of relative abundances fluctuations above and under methodological detection limits and most probably should not be considered as new “colonizations” and real “extinctions”. For bacteria, colonizations along the gradient are produced at the same order of magnitude than local extinctions, maintaining species richness and diversity as salinity increases. However, the increasing selective pressure entails that novel colonizers are not habitat generalists but salinity adapted organism. As these organisms require of specific osmoadaptation mechanisms, which are usually evolutionarily conserved (Roeßler and Müller, 2001), phylogenetic diversity decreases while species richness is maintained. Conversely, eukaryal colonization is one order-of-magnitude lower than extinction, evidencing not a taxa replacement but a progressive filtering of less adapted organisms instead. The increasing physiological stress drives the observed transition from habitat generalist to saline specific groups, together with a progressive change in resources allocation from cellular growth to stress tolerance mechanisms. This generalized process operates both on the community composition and on life strategies (Ho et al., 2017) resulting in an overall

reduction of interspecies competition.

After salinity exceeds an environmental threshold located at c.a. 5% of salts, only organism with proper osmoadaptative mechanisms successfully develop. Additionally, specific genomic traits seems to become determinants for the successful development of bacterial groups and are therefore selected. As a consequence of these eco-physiological limitations, beyond this threshold environmental selection overcomes dispersal processes shaping community composition for both bacteria and eukarya. Furthermore, the high selective pressure linked to the saline stress minimizes colonization success of potential colonizers coming from less saline ponds or external biomes. This way, the potential effect of dispersal processes such as mass effect or homogenizing dispersal is reduced, as well as the contribution of allochthonous microbes. Beyond this point, although salinity levels might keep rising, this increasing stress is not linearly translated into changes in the biological assembly. This suggests that novel colonizers, which already present osmoadaptation mechanisms, have wider environmental niches and can therefore cope with a broader range of salt concentrations. Thereby, as the colonization of novel groups is strictly controlled by the strict requirements for coping with high osmotic stress, turnover rates decrease resulting on higher temporal stability for both bacterial and microeukaryal communities. Once salinity exceeds the 15%, extreme halophiles relying on the “salt-in” osmoadaptation mechanism exponentially increase its relative abundance due its competitive advantage over moderate halophiles. However, moderate halophile groups depending on the “salt-out” mechanism maintain an important role even at the most saline end, probably due to time-lags in environmental selection. Once the hydrological cycle comes to its end and ephemeral lakes completely dry out, sediments arise again as a key compartment due its role as seed bank, which will drive water masses recolonization following the new hydrological cycle.

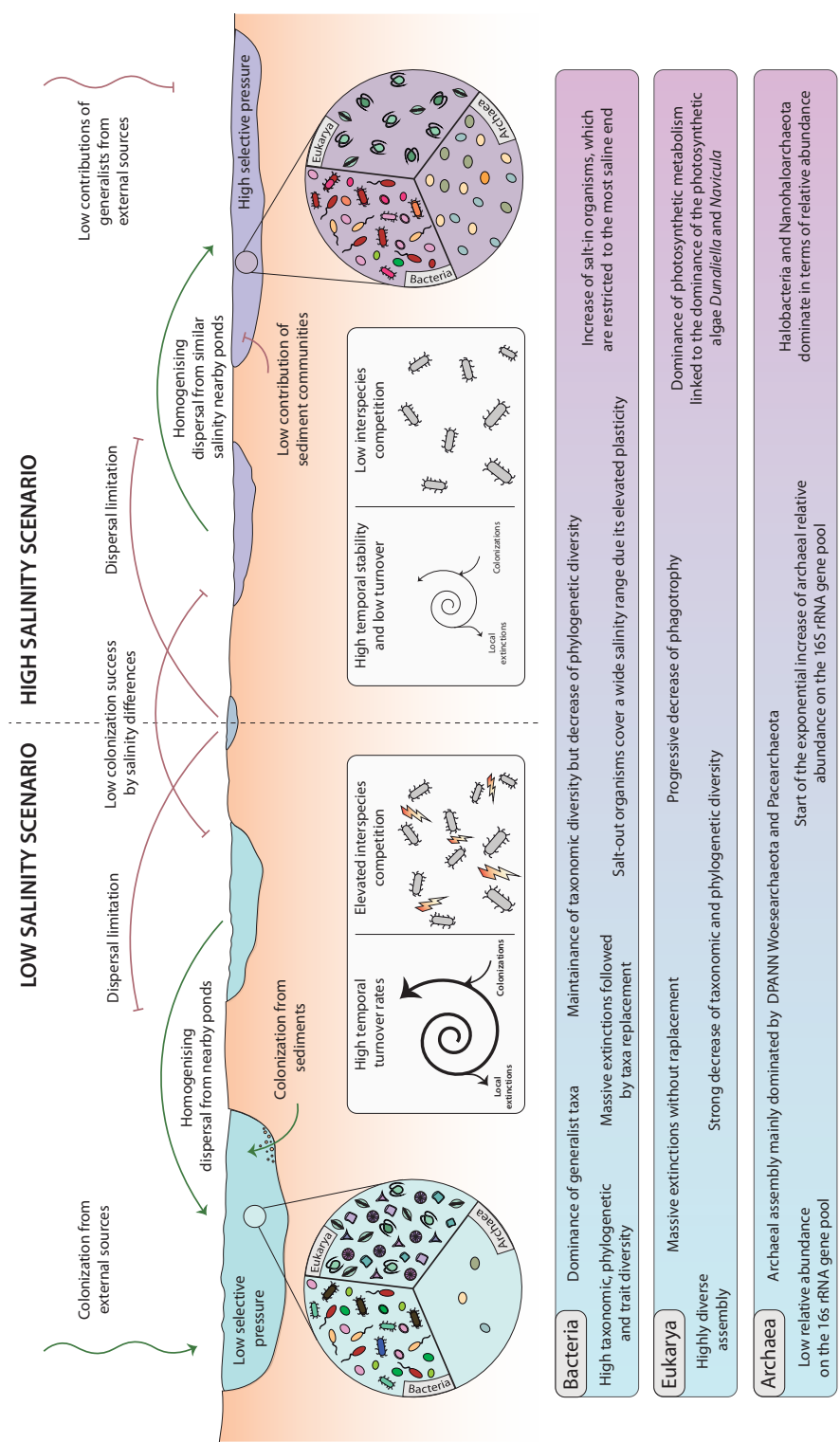


Figure 9.1: Simplified representation of the main processes driving community assembly under both low and high salinities scenarios, and the main milestones for the three microbial communities inhabiting the studied system.

9.2 Assembly processes and the physiological explanation of emerging ecological patterns

Understanding the mechanisms driving biological communities assembly is one of the fundamental questions in community ecology. The initial attempt to categorize the influence and dominance of the different assembly processes (Vellend, 2010) turned into a debate to determine the prevalence of deterministic or stochastic processes since the formulation of the neutral theory (Hubbell, 2001). As a counterpoint to this dichotomic perspective, Gravel and collaborators (2006) proposed that these scenarios are the extremes of a continuous landscape in which both processes interact, with variable strengths, to shape community assembly. Therefore, it seems reasonable that the focus should move from determining the prevalence of a single assembly process to understanding the different factors mediating the balance point between assembly forces, assuming that all processes simultaneously act (and interact). To reach this comprehensive understanding, the metacommunity framework (Leibold et al., 2004) arose as a useful tool to evaluate processes acting at variable scales and their potential interactions. Assessing the variable balance between assembly processes and how they interact to determine biological communities assembly is extremely complex as it requires of the simultaneous quantification/manipulation of several parameters as dispersal rates, selective pressure or community size. These difficulties increase when dealing with microbial communities due the associated observational and methodological limitations, and only a limited number of studies have evaluated the effect of manipulative variations in these processes on the assemblage, mostly on simplified microbial communities (Bell, 2010; Evans et al., 2017).

The study of dynamics behaviour at larger scales requires of natural systems reproducing measurable variability of the factors driving assembly processes. Detecting and quantifying dispersal magnitudes among different natural microbial communities is methodologically challenging. Instead, ephemeral saline lakes in the Monegros desert allowed us to detect and measure a selective pressure associated to the wide environmental gradient, which influenced the equilibrium among multiple assembly processes. In

this scenario, the magnitude of the potential dispersal among local communities remained constant while the variable selective pressure exerted by salinity modulated its net effect over communities assembly. Moreover, the nature of the environmental gradient allow us not only to easily link it to variable stress levels, but to understand the processes and mechanisms at the cellular level scaling up to high-order ecological patterns. Indeed, the observed patterns will most likely differ from those potentially emerging under increasing stress promoted by different factors as not all stressor promote an equivalent effect/response in the cells. The well-known effects of increasing osmotic stress at a cellular level, the metabolic and ecologic idiosyncrasy of the halophile and halotolerant groups, and the idiosyncratic osmoadaptation mechanisms to deal with the osmotic stress levels allowed us to accurately interpret some of the observed patterns. One of the most remarkable was the transition from stochastic to deterministic assembly processes. This transition, at 5% salinity, matches the previously described limit between moderate and hypersaline waters based on the presence of halophile groups (Edgerton and Brimblecombe, 1981; Imhoff, 2001; Kushner, 1978). Our findings support that, beyond this limit, the presence of specific osmoadaptation mechanisms to cope with the elevated osmotic pressure becomes essential and limits colonization success. These requirements limit the colonization of habitat generalists groups not only from less saline sources but also from external biomes as soils or sediments. The observed changes on community composition promoted by the increasing stress implies changes in organisms identities, community metabolic functioning and potential interspecies interactions. The well-described bioenergetic costs associated to maintain the osmoadaptive machinery (Oren, 1999), which forces a transition in resources allocation from growth to stress-tolerance mechanisms (Pedrós-Alió et al., 2000; Rath et al., 2019; Stevenson et al., 2015) allowed us to mechanistically explain the observed decrease in competition. Likewise, the lack of described osmotic stress amelioration mechanisms supported the reported lack of stress-promoted facilitation along the salinity gradient. Overall, our findings reinforce the importance of understanding processes operating at an individual level to successfully interpret the emerging high-order patterns. Additionally, our work emphasizes the suitability of highly

dynamic ephemeral saline systems as a perfect scenario for community ecology studies, both due its environmental idiosyncrasy and the extensive existing knowledge on the complex relationship of microorganisms with salinity.

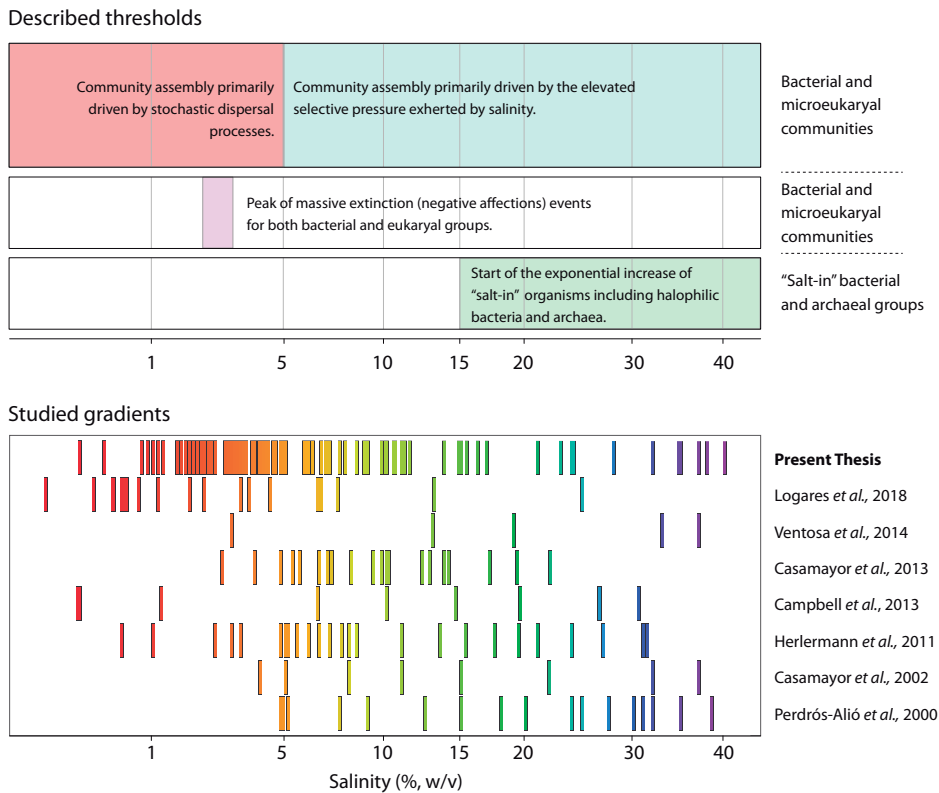


Figure 9.2: Described environmental thresholds along the salt gradient, and salinities covered in the present thesis compared some of the most comprehensive studies covering wide environmental gradient in the literature.

9.3 Multiscale community assembly

Considering communities assembly along multiple scales (e.g. temporal, spatial, phylogenetic) unveils patterns and processes which are not observed in more simplistic scenarios, and a more complex scenario where the interaction of the different processes have a variable effect in relation to the considered scale has to be considered for a more comprehensive and integrative mechanistic understanding of communities functioning.

9.3.1 Temporal scale

Although the importance of temporal dynamics in ecological communities has been considered since long (Magurran, 2008), ecological monitoring programs in general are only now starting to generate long-term datasets (Hortal et al., 2015). Furthermore, microbial temporal datasets are usually analyzed in terms of beta-diversity patterns linked to environmental fluctuations occurred over time (Eiler et al., 2012; Guo et al., 2018; Höfle et al., 1999; Shade et al., 2007), with little consideration of the temporal dimension itself. Recently, Ontiveros and collaborators (2019; 2021) developed a dynamic approach rooted on the Equilibrium Theory of Island Biogeography (ETIB) (MacArthur and Wilson, 1967) which, by exploring presence-absence patterns in temporal datasets, provided insights into the individual colonization and extinction dynamics, general dynamics of the community and turnover rates in relation to the environmental changes. This is especially relevant because, as stated by Linsdtröm and Langenheder (2012), reaching a mechanistic understanding about communities assembly relies not only on quantitatively describe the effect of environmental processes but also its interaction with other assembly mechanisms. This mechanistic approach unveiled the effect of increasing selective pressures over populations temporal dynamics and dispersal processes.

The temporal dimension also captured the role of the sediments in ephemeral ponds functioning. After the desiccation stage communities are essentially reassembled from the local inactive seed bank on the sediments (Sisson et al., 2018; Székely and Langenheder, 2017). This is possible due the capability of microbes to enter a dormant state, i.e. a reversible state of low

metabolic activity, which allow them to survive within sediments and have an impact over future communities assembly. After ponds re-inundation, organisms undergo an initial phase of recovery, followed by a secondary succession process which is driven by local environmental features and environmental selection (Fierer et al., 2010). This “continuous periodicity” allows to explain, for instance, the presence of extreme halophile at the beginning of the hydrological cycle when environmental characteristic are not favorable.

Explicitly considering the temporal dimension might also allow unveils processes that are not observable from a static point of view and that add complexity to the study of communities at a given time. In systems with high environmental variability, time-lags in environmental selection (Monger et al., 2015) might play a major role on determining the observed assembly at a given time. While these effects have been evaluated following an abrupt salinity decrease (Andersson et al., 2014), little is known about the potential effect that time-lags in environmental selection might have over a community exposed to a progressive and maintained salinity increase. The consideration of such processes might help to explain, for instance, the higher bacterial diversity found here at the most saline end compared more stable environments (Ventosa et al., 2014a).

9.3.2 Spatial scale

Dispersal of organisms across space is a key component for the assembly of biological communities. This is especially relevant for microbes that successfully disperse across long distances, exhibiting high survival rates due to the presence of resistance forms, and the capability of a single individual to successfully colonize and quickly develop (Brown, 1995). The description of biogeographical patterns from few meter (Lear et al., 2014) to a planetary scale (Finlay, 2002) not necessarily challenges the ubiquitous assumption (Baas Becking, 1934) but emphasizes the complex outcome that dispersal processes produced at different scales. Furthermore, dispersal processes among different biomes, e.g. between aquatic and terrestrial systems (Ruiz-González et al., 2015), increases assembly complexity.

The potential effect of dispersal processes operating at multiple spatial

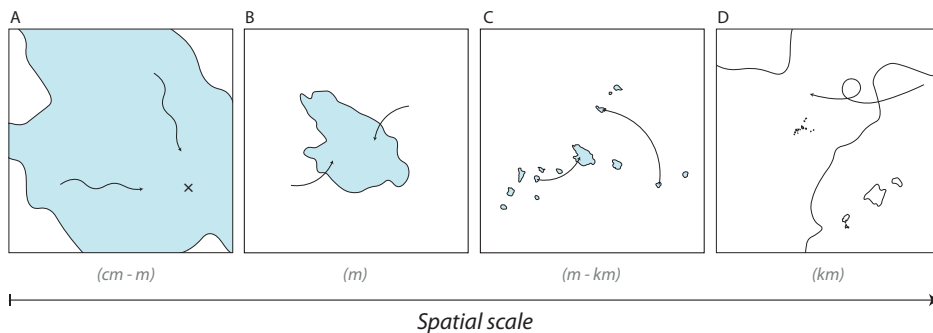


Figure 9.3: Spatial scales at which dispersal processes potentially occur: (a) within lakes, (b) from surrounding biomes, (c) among lakes and (d) long-range dispersal.

scales might have an unequal effect for the different populations. As proposed by Barberán and collaborators (2014), the idiosyncrasy of microorganisms can lead to a blurring between regional and global species pools, which would reduce the weight of processes acting at a regional scale in the assembly of local communities. However, this microbial specific framework still consider the importance of the regional pool for microbial groups with lower dispersal capability, either due to the lack of resistance forms or an infrequent distribution of potential sinks and sources. This dispersal limitation of organisms with a high specificity to extreme or rare environment has been already reported for hyperthermophile microorganisms (Whitaker et al., 2003), and might therefore apply to extreme halophiles. In this scenario, the arrival of halophile groups would rely on long-range dispersal events from saline or hypersaline distant sources, which would be produced at a lower rates than for habitat generalists which present a wider repertory of potential sources. Thus, long-range events would play a major role on the progressive configuration of the regional pool of halophiles than on the recolonization of water masses by hypersaline specialists along successive hydroperiods. The periodic recolonization mostly relies on the previously configured regional pool, which is possible due to both the prominent role of sediments as a seed bank and the cell dormancy which reduces the extinction risk at a regional scale (Lennon and Jones, 2011).

Conclusions

Chapter 2: *Archaeal community*

- The aquatic microbiome inhabiting the Monegros desert lacustrine system presented a rich archaeal assembly dominated by members of the DPANN superphylum (including Woesearchaeota, Pacearchaeota and Nanohaloarchaeota) and Halobacteria.
- The relative abundance of the archaeal domain exponentially increased beyond the limit between hypersaline and highly hypersaline waters, reaching a remarkable proportion of the prokaryotic amplicon pool at the most saline end.
- The dynamic approach linked the higher resilience to environmental variability of DPANN members to intermediate colonization and local extinction rates.

Chapter 3: *Bacterial and microeukaryal communities*

- The salinity gradient studied promoted bacterial taxonomic replacement, with a peak in colonization and extinction rates at the boundary between moderate and hypersaline waters. Colonization by better adapted groups maintained an elevated species richness and diversity, differentiating the observed assemblage from those found in more stable environments according to the “salt-in/ salt-out” expected strategies.
- Conversely, microeukaryotes suffered of a progressive filtering of less adapted organisms, with a peak of local extinctions at salinities c.a. 3%, and a lack of replacement leading to a decrease of the richness and diversity indexes. As a consequence, the microeukaryal assemblage at higher salinities was essentially dominated by the photosynthetic algae *Dunaliella* and *Navicula*.

Chapter 4: *Genomic trait based approach*

- A consistent salinity threshold at ca. 5‰ appeared for the prokaryotic community assembly forces captured at the genomic level. Beyond this point, the strong environmental filter imposed by salinity prevailed as the dominant assembly force. At lower salinities, the assembly was potentially dominated by a combination of biotic and abiotic factors.
- The new method provides a conceptual approach which allow to identify environmental thresholds in community assembly and to quantitatively asses the ecological impact of environmental changes.

Chapter 5: *Community assembly processes*

- The importance of the combined and simultaneous effect of dispersal and selection (together with ecological drift), changed along the salinity gradient for bacteria and microeukaryotes.
- There was a transition of species assembly at the of 5‰ salinity threshold from a dispersal-dominated to a selection-dominated scenario.
- The dynamic approach unveiled the mechanistic explanation that reduced communities turnover by the decrease of colonization and extinction rates. Microeukaryotes showed unbalanced colonization and extinction rates that yielded richness decrease. Conversely, bacteria showed equivalent rates and richness remained constant along the gradient.

Chapter 6: *Network analysis*

- The combination of a probabilistic method to determine significant co-occurrence and co-exclusions and a conceptual framework to classify the nature of the observed associations unveiled potential biological interactions along the gradient, and reduced potential artifacts related to environmental and spatial factors.

- The analysis of the co-occurrence and co-exclusion networks showed lower interspecies competition as the osmotic stress increased. Conversely, no stress-promoted facilitation was observed along the salinity gradient, which was linked to the lack of known facilitative mechanisms associated to the osmotic stress in aquatic environments.
- The prominent role of groups with potential parasitic or symbiotic lifestyle within the co-occurrence networks may help to investigate the potential ecology of symbionts.

Chapter 7: *The influence of sediments on the aquatic microbiome*

- Sediments play a prominent role in the ephemeral ponds ecology, acting as a seed bank during desiccation stages and driving aquatic colonization after ponds re-inundation at the beginning of the hydrological cycle.
- Under low salinities, aquatic communities are dominated by habitat generalist which successfully develop on both water and sediment compartments. Increasing saline stress promotes a severe environmental filtering which reduced the contribution of external sources and drive the transition from habitat generalists to specialists groups.

- Adler, P. B., Fajardo, A., Kleinhesselink, A. R., and Kraft, N. J. B. (2013). Trait-based tests of coexistence mechanisms. *Ecology Letters*, 16(10):1294–1306.
- Alcorlo, P. (1999). *Redes tróficas en lagunas salinas temporales de la comarca de Los Monegros (Zaragoza)*. PhD thesis, Universidad Autónoma de Madrid.
- Alcorlo, P., Baltanás, A., and Montes, C. (2001). *Food-web structure in two shallow salt lakes in Los Monegros (NE Spain): energetic vs dynamic constraints*, pages 307–316. Springer Netherlands, Dordrecht.
- Allison, S. D. and Martiny, J. B. H. (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences*, 105(1):11512–11519.
- Alonso, D., Etienne, R. S., and McKane, A. J. (2006). The merits of neutral theory. *Trends in Ecology & Evolution*, 21(8):451–457.
- Álvarez, S., Díaz, P., López-Archilla, A. I., and Guerrero, M. C. (2006). Phytoplankton composition and dynamics in three shallow temporary salt lakes (Monegros, Spain). *Journal of Arid Environments*, 65(4):553–571.
- Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., and Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLOS ONE*, 4(7):e6372.
- Ammon, A., Daniel, M., Navas-Molina, J., Evguenia, K., Morton, J., Zhenjiang, Z. X., Kightley, E., Thompson, L., Hyde, E., Antonio, G., Rob, K., and Gilbert, J. (2017). Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems*, 2(2):e00191–16.
- Anantharaman, K., Brown, C. T., Hug, L. A., Sharon, I., Castelle, C. J., Probst, A. J., Thomas, B. C., Singh, A., Wilkins, M. J., Karaoz, U., Brodie, E. L., Williams, K. H., Hubbard, S. S., and Banfield, J. F. (2016). Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nature Communications*, 7:13219.
- Andersson, M. G. I., Berga, M., Lindström, E. S., and Langenheder, S. (2014). The spatial structure of bacterial communities is influenced by historical environmental conditions. *Ecology*, 95(5):1134–1140.
- Andrei, A. S., Banciu, H. L., and Oren, A. (2012). Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. *FEMS Microbiology Letters*, 330(1):1–9.
- Andrei, A. S., Robeson, M. S., Baricz, A., Coman, C., Muntean, V., Ionescu, A., Etiope, G., Alexe, M., Sicora, C. I., Podar, M., and Banciu, H. L. (2015). Contrasting taxonomic stratification of microbial communities in two hypersaline

- meromictic lakes. *The ISME Journal*, 9(12):2642–2656.
- Antón, J., Oren, A., Benlloch, S., Rodríguez-Valera, F., Amann, R., and Rosselló-Mora, R. (2002). *Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds. *International Journal of Systematic and Evolutionary Microbiology*, 52(2):485–491.
- Araújo, M. B. and Rozenfeld, A. (2014). The geographic scaling of biotic interactions. *Ecography*, 37(5):406–415.
- Araújo, M. B., Rozenfeld, A., Rahbek, C., and Marquet, P. A. (2011). Using species co-occurrence networks to assess the impacts of climate change. *Ecography*, 34(6):897–908.
- Arndt, H. (1993). Rotifers as predators on components of the microbial web (bacteria, heterotrophic flagellates, ciliates) — a review. *Hydrobiologia*, 255(1):231–246.
- Arthur, R. M. (1969). Species packing, and what competition minimizes. *Proceedings of the National Academy of Sciences of the United States of America*, 64(4):1369–1371.
- Auguet, J. C., Barberán, A., and Casamayor, E. O. (2010). Global ecological patterns in uncultured Archaea. *The ISME Journal*, 4:182.
- Baas Becking, L. G. M. (1934). *Geobiologie of Inleiding tot de Milieukunde*. Van Stockum & Zoon, The Hague.
- Baker, M. E. and King, R. S. (2013). Of TITAN and straw men: an appeal for greater understanding of community data. *Freshwater Science*, 32(2):489–506.
- Baltanás, Á. (2001). *Candelacypris* n. gen.(crustacea, ostracoda): a new genus from iberian saline lakes, with a redescription of *eucypris aragonica brehm* & *margalef*, 1948. 101:183–192.
- Barberán, A., Bates, S. T., Casamayor, E. O., and Fierer, N. (2012a). Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal*, 6(2):343–351.
- Barberán, A. and Casamayor, E. O. (2010). Global phylogenetic community structure and β -diversity patterns in surface bacterioplankton metacommunities. *Aquatic Microbial Ecology*, 59(1):1–10.
- Barberán, A. and Casamayor, E. O. (2014). A phylogenetic perspective on species diversity, β -diversity and biogeography for the microbial world. *Molecular Ecology*, 23(23):5868–5876.
- Barberán, A., Casamayor, E. O., and Fierer, N. (2014a). The microbial contribution to macroecology. *Frontiers in Microbiology*, 5:1–8.
- Barberán, A., Fernández-Guerra, A., Bohannan, B. J. M., and Casamayor, E. O. (2012b). Exploration of community traits as ecological markers in microbial metagenomes. *Molecular Ecology*, 21(8):1909–1917.

- Barberán, A., Ramirez, K. S., Leff, J. W., Bradford, M. A., Wall, D. H., and Fierer, N. (2014b). Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecology Letters*, 17(7):794–802.
- Bartrons, M., Catalan, J., and Casamayor, E. O. (2012). High bacterial diversity in epilithic biofilms of oligotrophic mountain lakes. *Microbial Ecology*, 64(4):860–869.
- Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *Third International AAAI Conference on Weblogs and Social Media*.
- Bastolla, U., Fortuna, M. A., Pascual-García, A., Ferrera, A., Luque, B., and Bascompte, J. (2009). The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature*, 458(7241):1018–1020.
- Battle, K., Sébastien, T., Samuel, D., P. A. Saby, N., Walid, H., Mélanie, L., Virginie, N., Claudy, J., Dominique, A., Wincker, P., Cruaud, C., Bispo, A., Maron, P.-A., Chemidlin, P. B. N., and Ranjard, L. (2018). Biogeography of soil bacteria and archaea across France. *Science Advances*, 4(7):eaat1808.
- Bay, S. K., McGeoch, M. A., Gillor, O., Wieler, N., Palmer, D. J., Baker, D. J., Chown, S. L., Greening, C., and Gilbert, J. A. (2020). Soil bacterial communities exhibit strong biogeographic patterns at fine taxonomic resolution. *mSystems*, 5(4):e00540–20.
- Bell, T. (2010). Experimental tests of the bacterial distance–decay relationship. *The ISME Journal*, 4(11):1357–1365.
- Belmaker, J. and Jetz, W. (2011). Cross-scale variation in species richness–environment associations. *Global Ecology and Biogeography*, 20(3):464–474.
- Belmaker, J., Zarnetske, P., Tuanmu, M. N., Zonneveld, S., Record, S., Strecker, A., and Beaudrot, L. (2015). Empirical evidence for the scale dependence of biotic interactions. *Global Ecology and Biogeography*, 24(7):750–761.
- Berga, M., Székely, A. J., and Langenheder, S. (2012). Effects of disturbance intensity and frequency on bacterial community composition and function. *PLOS ONE*, 7(5):e36959.
- Berryman, A. (1992). The origins and evolution of predator-prey theory. *Ecology*, 73(5):1530–1535.
- Bertness, M. D. and Callaway, R. (1994). Positive interactions in communities. *Trends in Ecology & Evolution*, 9(5):191–193.
- Bertness, M. D. and Ewanchuk, P. J. (2002). Latitudinal and climate-driven variation in the strength and nature of biological interactions in New England salt marshes. *Oecologia*, 132(3):392–401.
- Bever, J. D. (1994). Feedback between plants and their soil communities in an old field community. *Ecology*, 75(7):1965–1977.

- Blanquart, F. and Gandon, S. (2013). Time-shift experiments and patterns of adaptation across time and space. *Ecology Letters*, 16(1):31–38.
- Blois, J. L., Gotelli, N. J., Behrensmeier, A. K., Faith, J. T., Lyons, S. K., Williams, J. W., Amatangelo, K. L., Bercovici, A., Du, A., Eronen, J. T., Graves, G. R., Jud, N., Labandeira, C., Looy, C. V., McGill, B., Patterson, D., Potts, R., Riddle, B., Terry, R., Tóth, A., Villaseñor, A., and Wing, S. (2014). A framework for evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil pollen associations across the late Quaternary. *Ecography*, 37(11):1095–1108.
- Bobay, L.-M. (2020). *The prokaryotic species concept and challenges*, pages 21–49. Springer International Publishing, Cham.
- Booth, I. R. and Louis, P. (1999). Managing hypoosmotic stress: Aquaporins and medianosensitive channels in *Escherichia coli*. *Current Opinion in Microbiology*, 2(2):166–169.
- Brenner, D. J., Staley, J. T., and Krieg, N. R. (2015). Classification of procaryotic organisms and the concept of bacterial speciation. *Bergey’s Manual of Systematics of Archaea and Bacteria*, pages 27–31.
- Brown, A. D. (1976). Microbial water stress. *Bacteriological reviews*, 40(4):803–846.
- Brown, J. H. (1995). *Macroecology*, page 270. Chicago.
- Brown, J. H. and Kodric-Brown, A. (1977). Turnover Rates in Insular Biogeography: Effect of Immigration on Extinction. *Ecology*, 58(2):445–449.
- Buckley, L. B., Davies, T. J., Ackerly, D. D., Kraft, N. J. B., Harrison, S. P., Anacker, B. L., Cornell, H. V., Damschen, E. I., Grytnes, J.-A., Hawkins, B. A., McCain, C. M., Stephens, P. R., and Wiens, J. J. (2010). Phylogeny, niche conservatism and the latitudinal diversity gradient in mammals. *Proceedings of the Royal Society B: Biological Sciences*, 277(1691):2131–2138.
- Bunin, G. (2017). Ecological communities with Lotka-Volterra dynamics. *Physical review. E*, 95(4-1):42414.
- Burke, C., Steinberg, P., Rusch, D., Kjelleberg, S., and Thomas, T. (2011). Bacterial community assembly based on functional genes rather than species. *Proceedings of the National Academy of Sciences*, 108(34):14288–14293.
- Buttigieg, P. L., Morrison, N., Smith, B., Mungall, C. J., Lewis, S. E., and Consortium, t. E. (2013). The environment ontology: contextualising biological and biomedical entities. *Journal of Biomedical Semantics*, 4(1):43.
- Cáliz, J., Triadó-Margarit, X., Camarero, L., and Casamayor, E. O. (2018). A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences*, 115(48):12229–12234.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and

- Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13(7):581–583.
- Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., Paolini, L., Pugnaire, F. I., Newingham, B., Aschehoug, E. T., Armas, C., Kikodze, D., and Cook, B. J. (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417(6891):844–848.
- Cao, H., Chen, R., Wang, L., Jiang, L., Yang, F., Zheng, S., Wang, G., and Lin, X. (2016). Soil pH, total phosphorus, climate and distance are the major factors influencing microbial activity at a regional spatial scale. *Scientific Reports*, 6(1):25815.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Meth*, 7(5):335–336.
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., and Knight, R. (2011). Global patterns of 16s rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(1):4516–4522.
- Caporaso, J. G., Paszkiewicz, K., Field, D., Knight, R., and Gilbert, J. A. (2012). The Western English Channel contains a persistent microbial seed bank. *The ISME Journal*, 6(6):1089–1093.
- Carscadden, K. A., Emery, N. C., Arnillas, C. A., Cadotte, M. W., Afkhami, M. E., Gravel, D., Livingstone, S. W., and Wiens, J. J. (2020). Niche breadth: causes and consequences for ecology, evolution, and conservation. *The Quarterly Review of Biology*, 95(3):179–214.
- Carver, S., Storey, A., Spafford, H., Lynas, J., Chandler, L., and Weinstein, P. (2009). Salinity as a driver of aquatic invertebrate colonisation behaviour and distribution in the wheatbelt of Western Australia. *Hydrobiologia*, 617(1):75–90.
- Casamayor, E. O., Triadó-Margarit, X., and Castañeda, C. (2013). Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. *FEMS Microbiology Ecology*, 85(3):503–518.
- Castañeda, C., Herrero, J., and Conesa, J. A. (2013). Distribution, morphology and habitats of saline wetlands: a case study from Monegros, Spain. *Geologica Acta: an international earth science journal.*, 11:371–388.
- Castelle, C. J. and Banfield, J. F. (2018). Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell*, 172(6):1181–1197.

- Castelle, C. J., Wrighton, K. C., Thomas, B. C., Hug, L. A., Brown, C. T., Wilkins, M. J., Frischkorn, K. R., Tringe, S. G., Singh, A., Markillie, L. M., Taylor, R. C., Williams, K. H., and Banfield, J. F. (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Current Biology*, 25(6):690–701.
- Catalan, J., Barbieri, M. G., Bartumeus, F., Bituššák, P., Botev, I., Brancelj, A., Cogălniceanu, D. A. N., Manca, M., Marchetto, A., Ognjanova-Rumenova, N., Pla, S., Rieradevall, M., Sorvari, S., Štefková, E., Stuchlík, E., and Ventura, M. (2009). Ecological thresholds in European alpine lakes. *Freshwater Biology*, 54(12):2494–2517.
- Caumette, P. (1993). Ecology and physiology of phototrophic bacteria and sulfate-reducing bacteria in marine salterns. *Experientia*, 49:473–481.
- Cavalazzi, B., Barbieri, R., Gómez, F., Capaccioni, B., Olsson-Francis, K., Pondrelli, M., Rossi, A., Hickman-Lewis, K., Agangi, A., Gasparotto, G., Glamoclija, M., Ori, G., Rodriguez, N., and Hagos, M. (2019). The dallol geothermal area, northern afar (ethiopia)—an exceptional planetary field analog on earth. *Astrobiology*, 19(4):553–578. PMID: 30653331.
- Chalmandrier, L., Pansu, J., Zinger, L., Boyer, F., Coissac, E., Génin, A., Gielly, L., Lavergne, S., Legay, N., Schilling, V., Taberlet, P., Münkemüller, T., and Thuiller, W. (2019). Environmental and biotic drivers of soil microbial β -diversity across spatial and phylogenetic scales. *Ecography*, 42(12):2144–2156.
- Chase, J. M. (2014). Spatial scale resolves the niche versus neutral theory debate. *Journal of Vegetation Science*, 25(2):319–322.
- Chen, Y. J., Leung, P. M., Wood, J. L., Bay, S. K., Hugenholtz, P., Kessler, A. J., Shelley, G., Waite, D. W., Franks, A. E., Cook, P. L. M., and Greening, C. (2021). Metabolic flexibility allows bacterial habitat generalists to become dominant in a frequently disturbed ecosystem. *The ISME Journal*, (15):2986–3004.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31:343–366.
- Cohan, F. M. (2002). What are bacterial species? *Annual Review of Microbiology*, 56(1):457–487.
- Comte, J., Langenheder, S., Berga, M., and Lindström, E. S. (2017). Contribution of different dispersal sources to the metabolic response of lake bacterioplankton following a salinity change. *Environmental Microbiology*, 19(1):251–260.
- Comte, J., Lindström, E. S., Eiler, A., and Langenheder, S. (2014). Can marine bacteria be recruited from freshwater sources and the air? *The ISME Journal*, 8(12):2423–2430.
- Conesa, J. A., Castañeda, C., and Pedrol, J. (2011). *Las saladas de Monegros y su entorno: hábitats y paisaje vegetal*. Consejo de Protección de la Naturaleza de

- Aragón.
- Connor, N., Barberán, A., and Clauset, A. (2017). Using null models to infer microbial cooccurrence networks. *PLoS ONE*, 12(5):1–23.
- Cotgreave, P. (1994). The relation between body size and abundance in a bird community: the effects of phylogeny and competition. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 256(1346):147–149.
- Cowan, D., Ramond, J.-B., Makhallanyane, T., and De Maayer, P. (2015). Metagenomics of extreme environments. *Current Opinion in Microbiology*, 25:97–102.
- Crump, B. C., Amaral-Zettler, L. A., and Kling, G. W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The ISME Journal*, 6(9):1629–1639.
- Csardi, G. and Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5):1–9.
- Dal Molin, A. and Di Camillo, B. (2018). How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives. *Briefings in Bioinformatics*, pages 1384–1394.
- Damuth, J. (1981). Population density and body size in mammals. *Nature*, 290(5808):699–700.
- Darcy, J. L., Gendron, E. M. S., Sommers, P., Porazinska, D. L., and Schmidt, S. K. (2018). Island biogeography of cryoconite hole bacteria in antarctica’s taylor valley and around the world. *Frontiers in Ecology and Evolution*, 6.
- Darwin, C. R. (1859). *On the origin of species by means of natural selection*. London.
- Debray, R., Herbert, R. A., Jaffe, A. L., Crits-Christoph, A., Power, M. E., and Koskella, B. (2021). Priority effects in microbiome assembly. *Nature Reviews Microbiology*, (20):109–121.
- Demergasso, C., Escudero, L., Casamayor, E. O., Chong, G., Balagué, V., and Pedrós-Alió, C. (2008). Novelty and spatio-temporal heterogeneity in the bacterial diversity of hypersaline Lake Tebenquiche (Salar de Atacama). *Extremophiles*, 12(4):491–504.
- Dennis, P. P. and Shimmin, L. C. (1997). Evolutionary divergence and salinity-mediated selection in halophilic archaea. *Microbiology and Molecular Biology Reviews*, 61(1):90–104.
- Dillon, J., Carlin, M., Gutierrez, A., Nguyen, V., and McLain, N. (2013). Patterns of microbial diversity along a salinity gradient in the guerrero negro solar saltern, baja ca sur, mexico. *Frontiers in Microbiology*, 4.
- Dillon, J., M. McMath, L., and L. Trout, A. (2009). *Seasonal changes in bacterial diversity in the Salton Sea*, volume 632, pages 49–64.
- Dini-Andreote, F., Stegen, J. C., van Elsas, J. D., and Salles, J. F. (2015). Disentan-

- gling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11).
- Dornelas, M., Gotelli Nicholas, J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., and Magurran Anne, E. (2014). Assemblage time series reveal biodiversity change but not systematic loss. *Science*, 344(6181):296–299.
- Dufrene, M. and Legendre, P. (1997). Species Assemblages and Indicator Species: The Need for a Flexible Asymmetrical Approach. *Ecological Monographs*, 67(3):345–366.
- Dupont, C. L., Larsson, J., Yooseph, S., Ininbergs, K., Goll, J., Asplund-Samuelsson, J., McCrow, J. P., Celepli, N., Allen, L. Z., Ekman, M., Lucas, A. J., Hagström, Å., Thiagarajan, M., Brindefalk, B., Richter, A. R., Andersson, A. F., Tenney, A., Lundin, D., Tovchigrechko, A., Nylander, J. A. A., Bami, D., Badger, J. H., Allen, A. E., Rusch, D. B., Hoffman, J., Norrby, E., Friedman, R., Pinhassi, J., Venter, J. C., and Bergman, B. (2014). Functional tradeoffs underpin salinity-driven divergence in microbial community composition. *PLOS ONE*, 9(2):e89549.
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19):2460–2461.
- Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Meth*, 10(10):996–998.
- Edgar, R. C. (2016). Unoise2: improved error-correction for illumina 16s and its amplicon sequencing. *bioRxiv*.
- Edgar, R. C. and Flyvbjerg, H. (2015). Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, 31(21):3476–3482.
- Edgerton, M. E. and Brimblecombe, P. (1981). Thermodynamics of halobacterial environments. *Canadian Journal of Microbiology*, 27(9):899–909.
- Edwards, K. F., Litchman, E., and Klausmeier, C. A. (2013). Functional traits explain phytoplankton responses to environmental gradients across lakes of the United States. *Ecology*, 94(7):1626–1635.
- Eiler, A., Heinrich, F., and Bertilsson, S. (2012). Coherent dynamics and association networks among lake bacterioplankton taxa. *The ISME Journal*, 6(2):330–342.
- El Hidri, D., Guesmi, A., Najjari, A., Cherif, H., Ettoumi, B., Hamdi, C., Boudabous, A., and Cherif, A. (2013). Cultivation-dependent assessment, diversity, and ecology of haloalkaliphilic bacteria in arid saline systems of southern Tunisia. *BioMed research international*, 2013:648141.
- Enke, T. N., Datta, M. S., Schwartzman, J., Cermak, N., Schmitz, D., Barriere, J., Pascual-García, A., and Cordero, O. X. (2019). Modular assembly of polysaccharide-degrading marine microbial communities. *Current Biology*,

- 29(9):1528–1535.
- Erdős, P. and Rényi, A. (1960). *On the Evolution of Random Graphs*, pages 17–61.
- Estrada, M., Henriksen, P., Gasol, J. M., Casamayor, E. O., and Pedrós-Alió, C. (2004). Diversity of planktonic photoautotrophic microorganisms along a salinity gradient as depicted by microscopy, flow cytometry, pigment analysis and DNA-based methods. *FEMS Microbiology Ecology*, 49(2):281–293.
- Etienne, R. S. and Alonso, D. (2005). A dispersal-limited sampling theory for species and alleles. *Ecology Letters*, 8(11):1147–1156.
- Etienne, R. S. and Alonso, D. (2007). Neutral community theory: how stochasticity and dispersal-limitation can explain species coexistence. *Journal of Statistical Physics*, 128(1):485–510.
- Evans, S., Martiny, J. B. H., and Allison, S. D. (2017). Effects of dispersal and selection on stochastic assembly in microbial communities. *The ISME Journal*, 11(1):176–185.
- Faci González, J. M. and Martínez-Cob, A. (1991). Cálculo de la evapotranspiración de referencia en Aragón.
- Faust, K. and Raes, J. (2012). Microbial interactions: from networks to models. *Nature Reviews Microbiology*, 10(8):538–550.
- Fernández, A. B., Ghai, R., Martin-Cuadrado, A. B., Sánchez-Porro, C., Rodríguez-Valera, F., and Ventosa, A. (2014). Prokaryotic taxonomic and metabolic diversity of an intermediate salinity hypersaline habitat assessed by metagenomics. *FEMS Microbiology Ecology*, 88(3):623–635.
- Fernández-Guerra, A. and Casamayor, E. O. (2012). Habitat-associated phylogenetic community patterns of microbial ammonia oxidizers. *PLOS ONE*, 7(10):e47330.
- Ferrera, I., Massana, R., Casamayor, E. O., Balagué, V., Sánchez, O., Pedrós-Alió, C., and Mas, J. (2004). High-diversity biofilm for the oxidation of sulfide-containing effluents. *Applied Microbiology and Biotechnology*, 64(5):726–734.
- Fierer, N. and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 103(3):626–631.
- Fierer, N., Nemergut, D., Knight, R., and Craine, J. M. (2010). Changes through time: integrating microorganisms into the study of succession. *Research in Microbiology*, 161(8):635–642.
- Filker, S., Forster, D., Weinisch, L., Mora-Ruiz, M., González, B., Farías, M. E., Rosselló-Móra, R., and Stoeck, T. (2017). Transition boundaries for protistan species turnover in hypersaline waters of different biogeographic regions. *Environmental Microbiology*, 19(8):3186–3200.
- Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species.

- Science*, 296(5570):1061–1063.
- Fodelianakis, S., Lorz, A., Valenzuela-Cuevas, A., Barozzi, A., Booth, J. M., and Daffonchio, D. (2019). Dispersal homogenizes communities via immigration even at low rates in a simplified synthetic bacterial metacommunity. *Nature Communications*, 10(1):1314.
- Foster, D., Swanson, F., Aber, J., Burke, I., Brokaw, N., Tilman, D., and Knapp, A. (2003). The importance of land-use legacies to ecology and conservation. *BioScience*, 53(1):77–88.
- Foster, K. and Bell, T. (2012). Competition, not cooperation, dominates interactions among culturable microbial species. *Current Biology*, 22(19):1845–1850.
- Fox, J. W. and Harder, L. D. (2015). Using a “time machine” to test for local adaptation of aquatic microbes to temporal and spatial environmental variation. *Evolution*, 69(1):136–145.
- Freilich, M. A., Wieters, E., Broitman, B. R., Marquet, P. A., and Navarrete, S. A. (2018). Species co-occurrence networks: Can they reveal trophic and non-trophic interactions in ecological communities? *Ecology*, 99(3):690–699.
- Friedman, J. and Alm, E. J. (2012). Inferring correlation networks from genomic survey data. *PLOS Computational Biology*, 8(9):e1002687.
- Fukami, T. (2015). Historical contingency in community assembly: integrating niches, species pools, and priority effects. *Annual Review of Ecology, Evolution, and Systematics*, 46(1):1–23.
- Fukami, T., Beaumont, H. J. E., Zhang, X.-X., and Rainey, P. B. (2007). Immigration history controls diversification in experimental adaptive radiation. *Nature*, 446(7134):436–439.
- Galand, P. E., Gutiérrez-Provecho, C., Massana, R., Gasol, J. M., and Casamayor, E. O. (2010). Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnology and Oceanography*, 55(5):2117–2125.
- Galinski, E. (1995). Osmoadaptation in bacteria. volume 37 of *Advances in Microbial Physiology*, pages 273–328. Academic Press.
- Galinski, E. A. and Trüper, H. G. (1994). Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiology Reviews*, 15(2-3):95–108.
- Garzon-Lopez, C. X., Jansen, P. A., Bohlman, S. A., Ordonez, A., and Olff, H. (2013). Effects of sampling scale on patterns of habitat association in tropical trees. *Journal of Vegetation Science*, 25(2):349–362.
- Gasol, J. M., Casamayor, E. O., Joint, I., Garde, K., Gustavson, K., Benlloch, S., Díez, B., Schauer, M., Massana, R., and Pedrós-Alió, C. (2004). Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquatic Microbial Ecology*, 34(2):193–206.

- Ghai, R., Pašić, L., Fernández, A. B., Martin-Cuadrado, A.-B., Mizuno, C. M., McMahon, K. D., Papke, R. T., Stepanauskas, R., Rodriguez-Brito, B., Rohwer, F., Sánchez-Porro, C., Ventosa, A., and Rodríguez-Valera, F. (2011). New abundant microbial groups in aquatic hypersaline environments. *Scientific Reports*, 1(1):135.
- Ghoul, M. and Mitri, S. (2016). The ecology and evolution of microbial competition. *Trends in Microbiology*, 24(10):833–845.
- Goberna, M., Montesinos-Navarro, A., Valiente-Banuet, A., Colin, Y., Gómez-Fernández, A., Donat, S., Navarro-Cano, J. A., and Verdú, M. (2019). Incorporating phylogenetic metrics to microbial co-occurrence networks based on amplicon sequences to discern community assembly processes. *Molecular Ecology Resources*, 19(6):1552–1564.
- Goberna, M., Navarro-Cano, J. A., Valiente-Banuet, A., García, C., and Verdú, M. (2014). Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecology Letters*, 17(10):1191–1201.
- Gong, J., Qing, Y., Guo, X., and Warren, A. (2014). “*Candidatus* Sonnebornia yantaiensis”, a member of candidate division OD1, as intracellular bacteria of the ciliated protist *Paramecium bursaria* (Ciliophora, Oligohymenophorea). *Systematic and Applied Microbiology*, 37(1):35–41.
- Gotelli, N. J. and Graves, G. R. (1996). *Null Models in Ecology*. Smithsonian Institution Press.
- Graham, C. H., Storch, D., and Machac, A. (2018). Phylogenetic scale in ecology and evolution. *Global Ecology and Biogeography*, 27(2):175–187.
- Granda, E., Escudero, A., de la Cruz, M., and Valladares, F. (2012). Juvenile–adult tree associations in a continental Mediterranean ecosystem: no evidence for sustained and general facilitation at increased aridity. *Journal of Vegetation Science*, 23(1):164–175.
- Gravel, D., Canham, C. D., Beaudet, M., and Messier, C. (2006). Reconciling niche and neutrality: the continuum hypothesis. *Ecology Letters*, 9(4):399–409.
- Grossmann, L., Beisser, D., Bock, C., Chatzinotas, A., Jensen, M., Preisfeld, A., Psenner, R., Rahmann, S., Wodniok, S., and Boenigk, J. (2016). Trade-off between taxon diversity and functional diversity in European lake ecosystems. *Molecular Ecology*, 25(23):5876–5888.
- Gunde-Cimerman, N., Plemenitaš, A., and Oren, A. (2018). Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiology Reviews*, 42(3):353–375.
- Guo, X., Feng, J., Shi, Z., Zhou, X., Yuan, M., Tao, X., Hale, L., Yuan, T., Wang, J., Qin, Y., Zhou, A., Fu, Y., Wu, L., He, Z., Van Nostrand, J. D., Ning, D.,

- Liu, X., Luo, Y., Tiedje, J. M., Yang, Y., and Zhou, J. (2018). Climate warming leads to divergent succession of grassland microbial communities. *Nature Climate Change*, 8(9):813–818.
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29(8):1072–1075.
- Gutiérrez, F. and Gutiérrez, M. (2014). *Playa-lakes and yardangs in the Bujaraloz-Sástago endorheic area, Central Ebro Basin*, pages 145–153. Springer Netherlands, Dordrecht.
- Habtom, H., Pasternak, Z., Matan, O., Azulay, C., Gafny, R., and Jurkevitch, E. (2019). Applying microbial biogeography in soil forensics. *Forensic Science International: Genetics*, 38:195–203.
- Hagemann, M. (2016). *Coping with high and variable salinity: molecular aspects of compatible solute accumulation*, pages 359–372. Springer International Publishing, Cham.
- Hamm, J. N., Erdmann, S., Eloë-Fadrosch, E. A., Angeloni, A., Zhong, L., Brownlee, C., Williams, T. J., Barton, K., Carswell, S., Smith, M. A., Brazendale, S., Hancock, A. M., Allen, M. A., Raftery, M. J., and Cavicchioli, R. (2019). Unexpected host dependency of Antarctic Nanohaloarchaeota. *Proceedings of the National Academy of Sciences*, 116(29):14661–14670.
- Hanski, I. (1994). A practical model of metapopulation dynamics. *Journal of Animal Ecology*, 63(1):151–162.
- Harding, T., Roger, A. J., and Simpson, A. G. B. (2017). Adaptations to high salt in a halophilic protist: differential expression and gene acquisitions through duplications and gene transfers. *Frontiers in microbiology*, 8:944.
- He, C., Keren, R., Whittaker, M. L., Farag, I. F., Doudna, J. A., Cate, J. H. D., and Banfield, J. F. (2021). Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nature microbiology*, 6(3):354–365.
- He, Q., Bertness, M. D., and Altieri, A. H. (2013). Global shifts towards positive species interactions with increasing environmental stress. *Ecology Letters*, 16(5):695–706.
- Heidelberg, K., Nelson, W., Holm, J., Eisenkolb, N., Andrade, K., and Emerson, J. (2013). Characterization of eukaryotic microbial diversity in hypersaline Lake Tyrrell, Australia. *Frontiers in Microbiology*, 4:115.
- Helmus, M. R., Savage, K., Diebel, M. W., Maxted, J. T., and Ives, A. R. (2007). Separating the determinants of phylogenetic community structure. *Ecology Letters*, 10(10):917–925.
- Hernandez, D. J., David, A. S., Menges, E. S., Searcy, C. A., and Afkhami, M. E. (2021). Environmental stress destabilizes microbial networks. *The ISME Journal*,

- 15:1722–1734.
- Herrero, J. and Snyder, R. L. (1997). Aridity and irrigation in Aragon, Spain. *Journal of Arid Environments*, 35(3):535–547.
- Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *The American Naturalist*, 163(2):192–211.
- Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., Filstrup, C. T., Harpole, W. S., Hodapp, D., Larsen, S., Lewandowska, A. M., Seabloom, E. W., Van de Waal, D. B., and Ryabov, A. B. (2018). Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology*, 55(1):169–184.
- HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., and Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, 43(1):227–248.
- Ho, A., Di Lonardo, D. P., and Bodelier, P. L. E. (2017). Revisiting life strategy concepts in environmental microbial ecology. *FEMS Microbiology Ecology*, 93(3).
- Höfle, G. M., Haas, H., and Dominik, K. (1999). Seasonal dynamics of bacterio-plankton community structure in a eutrophic lake as determined by 5S rRNA analysis. *Applied and Environmental Microbiology*, 65(7):3164–3174.
- Hohenegger, J. (2014). Species as the basic units in evolution and biodiversity: recognition of species in the recent and geological past as exemplified by larger foraminifera. *Gondwana Research*, 25(2):707–728.
- Holmer, M. and Storkholm, P. (2001). Sulphate reduction and sulphur cycling in lake sediments: a review. *Freshwater Biology*, 46(4):431–451.
- Holyoak, M. (2005). *Metacommunities: a framework for large scale community ecology.*, page 513. University of Chicago Press, Chicago.
- Hongchen, J., Hailiang, D., Gengxin, Z., Bingsong, Y., R., C. L., and W., F. M. (2006). Microbial diversity in water and sediment of Lake Chaka, an athalas-sohaline lake in northwestern China. *Applied and Environmental Microbiology*, 72(6):3832–3845.
- Hortal, J., de Bello, F., Diniz-Filho, J. A. F., Lewinsohn, T. M., Lobo, J. M., and Ladle, R. J. (2015). Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46(1):523–549.
- Hubbell, S. P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.
- Hugoni, M., Taib, N., Debroas, D., Domaizon, I., Jouan Dufournel, I., Bronner, G., Salter, I., Agogu  , H., Mary, I., and Galand, P. E. (2013). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proceedings of the National Academy of Sciences*, 110(15):6004–6009.

- Imhoff, J. F. (2001). True marine and halophilic anoxygenic phototrophic bacteria. *Archives of Microbiology*, 176(4):243–254.
- Isaac, N. J. B., Storch, D., and Carbone, C. (2011). Taxonomic variation in size–density relationships challenges the notion of energy equivalence. *Biology Letters*, 7(4):615–618.
- Jacob, J. H., Hussein, E. I., Shakhathreh, M. A. K., and Cornelison, C. T. (2017). Microbial community analysis of the hypersaline water of the Dead Sea using high-throughput amplicon sequencing. *MicrobiologyOpen*, 6(5):e00500.
- Javor, B. (1989). *Dunaliella and Other Halophilic, Eucaryotic Algae*, pages 147–158. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Jiang, H., Dong, H., Zhang, G., Yu, B., Chapman, L. R., and Fields, M. W. (2006). Microbial diversity in water and sediment of Lake Chaka, an athalassohaline lake in northwestern China. *Applied and environmental microbiology*, 72(6):3832–3845.
- Johanna, S., Per, K.-S., Mikael, P., Björn, C., Anders, T., Per, L., Åke, H., and Lasse, R. (2012). Recruitment of members from the rare biosphere of marine bacterioplankton communities after an environmental disturbance. *Applied and Environmental Microbiology*, 78(5):1361–1369.
- Kan, J., Clingenpeel, S., Macur, R. E., Inskeep, W. P., Lovalvo, D., Varley, J., Gorby, Y., Mcdermott, T. R., and Nealson, K. (2011). Archaea in Yellowstone Lake. *The ISME Journal*, 5(11):1784–1795.
- Kempf, B. and Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Archives of Microbiology*, 170(5):319–330.
- Kim, J.-S., Makama, M., Petito, J., Park, N.-H., Cohan, F. M., and Dungan, R. S. (2012). Diversity of bacteria and archaea in hypersaline sediment from Death Valley National Park, California. *MicrobiologyOpen*, 1(2):135–148.
- Kis-Papo, T., Weig, A. R., Riley, R., Peršoh, D., Salamov, A., Sun, H., Lipzen, A., Wasser, S. P., Rambold, G., Grigoriev, I. V., and Nevo, E. (2014). Genomic adaptations of the halophilic Dead Sea filamentous fungus *Eurotium rubrum*. *Nature Communications*, 5(1):3745.
- Klappenbach, J. A., Dunbar, J. M., and Schmidt, T. M. (2000). rRNA operon copy number reflects ecological strategies of bacteria. *Applied and environmental microbiology*, 66(4):1328–1333.
- Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., and Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29(5):592–599.
- Kraft, N. J. B., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., Stegen, J. C., Vellend, M., Boyle, B., Anderson, M. J., Cornell, H. V., Davies, K. F., Freestone, A. L., Inouye, B. D., Harrison, S. P., and Myers, J. A.

- (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, 333(6050):1755–1758.
- Krause, S., Le Roux, X., Niklaus, P. A., Van Bodegom, P. M., Lennon, J. T., Bertilsson, S., Grossart, H.-P., Philippot, L., and Bodelier, P. L. E. (2014). Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Frontiers in Microbiology*, 5.
- Kushner, D. J. (1978). *Life in high salt and solute concentrations : halophilic bacteria*, pages 317–368. Academic Press.
- Kushner, D. J. (1985). *The Halobacteriaceae*, pages 171–214. Academic Press.
- Ladau, J. and Elie-Fadrosh, E. A. (2019). Spatial, temporal, and phylogenetic scales of microbial ecology. *Trends in Microbiology*, 27(8):662–669.
- Lang, F. (2007). Mechanisms and significance of cell volume regulation. *Journal of the American College of Nutrition*, 26(5):613S–623S.
- Langenheder, S., Lindström, E. S., and Tranvik, L. J. (2005). Weak coupling between community composition and functioning of aquatic bacteria. *Limnology and Oceanography*, 50(3):957–967.
- Langille, M. G. I., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., and Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, 31(9):814–821.
- Lankiewicz, T. S., Cottrell, M. T., and Kirchman, D. L. (2016). Growth rates and rRNA content of four marine bacteria in pure cultures and in the Delaware estuary. *The ISME Journal*, 10(4):823–832.
- Law, R. and Morton, R. D. (1996). Permanence and the assembly of ecological communities. *Ecology*, 77(3):762–775.
- Law, R., Weatherby, A. J., and Warren, P. H. (2000). On the Invasibility of Persistent Protist Communities. *Oikos*, 88(2):319–326.
- Lawrence, D. and Barraclough, T. G. (2016). Evolution of resource use along a gradient of stress leads to increased facilitation. *Oikos*, 125(9):1284–1295.
- Lawton, J. H. (1999). Are there general laws in ecology? *Oikos*, 84(2):177–192.
- Lear, G., Bellamy, J., Case, B. S., Lee, J. E., and Buckley, H. L. (2014). Fine-scale spatial patterns in bacterial community composition and function within freshwater ponds. *The ISME Journal*, 8(8):1715–1726.
- Legendre, P., Borcard, D., and Peres-Neto, P. R. (2005). Analyzing beta diversity: Partitioning the spatial variation of community composition data. *Ecological Monographs*, 75(4):435–450.
- Legendre, P. and Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2):271–280.

- Leibold, M. A. and Chase, J. M. (2018). *Metacommunity Ecology, Volume 59*. Princeton University Press.
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M., and Gonzalez, A. (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecology Letters*, 7(7):601–613.
- Lennon, J. T. and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology*, 9(2):119–130.
- León, M. J., Fernández, A. B., Ghai, R., Sánchez-Porro, C., Rodríguez-Valera, F., and Ventosa, A. (2014). From metagenomics to pure culture: Isolation and characterization of the moderately halophilic bacterium *Spiribacter salinus* gen. nov., sp. nov. *Applied and Environmental Microbiology*, 80(13):3850–3857.
- Levin, S. A. (1992). The problem of pattern and scale in ecology: The Robert H. MacArthur award lecture. *Ecology*, 73(6):1943–1967.
- Liao, J., Cao, X., Zhao, L., Wang, J., Gao, Z., Wang, M. C., and Huang, Y. (2016). The importance of neutral and niche processes for bacterial community assembly differs between habitat generalists and specialists. *FEMS Microbiology Ecology*, 92(11).
- Lindström, E. S. and Langenheder, S. (2012). Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports*, 4(1):1–9.
- Lipsewiers, Y. A., Hopmans, E. C., Sinninghe Damsté, J. S., and Villanueva, L. (2018). Potential recycling of thaumarchaeotal lipids by DPANN Archaea in seasonally hypoxic surface marine sediments. *Organic Geochemistry*, 119:101–109.
- Liu, X., Li, M., Castelle, C. J., Probst, A. J., Zhou, Z., Pan, J., Liu, Y., Banfield, J. F., and Gu, J.-D. (2018). Insights into the ecology, evolution, and metabolism of the widespread Woese archaeal lineages. *Microbiome*, 6(1):102.
- Llorens-Marès, T., Yooseph, S., Goll, J., Hoffman, J., Vila-Costa, M., Borrego, C. M., Dupont, C. L., and Casamayor, E. O. (2015). Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. *The ISME Journal*, 9(7):1648–1661.
- Logares, R., Tesson, S. V., Canbäck, B., Pontarp, M., Hedlund, K., and Rengefors, K. (2018). Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environmental Microbiology*, 20(6):2231–2240.
- Logue, J. B. and Lindström, E. S. (2010). Species sorting affects bacterioplankton community composition as determined by 16S rDNA and 16S rRNA fingerprints. *The ISME Journal*, 4(6):729–738.
- Losos, J. B. (2008). Phylogenetic niche conservatism, phylogenetic signal and the

- relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, 11(10):995–1003.
- Losos, J. B. and Ricklefs, R. E. (2010). *The Theory of Island Biogeography Revisited*. Princeton University Press.
- Louca, S., Parfrey, L. W., and Doebeli, M. (2016). Decoupling function and taxonomy in the global ocean microbiome. *Science*, 353(6305):1272–1277.
- Louca, S., Polz, M. F., Mazel, F., Albright, M. B. N., Huber, J. A., O’Connor, M. I., Ackermann, M., Hahn, A. S., Srivastava, D. S., Crowe, S. A., Doebeli, M., and Parfrey, L. W. (2018). Function and functional redundancy in microbial systems. *Nature Ecology & Evolution*, 2(6):936–943.
- Lozupone, C. A. and Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, 104(27):11436–11440.
- Luo, Z., Liu, J., Zhao, P., Jia, T., Li, C., and Chai, B. (2019). Biogeographic patterns and assembly mechanisms of bacterial communities differ between habitat generalists and specialists across elevational gradients. *Frontiers in Microbiology*, 10.
- MacArthur, R. H. (1965). Patterns of species diversity. *Biological Reviews*, 40(4):510–533.
- MacArthur, R. H. and Wilson, E. O. (1963). An equilibrium theory of insular zoogeography. *Evolution*, 17(4):373–387.
- MacArthur, R. H. and Wilson, E. O. (1967). *The Theory of Island Biogeography*. Monographs in population biology. Princeton University Press, Princeton, N.J.
- Machado, D., Maistrenko, O. M., Andrejev, S., Kim, Y., Bork, P., Patil, K. R., and Patil, K. R. (2021). Polarization of microbial communities between competitive and cooperative metabolism. *Nature Ecology & Evolution*, 5(2):195–203.
- Maestre, F. T., Callaway, R. M., Valladares, F., and Lortie, C. J. (2009). Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *Journal of Ecology*, 97(2):199–205.
- Maestre, F. T. and Cortina, J. (2004). Do positive interactions increase with abiotic stress? A test from a semi-arid steppe. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(5):331–333.
- Maestre, F. T., Cortina, J., Bautista, S., and Bellot, J. (2003). Does *Pinus halepensis* facilitate the establishment of shrubs in Mediterranean semi-arid afforestations? *Forest Ecology and Management*, 176(1):147–160.
- Maestre, F. T., Valladares, F., and Reynolds, J. F. (2005). Is the change of plant–plant interactions with abiotic stress predictable? A meta-analysis of field results in arid environments. *Journal of Ecology*, 93(4):748–757.
- Magurran, A. E. (2008). Diversity Over Time. *Folia Geobotanica*, 43(3):319–327.
- Mardis, E. R. (2013). Next-Generation Sequencing Platforms. *Annual Review of*

- Analytical Chemistry*, 6(1):287–303.
- Mariadassou, M., Pichon, S., and Ebert, D. (2015). Microbial ecosystems are dominated by specialist taxa. *Ecology Letters*, 18(9):974–982.
- Markowitz, V. M., Chen, I.-M. A., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., Ratner, A., Huang, J., Woyke, T., Huntemann, M., Anderson, I., Billis, K., Varghese, N., Mavromatis, K., Pati, A., Ivanova, N. N., and Kyrpides, N. C. (2014). IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Research*, 42(D1):D560–D567.
- Martin-Cuadrado, A.-B., Senel, E., Martínez-García, M., Cifuentes, A., Santos, F., Almansa, C., Moreno-Paz, M., Blanco, Y., García-Villadangos, M., del Cura, M. Á. G., Sanz-Montero, M. E., Rodríguez-Aranda, J. P., Rosselló-Móra, R., Antón, J., and Parro, V. (2019). Prokaryotic and viral community of the sulfate-rich crust from Peñahueca ephemeral lake, an astrobiology analogue. *Environmental Microbiology*, 21(10):3577–3600.
- Martinez-Romo, J., Araujo, L., Borge-Holthoefer, J., Arenas, A., Capitán, J. A., and Cuesta, J. A. (2011). Disentangling categorical relationships through a graph of co-occurrences. *Physical Review E*, 84(4):46108.
- Maurer, B. A. and Taper, M. L. (2002). Connecting geographical distributions with population processes. *Ecology Letters*, 5(2):223–231.
- Mayfield, M. M. and Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, 13(9):1085–1093.
- McCulloch, G. P., Irvine, K., Eckardt, F. D., and Bryant, R. (2008). Hydrochemical fluctuations and crustacean community composition in an ephemeral saline lake (Sua Pan, Makgadikgadi Botswana). *Hydrobiologia*, 596(1):31–46.
- McGill, B. (2010). Matters of Scale. *Science*, 328(5978):575–576.
- McGill, B. J., Enquist, B. J., Weiher, E., and Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21(4):178–185.
- Mees, F., Castañeda, C., and van Ranst, E. (2011). Sedimentary and diagenetic features in saline lake deposits of the Monegros region, northern Spain. *Catena*, 85(3):245–252.
- Menéndez-Serra, M., Ontiveros, V., Triadó-Margarit, X., Alonso, D., and Casamayor, E. (2020). Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain). *FEMS Microbiology Ecology*, 96(3):fiae019.
- Menéndez-Serra, M., Triadó-Margarit, X., and Casamayor, E. O. (2021). Ecological and metabolic thresholds in the bacterial, protist, and fungal microbiome of ephemeral saline lakes (Monegros Desert, Spain). *Microbial Ecology*, 82:885–896.
- Menéndez-Serra, M., Triadó-Margarit, X., Castañeda, C., Herrero, J., and

- Casamayor, E. O. (2019). Microbial composition, potential functional roles and genetic novelty in gypsum-rich and hypersaline soils of Monegros and Gallo-canta (Spain). *Science of The Total Environment*, 650:343–353.
- Meyer, K. M., Memiaghe, H., Korte, L., Kenfack, D., Alonso, A., and Bohannan, B. J. M. (2018). Why do microbes exhibit weak biogeographic patterns? *The ISME Journal*, 12(6):1404–1413.
- Milici, M., Deng, Z.-L., Tomasch, J., Decelle, J., Wos-Oxley, M. L., Wang, H., Jáuregui, R., Plumeier, I., Giebel, H.-A., Badewien, T. H., Wurst, M., Pieper, D. H., Simon, M., and Wagner-Döbler, I. (2016). Co-occurrence analysis of microbial taxa in the Atlantic Ocean reveals high connectivity in the free-living bacterioplankton.
- Mo, Y., Peng, F., Gao, X., Xiao, P., Logares, R., Jeppesen, E., Ren, K., Xue, Y., and Yang, J. (2021). Low shifts in salinity determined assembly processes and network stability of microeukaryotic plankton communities in a subtropical urban reservoir. *Microbiome*, 9(1):128.
- Monard, C., Gantner, S., Bertilsson, S., Hallin, S., and Stenlid, J. (2016). Habitat generalists and specialists in microbial communities across a terrestrial-freshwater gradient. *Scientific Reports*, 6(1):37719.
- Monger, C., Sala, O. E., Duniway, M. C., Goldfus, H., Meir, I. A., Poch, R. M., Throop, H. L., and Vivoni, E. R. (2015). Legacy effects in linked ecological-soil-geomorphic systems of drylands. *Frontiers in Ecology and the Environment*, 13(1):13–19.
- Monier, J.-M. and Lindow, S. E. (2005). Aggregates of resident bacteria facilitate survival of immigrant bacteria on leaf surfaces. *Microbial Ecology*, 49(3):343–352.
- Mony, C., Vandenkoornhuyse, P., Bohannan, B. J. M., Peay, K., and Leibold, M. A. (2020). A landscape of opportunities for microbial ecology research.
- Mor, J. A. C., Álamo, C. C., and Solanes, J. P. (2010). *Las saladas de Monegros y su entorno: hábitats y paisaje vegetal*. Publicaciones del Consejo de Protección de la Naturaleza de Aragón: Serie Conservación. Consejo de Protección de la Naturaleza de Aragón.
- Narasingarao, P., Podell, S., Ugalde, J. A., Brochier-Armanet, C., Emerson, J. B., Brocks, J. J., Heidelberg, K. B., Banfield, J. F., and Allen, E. E. (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *The ISME Journal*, 6(1):81–93.
- Navarro, J. B., Moser, D. P., Flores, A., Ross, C., Rosen, M. R., Dong, H., Zhang, G., and Hedlund, B. P. (2009). Bacterial succession within an ephemeral hypereutrophic Mojave Desert playa lake. *Microbial Ecology*, 57(2):307–320.
- Nawrocki, E. P., Jones, T. A., and Eddy, S. R. (2018). Group I introns are widespread in archaea. *Nucleic Acids Research*, pages gky414–gky414.

- Nekola, J. C. and White, P. S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26(4):867–878.
- Nemergut, D. R., Knelman, J. E., Ferrenberg, S., Bilinski, T., Melbourne, B., Jiang, L., Violle, C., Darcy, J. L., Prest, T., Schmidt, S. K., and Townsend, A. R. (2016). Decreases in average bacterial community rRNA operon copy number during succession. *The ISME journal*, 10(5):1147–1156.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167–256.
- Niku, J., Hui, F. K. C., Taskinen, S., and Warton, D. I. (2019). gllvm: Fast analysis of multivariate abundance data with generalized linear latent variable models in r. *Methods in Ecology and Evolution*, 10:2173–2182.
- Oftçerü, I. D., Lunn, M., Curtis, T. P., Wells, G. F., Criddle, C. S., Francis, C. A., and Sloan, W. T. (2010). Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences*, 107(35):15345–15350.
- Oh, D., Porter, K., Russ, B., Burns, D., and Dyll-Smith, M. (2010). Diversity of Haloquadratum and other haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. *Extremophiles*, 14(2):161–169.
- Okin, G. S., Gillette, D. A., and Herrick, J. E. (2006). Multi-scale controls on and consequences of aeolian processes in landscape change in arid and semi-arid environments. *Journal of Arid Environments*, 65(2):253–275.
- Oksanen, J., Kindt, R., Legendre, P., O ’ Hara, B., Henry, M., and Stevens, H. (2016). *vegan Community Ecology Package*.
- Ontiveros, V. J., Capitán, J. A., Arthur, R., Casamayor, E. O., and Alonso, D. (2019). Colonization and extinction rates estimated from temporal dynamics of ecological communities: The island r package. *Methods in Ecology and Evolution*, 10(7):1108–1117.
- Ontiveros, V. J., Capitán, J. A., Casamayor, E. O., and Alonso, D. (2021). The characteristic time of ecological communities. *Ecology*, 102(2):e03247.
- Oren, A. (1990). Thymidine incorporation in saltern ponds of different salinities: Estimation of in situ growth rates of halophilic archaeobacteria and eubacteria. *Microbial Ecology*, 19(1):43–51.
- Oren, A. (1999). Bioenergetic aspects of halophilism. *Microbiology and Molecular Biology Reviews*, 63(2):334–348.
- Oren, A. (2002). Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. *Journal of Industrial Microbiology and Biotechnology*, 28(1):56–63.
- Oren, A. (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline systems*, 4:2.

- Oren, A. (2011). Thermodynamic limits to microbial life at high salt concentrations. *Environmental Microbiology*, 13(8):1908–1923.
- Oren, A. (2013). Life at high salt concentrations. *The Prokaryotes: Prokaryotic Communities and Ecophysiology*, pages 421–440. Springer, Berlin, Heidelberg.
- Oren, A. (2014). The ecology of *Dunaliella* in high-salt environments. *Journal of Biological Research*, 21(1):1–8.
- Oren, A. (2015). Halophilic microbial communities and their environments. *Current Opinion in Biotechnology*, 33:119–124.
- Ortiz-Álvarez, R., Cáliz, J., Camarero, L., and Casamayor, E. O. (2020). Regional community assembly drivers and microbial environmental sources shaping bacterioplankton in an alpine lacustrine district (Pyrenees, Spain). *Environmental Microbiology*, 22(1):297–309.
- Ortiz-Alvarez, R. and Casamayor, E. O. (2016). High occurrence of Pacearchaeota and Woesearchaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environmental Microbiology Reports*, 8(2):210–217.
- Ortiz-Álvarez, R., Fierer, N., de Los Ríos, A., Casamayor, E. O., and Barberán, A. (2018a). Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession. *The ISME journal*, 12(7):1658–1667.
- Ortiz-Álvarez, R., Triadó-Margarit, X., Camarero, L., Casamayor, E. O., and Catalan, J. (2018b). High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms. *Scientific Reports*, 8(1):1–11.
- Pachiadaki, M. G., Kallionaki, A., Dählmann, A., De Lange, G. J., and Kormas, K. A. (2011). Diversity and spatial distribution of prokaryotic communities along a sediment vertical profile of a deep-sea mud volcano. *Microbial ecology*, 62(3):655–68.
- Paerl, H. W. and Pinckney, J. L. (1996). A mini-review of microbial consortia: Their roles in aquatic production and biogeochemical cycling. *Microbial Ecology*, 31(3):225–247.
- Palmer, M. W. (1994). Variation in species richness: Towards a unification of hypotheses. *Folia Geobotanica et Phytotaxonomica*, 29(4):511.
- Pandit, S. N., Kolasa, J., and Cottenie, K. (2009). Contrasts between habitat generalists and specialists: an empirical extension to the basic metacommunity framework. *Ecology*, 90(8):2253–2262.
- Patil, R. S. (2019). *Comparative Analysis of Denoising and Clustering Methods in Microbiome Analysis*. PhD thesis, Ann Arbor.
- Pattnaik, P., Mishra, S. R., Bharati, K., Mohanty, S. R., Sethunathan, N., and

- Adhya, T. K. (2000). Influence of salinity on methanogenesis and associated microflora in tropical rice soils. *Microbiological Research*, 155(3):215–220.
- Paul, S., Bag, S. K., Das, S., Harvill, E. T., and Dutta, C. (2008). Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biology*, 9(4):R70.
- Pedrocchi, C. (1998). *Ecología de los Monegros*. Instituto de Estudios Altoaragoneses, 430pp, Huesca.
- Pedrós-Alió, C., Calderón-Paz, J. I., MacLean, M. H., Medina, G., Marrasé, C., Gasol, J. M., and Guixa-Boixereu, N. (2000). The microbial food web along salinity gradients. *FEMS Microbiology Ecology*, 32(2):143–155.
- Peres-Neto, P. R., Dray, S., and ter Braak, C. J. F. (2017). Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. *Ecography*, 40(7):806–816.
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R. P., Martínez-Meyer, E., Nakamura, M., and Araújo, M. B. (2011). *Ecological niches and geographic distributions*. Princeton University Press.
- Piccardi, P., Vessman, B., and Mitri, S. (2019). Toxicity drives facilitation between 4 bacterial species. *Proceedings of the National Academy of Sciences of the United States of America*, 116(32):15979–15984.
- Plemenitas, A., Vaupotic, T., Lenassi, M., Kogej, T., and Gunde-Cimerman, N. (2008). Adaptation of extremely halotolerant black yeast *Hortaea werneckii* to increased osmolarity: a molecular perspective at a glance. *Studies in mycology*, 61:67–75.
- Podell, S., Emerson, J. B., Jones, C. M., Ugalde, J. a., Welch, S., Heidelberg, K. B., Banfield, J. F., and Allen, E. E. (2014). Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. *The ISME journal*, 8(5):979–90.
- Poisot, T., Canard, E., Mouquet, N., and Hochberg, M. E. (2012). A comparative study of ecological specialization estimators. *Methods in Ecology and Evolution*, 3(3):537–544.
- Por, F. D. (1972). Hydrobiological notes on the high-salinity waters of the Sinai Peninsula. *Marine Biology*, 14(2):111–119.
- Preston, F. W. (1960). Time and space and the variation of species. *Ecology*, 41(4):611–627.
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLOS ONE*, 5(3):e9490.
- Price, P. W. (1977). General Concepts on the Evolutionary Biology of Parasites. *Evolution*, 31(2):405–420.
- Prosser, J. I., Bohannon, B. J. M., Curtis, T. P., Ellis, R. J., Firestone, M. K.,

- Freckleton, R. P., Green, J. L., Green, L. E., Killham, K., Lennon, J. J., Osborn, A. M., Solan, M., van der Gast, C. J., and Young, J. P. W. (2007). The role of ecological theory in microbial ecology. *Nature Reviews Microbiology*, 5(5):384–392.
- Pruesse, E., Peplies, J., and Glöckner, F. O. (2012). SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics*, 28(14):1823–1829.
- Pueyo Mur, J. J. and Inglés Urpinell, M. (1987). Magnesite formation in recent playa lakes, Los Monegros, Spain. *Geological Society, London, Special Publications*, 36(1):119–122.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, 41:D590–6.
- Ramin, K. I. and Allison, S. D. (2019). Bacterial tradeoffs in growth rate and extracellular enzymes. *Frontiers in microbiology*, 10:2956.
- Rath, K. M., Murphy, D. N., and Rousk, J. (2019). The microbial community size, structure, and process rates along natural gradients of soil salinity. *Soil Biology and Biochemistry*, 138:107607.
- Reche, I., Pulido-Villena, E., Morales-Baquero, R., and Casamayor, E. O. (2005). Does ecosystem size determine aquatic bacterial richness? *Ecology*, 86(7):1715–1722.
- Ricklefs, R. E. (1987). Community Diversity: Relative Roles of Local and Regional Processes. *Science*, 235(4785):167–171.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., Darling, A., Malfatti, S., Swan, B. K., Gies, E. a., Dodsworth, J. a., Hedlund, B. P., Tsiamis, G., Sievert, S. M., Liu, W. T., Eisen, J. a., Hallam, S. J., Kyrpides, N. C., Stepanauskas, R., Rubin, E. M., Hugenholtz, P., and Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, 499(7459):431–7.
- Roberts, M. F. (2005). Organic compatible solutes of halotolerant and halophilic microorganisms. *Saline systems*, 1:5.
- Robertson, C. E., Spear, J. R., Harris, J. K., and Pace, N. R. (2009). Diversity and stratification of archaea in a hypersaline microbial mat. *Applied and environmental microbiology*, 75(7):1801–10.
- Rodriguez-Valera, F. (1993). Introduction to saline environments. In *The Biology of Halophilic Bacteria*, page 23. CRC Press.
- Rodriguez-Valera, F., Ventosa, A., Juez, G., and Imhoff, J. F. (1985). Variation of environmental features and microbial populations with salt concentrations in a

- multi-pond saltern. *Microbial Ecology*, 11(2):107–115.
- Roeßler, M. and Müller, V. (2001). Osmoadaptation in bacteria and archaea: common principles and differences. *Environmental Microbiology*, 3(12):743–754.
- Rose, R. W., Brüser, T., Kissinger, J. C., and Pohlschröder, M. (2002). Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Molecular Microbiology*, 45(4):943–950.
- Rosen, M. R. (1994). The importance of groundwater in playas: A review of playa classifications and the sedimentology and hydrology of playas. In *Paleoclimate and Basin Evolution of Playa Systems*. Geological Society of America.
- Rosindell, J., Hubbell, S. P., He, F., Harmon, L. J., and Etienne, R. S. (2012). The case for ecological neutral theory. *Trends in Ecology & Evolution*, 27(4):203–208.
- Rosselló-Mora, R. and Amann, R. (2001). The species concept for prokaryotes. *FEMS Microbiology Reviews*, 25(1):39–67.
- Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R., and Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal*, 4(10):1340–1351.
- Ruiz-González, C., Niño-García, J. P., and del Giorgio, P. A. (2015). Terrestrial origin of bacterial communities in complex boreal freshwater networks. *Ecology Letters*, 18(11):1198–1206.
- Ruiz-González, C., Niño-García, J. P., Kembel, S. W., and del Giorgio, P. A. (2017). Identifying the core seed bank of a complex boreal bacterial metacommunity. *The ISME Journal*, 11(9):2012–2021.
- Saccò, M., White, N. E., Harrod, C., Salazar, G., Aguilar, P., Cubillos, C. F., Meredith, K., Baxter, B. K., Oren, A., Anufrieva, E., Shadrin, N., Marambio-Alfaro, Y., Bravo-Naranjo, V., and Allentoft, M. E. (2021). Salt to conserve: a review on the ecology and preservation of hypersaline ecosystems. *Biological Reviews*, 96(6):2828–2850.
- Sala, O. E., Gherardi, L. A., Reichmann, L., Jobbágy, E., and Peters, D. (2012). Legacies of precipitation fluctuations on primary production: theory and data synthesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1606):3135–3144.
- Samper-Calvete, F. J. and García-Vera, M. A. (1998). Inverse modeling of groundwater flow in the semiarid evaporitic closed basin of Los Monegros, Spain. *Hydrogeology Journal*, 6(1):33–49.
- Schneider, D. C. (2001). The rise of the concept of scale in ecology: the concept of scale is evolving from verbal expression to quantitative expression. *BioScience*, 51(7):545–553.
- Scholz, B. and Liebezeit, G. (2012). Compatible solutes in three marine intertidal microphytobenthic Wadden Sea diatoms exposed to different salinities. *European*

- Journal of Phycology*, 47(4):393–407.
- Shade, A., Kent, A. D., Jones, S. E., Newton, R. J., Triplett, E. W., and McMahon, K. D. (2007). Interannual dynamics and phenology of bacterial communities in a eutrophic lake. *Limnology and Oceanography*, 52(2):487–494.
- Shmida, A. and Wilson, M. V. (1985). Biological determinants of species diversity. *Journal of Biogeography*, 12(1):1–20.
- Simberloff, D. S. (1969). Experimental zoogeography of islands: a model for insular colonization. *Ecology*, 50(2):296–314.
- Sinclair, L., Ijaz, U. Z., Jensen, L. J., Coolen, M. J. L., Gubry-Rangin, C., Chroňáková, A., Oulas, A., Pavloudi, C., Schnetzer, J., Weimann, A., Ijaz, A., Eiler, A., Quince, C., and Pafilis, E. (2016). Seqenv: linking sequences to environments through text mining. *PeerJ*, 4:e2690–e2690.
- Sisson, C., Gulla-Devaney, B., Katz, L. A., and Grattapanche, J.-D. (2018). Seed bank and seasonal patterns of the eukaryotic SAR (Stramenopila, Alveolata and Rhizaria) clade in a New England vernal pool. *Journal of Plankton Research*, 40(4):376–390.
- Šizling, A. L., Kunin, W. E., Šizlingová, E., Reif, J., and Storch, D. (2011). Between geometry and biology: the problem of universality of the species-area relationship. *The American Naturalist*, 178(5):602–611.
- Skirnisdottir, S., Hreggvidsson, G. O., Hjörleifsdottir, S., Marteinson, V. T., Petursdottir, S. K., Holst, O., and Kristjansson, J. K. (2000). Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats. *Applied and Environmental Microbiology*, 66(7):2835–2841.
- Sleator, R. D. and Hill, C. (2002). Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiology Reviews*, 26(1):49–71.
- Smith, T. W. and Lundholm, J. T. (2010). Variation partitioning as a tool to distinguish between niche and neutral processes. *Ecography*, 33(4):648–655.
- Soberon, J. M. and Martinez Del Rio, C. (1981). The dynamics of a plant-pollinator interaction. *Journal of Theoretical Biology*, 91(2):363–378.
- Sommers, P., Darcy, J. L., Porazinska, D. L., Gendron, E. M. S., Fountain, A. G., Zamora, F., Vincent, K., Cawley, K. M., Solon, A. J., Vimercati, L., Ryder, J., and Schmidt, S. K. (2019). Comparison of microbial communities in the sediments and water columns of frozen cryoconite holes in the McMurdo Dry Valleys, Antarctica.
- Spang, A., Caceres, E. F., and Ettema, T. J. (2017). Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science*, 357(6351).
- Stackebrandt, E. and Goebel, B. M. (1994). Taxonomic note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition

- in bacteriology.
- Steenwerth, K. L., Jackson, L. E., Calderón, F. J., Stromberg, M. R., and Scow, K. M. (2002). Soil microbial community composition and land use history in cultivated and grassland ecosystems of coastal California. *Soil Biology and Biochemistry*, 34(11):1599–1611.
- Stegen, J. C., Lin, X., Fredrickson, J. K., Chen, X., Kennedy, D. W., Murray, C. J., Rockhold, M. L., and Konopka, A. (2013). Quantifying community assembly processes and identifying features that impose them. *The ISME Journal*, 7(11):2069–2079.
- Stevenson, A., Cray, J. A., Williams, J. P., Santos, R., Sahay, R., Neuenkirchen, N., McClure, C. D., Grant, I. R., Houghton, J. D. R., Quinn, J. P., Timson, D. J., Patil, S. V., Singhal, R. S., Antón, J., Dijksterhuis, J., Hocking, A. D., Lievens, B., Rangel, D. E. N., Voytek, M. A., Gunde-Cimerman, N., Oren, A., Timmis, K. N., McGenity, T. J., and Hallsworth, J. E. (2015). Is there a common water-activity limit for the three domains of life? *The ISME Journal*, 9(6):1333–1351.
- Storch, D. and Šizling, A. L. (2008). The concept of taxon invariance in ecology: do diversity patterns vary with changes in taxonomic resolution? *Folia Geobotanica*, 43(3):329–344.
- Strona, G., Nappo, D., Boccacci, F., Fattorini, S., and San Miguel Ayanz, J. (2014). A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, 5(1):4114.
- Székely, A. J. and Langenheder, S. (2014). The importance of species sorting differs between habitat generalists and specialists in bacterial communities. *FEMS Microbiology Ecology*, 87(1):102–112.
- Székely, A. J. and Langenheder, S. (2017). Dispersal timing and drought history influence the response of bacterioplankton to drying–rewetting stress. *The ISME Journal*, 11(8):1764–1776.
- Telford, J. R., Vigdis, V., and Birks, H. J. B. (2006). Dispersal limitations matter for microbial morphospecies. *Science*, 312(5776):1015.
- Teng, W., Kuang, J., Luo, Z., and Shu, W. (2017). Microbial diversity and community assembly across environmental gradients in acid mine drainage. *Minerals*, 7(6).
- ter Braak, C. J. F., Cormont, A., and Dray, S. (2012). Improved testing of species traits–environment relationships in the fourth-corner problem. *Ecology*, 93(7):1525–1526.
- Thakur, M. P. and Wright, A. J. (2017). Environmental filtering, niche construction, and trait variability: the missing discussion. *Trends in Ecology & Evolution*, 12(32):884–886.

- 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., Zech Xu, Z., Jiang, L., Haroon, M. F., Kanbar, J., Zhu, Q., Jin Song, S., 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., and TheEarthMicrobiome-ProjectConsortium (2017). A communal catalogue reveals Earth’s multiscale microbial diversity. *Nature*, 551:457–463.
- Tilman, D. (1982). *Resource Competition and Community Structure*. (MPB-17), Volume 17. Princeton University Press.
- Triadó-Margarit, X., Capitán, J. A., Menéndez-Serra, M., Ortiz-Álvarez, R., Ontiveros, V. J., Casamayor, E. O., and Alonso, D. (2019). A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly. *The ISME Journal*.
- Triadó-Margarit, X. and Casamayor, E. O. (2013). High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. *FEMS Microbiology Ecology*, 85(1):27–36.
- Tripathi, B. M., Stegen, J. C., Kim, M., Dong, K., Adams, J. M., and Lee, Y. K. (2018). Soil pH mediates the balance between stochastic and deterministic assembly of bacteria. *The ISME Journal*, 12(4):1072–1083.
- Trüper, H. G. and Galinski, E. A. (1990). Biosynthesis and fate of compatible solutes in extremely halophilic phototrophic eubacteria. *FEMS Microbiology Reviews*, 6(2-3):247–254.
- Van Der Gast, C. J., Ager, D., and Lilley, A. K. (2008). Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. *Environmental Microbiology*, 10(6):1411–1418.
- Van der Gucht, K., Cottenie, K., Muylaert, K., Vloemans, N., Cousin, S., Declerck, S., Jeppesen, E., Conde-Porcuna, J. M., Schwenk, K., Zwart, G., Degans, H., Vyverman, W., and De Meester, L. (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proceedings of the National Academy of Sciences*, 104(51):20404–20409.
- van der Plas, F., Janzen, T., Ordonez, A., Fokkema, W., Reinders, J., Etienne, R. S., and Olff, H. (2015). A new modeling approach estimates the relative importance of different community assembly processes. *Ecology*, 96(6):1502–1515.
- van der Ploeg, R., Mäder, U., Homuth, G., Schaffer, M., Denham, E. L., Monteferrante, C. G., Miethke, M., Marahiel, M. A., Harwood, C. R., Winter, T., Hecker, M., Antelmann, H., and van Dijk, J. M. (2011). Environmental salinity

- determines the specificity and need for tat-pepdependent secretion of the YwbN protein in *Bacillus subtilis*. *PLOS ONE*, 6(3):e18140.
- van der Ploeg, R., Monteferrante, C. G., Piersma, S., Barnett, J. P., Kouwen, T. R. H. M., Robinson, C., and van Dijl, J. M. (2012). High-salinity growth conditions promote tat-independent secretion of tat substrates in *Bacillus subtilis*. *Applied and Environmental Microbiology*, 78(21):7733–7744.
- Vass, M. and Langenheder, S. (2017). The legacy of the past: effects of historical processes on microbial metacommunities . *Aquatic Microbial Ecology*, 79(1):13–19.
- Veech, J. A. (2013). A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, 22(2):252–260.
- Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly review of biology*, 85(2):183–206.
- Vellend, M. (2016). *The theory of ecological communities*. Princeton University Press.
- Ventosa, A., de la Haba, R. R., Sánchez-Porro, C., and Papke, R. T. (2015). Microbial diversity of hypersaline environments: A metagenomic approach. *Current Opinion in Microbiology*, 25:80–87.
- Ventosa, A., Fernández, A. B., León, M. J., Sánchez-Porro, C., and Rodriguez-Valera, F. (2014a). The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles*, 18(5):811–824.
- Ventosa, A., Fernández, A. B., León, M. J., Sánchez-Porro, C., and Rodriguez-Valera, F. (2014b). The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles*, 18(5):811–824.
- Ventosa, A., Rodriguez-Valera, F., Poindexter, J. S., and Reznikoff, W. S. (1984). Selection for moderately halophilic bacteria by gradual salinity increases. *Canadian Journal of Microbiology*, 30(10):1279–1282.
- Viana, D. S. and Chase, J. M. (2019). Spatial scale modulates the inference of metacommunity assembly processes. *Ecology*, 100(2):e02576.
- von Heijne, G. (1985). Signal sequences: The limits of variation. *Journal of Molecular Biology*, 184(1):99–105.
- Veeland, R. H., Litchfield, C. D., Martin, E. L., and Elliot, E. (1980). *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. *International Journal of Systematic Bacteriology*, 30(2):485–495.
- Waiser, M. and Robarts, R. (2009). Saline inland waters. In Likens, G. E., editor, *Encyclopedia of Inland Waters*, pages 634–644. Academic Press, Oxford.
- Wang, S., Hou, W., Dong, H., Jiang, H., Huang, L., Wu, G., Zhang, C., Song, Z., Zhang, Y., Ren, H., Zhang, J., and Zhang, L. (2013). Control of temperature on microbial community structure in hot springs of the Tibetan Plateau. *PLOS*

- ONE*, 8(5):e62901.
- Wang, Y., Li, C., Kou, Y., Wang, J., Tu, B., Li, H., Li, X., Wang, C., and Yao, M. (2017). Soil pH is a major driver of soil diazotrophic community assembly in Qinghai-Tibet alpine meadows. *Soil Biology and Biochemistry*, 115:547–555.
- Wattam, A. R., Abraham, D., Dalay, O., Disz, T. L., Driscoll, T., Gabbard, J. L., Gillespie, J. J., Gough, R., Hix, D., Kenyon, R., Machi, D., Mao, C., Nordberg, E. K., Olson, R., Overbeek, R., Pusch, G. D., Shukla, M., Schulman, J., Stevens, R. L., Sullivan, D. E., Vonstein, V., Warren, A., Will, R., Wilson, M. J. C., Yoo, H. S., Zhang, C., Zhang, Y., and Sobral, B. W. (2014). PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, 42(D1):D581–D591.
- Webb, C. O., Ackerly, D. D., McPeck, M. A., and Donoghue, M. J. (2002). Phylogenies and Community Ecology. *Annual Review of Ecology and Systematics*, 33(1):475–505.
- Weinisch, L., Kühner, S., Roth, R., Grimm, M., Roth, T., Netz, D. J. A., Pierik, A. J., and Filker, S. (2018). Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate *Schmidingerothrix salinarum*. *PLOS Biology*, 16(1):e2003892.
- Wennekes, P. L., Rosindell, J., and Etienne, R. S. (2012). The neutral—niche debate: a philosophical perspective. *Acta Biotheoretica*, 60(3):257–271.
- Whitaker, R. J., Grogan, D. W., and Taylor, J. W. (2003). Geographic barriers isolate endemic populations of hyperthermophilic Archaea. *Science*, 301(5635):976–978.
- White, E. R. and Hastings, A. (2020). Seasonality in ecology: Progress and prospects in theory. *Ecological Complexity*, 44:100867.
- Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York.
- Wiggins, G., Mackay, R., and Smith, I. (1980). Evolutionary and ecological strategies of animals in temporary pools. *Archiv fur Hydrobiologie*.
- Williams, T. J., Allen, M. A., DeMaere, M. Z., Kyrpides, N. C., Tringe, S. G., Woyke, T., and Cavicchioli, R. (2014). Microbial ecology of an Antarctic hypersaline lake: genomic assessment of ecophysiology among dominant haloarchaea. *The ISME Journal*, 8(8):1645–1658.
- Williams, W. D. (1996). The largest, highest and lowest lakes of the world: Saline lakes. *SIL Proceedings, 1922-2010*, 26(1):61–79.
- Wisn, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., Dormann, C. F., Forchhammer, M. C., Grytnes, J. A., Guisan, A., Heikkinen, R. K., Høye, T. T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.-C., Normand, S., Öckinger, E., Schmidt, N. M., Termansen, M., Timmermann, A., Wardle,

- D. A., Aastrup, P., and Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, 88(1):15–30.
- Woese, C. R. and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proceedings of the National Academy of Sciences*, 74(11):5088–5090.
- Wood, C. M. (2011). Rapid regulation of Na^+ and Cl^- flux rates in killifish after acute salinity challenge. *Journal of Experimental Marine Biology and Ecology*, 409(1):62–69.
- Wu, L., Ning, D., Zhang, B., Li, Y., Zhang, P., Shan, X., Zhang, Q., Brown, M., Li, Z., Van Nostrand, J. D., Ling, F., Xiao, N., Zhang, Y., Vierheilig, J., Wells, G. F., Yang, Y., Deng, Y., Tu, Q., Wang, A., Acevedo, D., Agullo-Barcelo, M., Alvarez, P. J. J., Alvarez-Cohen, L., Andersen, G. L., de Araujo, J. C., Boehnke, K., Bond, P., Bott, C. B., Bovio, P., Brewster, R. K., Bux, F., Cabezas, A., Cabrol, L., Chen, S., Criddle, C. S., Deng, Y., Etchebehere, C., Ford, A., Frigon, D., Gómez, J. S., Griffin, J. S., Gu, A. Z., Habagil, M., Hale, L., Hardeman, S. D., Harmon, M., Horn, H., Hu, Z., Jauffur, S., Johnson, D. R., Keller, J., Keucken, A., Kumari, S., Leal, C. D., Lebrun, L. A., Lee, J., Lee, M., Lee, Z. M. P., Li, Y., Li, Z., Li, M., Li, X., Ling, F., Liu, Y., Luthy, R. G., Mendonça-Hagler, L. C., de Menezes, F. G. R., Meyers, A. J., Mohebbi, A., Nielsen, P. H., Ning, D., Oehmen, A., Palmer, A., Parameswaran, P., Park, J., Patsch, D., Reginatto, V., de los Reyes, F. L., Rittmann, B. E., Robles, A. N., Rossetti, S., Shan, X., Sidhu, J., Sloan, W. T., Smith, K., de Sousa, O. V., Stahl, D. A., Stephens, K., Tian, R., Tiedje, J. M., Tooker, N. B., Tu, Q., Van Nostrand, J. D., De los Cobos Vasconcelos, D., Vierheilig, J., Wagner, M., Wakelin, S., Wang, A., Wang, B., Weaver, J. E., Wells, G. F., West, S., Wilmes, P., Woo, S.-G., Wu, L., Wu, J.-H., Wu, L., Xi, C., Xiao, N., Xu, M., Yan, T., Yang, Y., Yang, M., Young, M., Yue, H., Zhang, B., Zhang, P., Zhang, Q., Zhang, Y., Zhang, T., Zhang, Q., Zhang, W., Zhang, Y., Zhou, H., Zhou, J., Wen, X., Curtis, T. P., He, Q., He, Z., Brown, M., Zhang, T., He, Z., Keller, J., Nielsen, P. H., Alvarez, P. J. J., Criddle, C. S., Wagner, M., Tiedje, J. M., He, Q., Curtis, T. P., Stahl, D. A., Alvarez-Cohen, L., Rittmann, B. E., Wen, X., Zhou, J., and Consortium, G. W. M. (2019). Global diversity and biogeography of bacterial communities in wastewater treatment plants. *Nature Microbiology*, 4:1183–1195.
- Wu, Q. L., Zwart, G., Schauer, M., van Agterveld, M. P. K., and Hahn, M. W. (2006). Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the tibetan plateau, china. *Applied and Environmental Microbiology*, 72(8):5478–5485.
- Wu, W., Lu, H. P., Sastri, A., Yeh, Y. C., Gong, G. C., Chou, W. C., and Hsieh,

- C. H. (2018). Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME Journal*, 12(2):485–494.
- Xu, J., Gao, W., Zhao, B., Chen, M., Ma, L., Jia, Z., and Zhang, J. (2021). Bacterial community composition and assembly along a natural sodicity/salinity gradient in surface and subsurface soils. *Applied Soil Ecology*, 157:103731.
- Yechieli, Y. and Wood, W. W. (2002). Hydrogeologic processes in saline systems: playas, sabkhas, and saline lakes. *Earth-Science Reviews*, 58(3):343–365.
- Yeoh, Y. K., Sekiguchi, Y., Parks, D. H., and Hugenholtz, P. (2016). Comparative genomics of candidate phylum TM6 suggests that parasitism is widespread and ancestral in this lineage. *Molecular Biology and Evolution*, 33(4):915–927.
- Zajc, J., Kogej, T., Galinski, E. A., Ramos, J., and Gunde-Cimerman, N. (2014). Osmoadaptation strategy of the most halophilic fungus, *Wallemia ichthyophaga*, growing optimally at salinities above 15% NaCl. *Applied and Environmental Microbiology*, 80(1):247–256.
- Zeileis, A., Kleiber, C., Krämer, W., and Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44(1):109–123.
- Zhalnina, K., Dias, R., de Quadros, P. D., Davis-Richardson, A., Camargo, F. A. O., Clark, I. M., McGrath, S. P., Hirsch, P. R., and Triplett, E. W. (2015). Soil pH determines microbial diversity and composition in the Park Grass Experiment. *Microbial Ecology*, 69(2):395–406.
- Zhang, C., Li, J., Wang, J., Liu, G., Wang, G., Guo, L., and Peng, S. (2019). Decreased temporary turnover of bacterial communities along soil depth gradient during a 35-year grazing exclusion period in a semiarid grassland. *Geoderma*, 351:49–58.

Chapter 2 Supporting Information

A.1 Supplementary Tables

Table A.1: Detailed information for each sample consisting on sampling date, sampling site, water height (cm), salinity (% w/v), conductivity (mS/cm), water temperature (°C), pH, redox conditions (mV), dissolved oxygen (mg/l) and Chla ($\mu\text{g/l}$).

Code	Date	Lagoon	Height	Salinity	Cond.	Temp.	pH	Redox	D.O.	Chla
MON_002	Piñol	12/11/12	13	1.2	16	13.5	8.6	315	11.3	5.4
MON_004	Rebollón	12/11/12	5	1	7.8	14.3	8.2	340	10.4	1.3
MON_005	Rollico	12/11/12	10	1.1	11.6	12.5	8.6	345	10.5	1.7
MON_009	A. Alta	13/11/12	10	1	4.7	8.6	7.9	326	11.8	0.5
MON_008	A. Baja	13/11/12	5	0.9	4.6	6.3	8.2	339	12	1.2
MON_011	Camarón	13/11/12	40	1.2	14.2	11.7	8.3	305	12.2	0.5
MON_010	Pito	13/11/12	15	1.5	18.3	13.9	7.8	330	11.7	6.2
MON_012	Playa	13/11/12	10	2.1	26.3	14.4	8.8	306	12	1.2
MON_013	Salineta	13/11/12	13	6.9	66.3	13.1	9.2	278	10.6	79.8
MON_007	Salobral	13/11/12	20	3.3	31.7	5.9	9.0	341	9.4	2.9
MON_014	Guallar	4/12/12	10	4	40	10.5	8.1	350	10.6	3.7
MON_017	Muerte	4/12/12	2	3.4	37.8	12	8.1	312	10.3	5.2
MON_020	Pez	4/12/12	4	2.1	21.7	8.1	8.0	287	11.1	4.2
MON_015	Piñol	4/12/12	10	2	21.6	12	8.3	305	10.6	8.4
MON_019	Rebollón	4/12/12	5	1.8	15.7	9.1	8.1	280	10.9	1.3
MON_018	Rollico	4/12/12	10	2	20.4	10	8.1	318	10.3	6.7
MON_022	A. Alta	5/12/12	5	1.2	8.1	7.8	7.9	290	11.3	0.6
MON_021	A. Baja	5/12/12	2	1.2	7.8	5.9	7.9	242	11.9	0.7
MON_023	Pito	5/12/12	4	2.3	25	9.2	8.0	302	10.7	2.3
MON_024	Playa	5/12/12	15	3	31	8.5	8.2	303	11.5	0.4
MON_025	Salineta	5/12/12	25	7.8	77.1	8.2	9.0	286	12.1	237.4
MON_026	Salobral	5/12/12	30	3.2	31.3	8.0	8.8	275	13.5	1.7
MON_030	Camarón	9/1/13	40	2.5	27	7.0	8.3	250	12.6	6.0
MON_027	Guallar	9/1/13	8	6	65.3	4.5	8.4	225	12.8	0.9
MON_029	Muerte	9/1/13	5	5.8	65.6	7.8	8.2	260	12.3	0.9
MON_033	Pez	9/1/13	3	3.3	38	8.3	8.2	280	12.7	3.6
MON_028	Piñol	9/1/13	10	4	32.5	5.5	8.7	206	13.3	1.4
MON_032	Rebollón	9/1/13	3	2.4	27.2	9.5	8.8	245	14.6	1.0
MON_031	Rollico	9/1/13	5	3.3	37.8	9.5	8.4	238	12.2	0.7
MON_035	A. Alta	10/1/13	5	1.7	15.6	10	8.0	285	11.1	3.0
MON_034	A. Baja	10/1/13	3	1.8	15.6	9.5	8.0	256	11.6	0.8
MON_036	Pito	10/1/13	10	3	36.5	12	8.0	295	10.5	4.2
MON_037	Playa	10/1/13	30	4.5	49.7	11.3	8.4	270	11.6	0.5
MON_038	Salineta	10/1/13	50	10	94.3	12	8.7	275	11.8	71.3
MON_039	Salobral	10/1/13	40	3.2	32	11	8.9	280	13.8	2.9
MON_040	Guallar	22/2/13	10	10.5	118.9	9.4	8.2	244	10.6	1.6

Table A.1 Continued from previous page

Code	Date	Lagoon	Height	Salinity	Cond.	Temp.	pH	Redox	D.O.	Chla
MON_041	Playa	22/2/13	40	5.9	67	11	8.5	230	11.3	3.3
MON_042	Salineta	22/2/13	30	15	128.8	9.4	8.4	246	10.4	35.8
MON_043	Salobral	22/2/13	30	3.7	35.6	10.1	8.9	210	11.5	3.8
MON_046	Camarón	23/2/13	30	3.1	34.3	2.3	8.0	288	12.4	1.8
MON_045	Pez	23/2/13	10	2.3	26.6	2.0	8.0	285	12.5	2.1
MON_044	Pito	23/2/13	10	4.8	53.4	1.5	7.9	293	12.7	3.8
MON_050	Camarón	20/3/13	20	3.2	39.8	15.3	8.4	288	10.3	5.7
MON_047	Guallar	20/3/13	10	6.8	80.6	13.6	8.2	245	10.3	5.9
MON_053	Pez	20/3/13	5	4	50.1	14	8.3	303	10.1	6.6
MON_048	Piñol	20/3/13	8	4	49.6	15	8.6	275	10.5	9.6
MON_052	Rebollón	20/3/13	5	2.2	26.2	14.8	9.1	258	10.5	13.2
MON_051	Rollico	20/3/13	5	3.2	40.3	15.4	8.2	245	10	13.7
MON_055	A. Alta	21/3/13	2	1	10.4	15	8.4	265	12.5	0.6
MON_054	A. Baja	21/3/13	2	1.5	13.3	11.2	8.6	258	11.8	0.9
MON_056	Pito	21/3/13	5	3.8	45.4	13.9	8.0	240	11	6.1
MON_057	Playa	21/3/13	30	5.8	67.5	16.3	8.6	250	12.4	5.9
MON_058	Salineta	21/3/13	30	11	107.5	16.1	8.4	270	10.1	40.9
MON_059	Salobral	21/3/13	50	3	31.3	16	8.9	230	10.8	15.5
MON_063	Camarón	16/4/13	10	4.3	68	29.5	9.0	220	12.3	4.2
MON_060	Guallar	16/4/13	3	16.9	166.4	23.3	7.9	270	8.8	2.9
MON_062	Muerte	16/4/13	1	32	207.4	32	7.3	290	7.4	2.1
MON_064	Pez	16/4/13	2	13.8	165.5	31	7.8	260	8.4	6.4
MON_061	Piñol	16/4/13	5	7.8	102.2	24.9	8.8	235	10.6	6.3
MON_065	Pito	17/4/13	3	7	94.1	22.1	8.1	325	9.7	6.6
MON_066	Playa	17/4/13	40	10.1	124.4	23.5	8.1	312	9.4	3.8
MON_067	Salineta	17/4/13	15	24	156.8	28	8.1	188	9.0	23.8
MON_068	Salobral	17/4/13	40	2.9	34.3	25.8	9.3	170	8.2	4.1
MON_071	Camarón	26/5/13	5	9.8	117.3	22.5	8.2	220	8.8	3.9
MON_069	Guallar	26/5/13	0.5	23	197.6	26.5	7.6	270	8.8	10
MON_077	Molinos	26/5/13	NA	0.1	0.9	17.4	8.3	195	9.8	3.0
MON_070	Piñol	26/5/13	1	8.4	103	25.1	8.6	220	8.9	27.9
MON_074	Playa	26/5/13	40	10.6	117.6	17.5	8.7	225	9.2	14
MON_075	Salineta	26/5/13	5	23	152	19.1	8.2	235	12.9	85.7
MON_076	Salobral	26/5/13	40	3.4	36.2	20	9.5	180	14.2	4.4
MON_073	Pez	27/5/13	2	6	79.1	21.5	8.1	200	9.1	6.7
MON_072	Rollico	27/5/13	2	5.8	77.6	21.5	8.1	205	9.1	2.8
MON_078	Pito	12/6/13	0.2	28	NA	34.5	7.5	260	7.5	1.8
MON_079	Playa	12/6/13	25	14.8	169.3	27.9	8.1	225	8.8	11.7
MON_080	Salineta	12/6/13	5	35	150.3	34	7.9	220	9.5	54.3
MON_081	Salobral	12/6/13	40	3	42.4	30.3	9.1	190	11.5	2.8
MON_083	Playa	8/7/13	30	16.2	168.1	30	8.2	230	7.3	4.0
MON_084	Salineta	8/7/13	1	38	142	32.5	7.9	230	7.9	5.6
MON_085	Salobral	8/7/13	30	3.8	47.1	29.6	8.8	122	8.1	5.4
MON_088	Salineta	9/10/13	1	40	143.3	18.8	7.7	265	8.9	2.4
MON_093	Camarón	27/11/13	15	2	22.7	10	8.3	270	10.9	9.5
MON_090	Guallar	27/11/13	10	2.8	26.2	2.0	8.1	250	12.1	1.1
MON_092	Muerte	27/11/13	5	2	20.3	7.5	8.3	270	11.2	1.4
MON_096	Pez	27/11/13	2	1.1	8.5	5.2	8.0	320	11.6	1.1
MON_091	Piñol	27/11/13	8	1.6	14.4	5.0	8.0	245	11.8	0.9
MON_095	Rebollón	27/11/13	15	1	5.1	7.5	8.1	305	11.6	1.8
MON_094	Rollico	27/11/13	5	0.8	5.6	9.0	8.1	305	11	1.9
MON_098	A. Alta	28/11/13	20	0.9	3.6	3.0	8.1	320	12.1	3.4
MON_097	A. Baja	28/11/13	4	0.9	4.3	1.5	9.2	280	11.7	5.5
MON_099	Pito	28/11/13	10	1.2	12	3.5	8.0	295	12.6	1.1
MON_100	Playa	28/11/13	40	5	58.1	4.0	8.0	280	11.7	0.8
MON_101	Salineta	28/11/13	10	7.5	67.5	7.4	NA	315	11.6	3.7

Table A.1 Continued from previous page

Code	Date	Lagoon	Height	Salinity	Cond.	Temp.	pH	Redox	D.O.	Chla
MON_106	Camarón	17/12/13	5	3.1	38.5	9.1	8.6	235	14.1	1.0
MON_113	Playa	18/12/13	25	6	78.3	8.5	8.4	144	12	8.0
MON_108	Rebollón	18/12/13	3	1.1	12.4	10	9.1	242	12.6	27.1
MON_107	Rollico	18/12/13	5	1	13.3	11.3	8.8	230	11.5	2.8
MON_114	Salineta	18/12/13	10	9	85.6	11	8.4	245	15.7	54.7
MON_111	A. Alta	19/12/13	5	0.9	5.3	7.5	8.2	275	10.9	6.8
MON_110	A. Baja	19/12/13	3	0.9	6.4	7.9	8.3	270	10.5	11.4
MON_112	Pito	19/12/13	5	1.7	20	6.7	8.3	275	11.6	1.4
MON_119	Camarón	22/1/14	3	4.6	59.2	11.4	8.3	270	10.8	4.3
MON_116	Guallar	22/1/14	6	5.8	67.4	8.5	8.4	265	11.6	1.7
MON_128	Molinos	22/1/14	40	0.3	0.7	8.0	9.1	240	12.1	9.7
MON_118	Muerte	22/1/14	3	5.9	71.1	9.7	8.3	275	11.3	2.9
MON_117	Piñol	22/1/14	10	4.1	49	9.0	8.3	272	11.8	5.4
MON_126	Playa	22/1/14	30	6.6	86.4	10.8	8.6	285	11.5	6.0
MON_124	A. Alta	23/1/14	8	1	8.7	8.5	8.0	248	11.8	0.9
MON_123	A. Baja	23/1/14	2	1.1	10.3	8.0	8.0	244	11.6	1.4
MON_122	Pez	23/1/14	6	2.9	35	11.5	8.0	300	10.5	4.3
MON_125	Pito	23/1/14	8	2.8	34.5	11.2	8.0	304	10.8	2.7
MON_120	Rollico	23/1/14	6	2.1	28.7	11.3	8.0	290	10.6	10.6
MON_127	Salineta	23/1/14	15	11.2	105.2	12.2	8.8	278	11.1	201.5
MON_129	Camarón	14/3/14	1	15.4	164.7	10	8.0	242	9.7	46
MON_130	Playa	14/3/14	25	11.5	134.9	10	8.4	215	9.9	29.6
MON_131	Salineta	14/3/14	10	24.2	167.9	10	8.0	235	10.7	95.5
MON_133	Playa	30/4/14	20	21	194.5	25	8.0	215	10.1	23.6
MON_134	Salineta	30/4/14	1	37	129.3	24	7.8	235	7.8	87.3
MON_139	Camarón	22/12/14	10	4.3	51.3	8.5	8.26	262	11.6	2.1
MON_136	Guallar	22/12/14	10	3.95	42	7.9	8.35	145	10.65	2.5
MON_138	Muerte	22/12/14	6	4	44.8	8.4	8.61	224	12.8	1.5
MON_142	Pez	22/12/14	5	1.6	17.68	8.1	8.4	279	13.1	3.2
MON_137	Piñol	22/12/14	8	2.9	31.4	8.1	8.39	187	11.9	2.3
MON_141	Rebollón	22/12/14	3	1.5	14.57	8.9	8.8	225	14.3	1.6
MON_140	Rollico	22/12/14	4	1.8	18.36	8.7	8.28	261	12.8	1.0
MON_148	A. Alta	23/12/14	25	1	6.95	6.3	8.25	299	11.1	1.2
MON_147	A. Baja	23/12/14	3	0.9	5.95	7.2	8.34	296	13.1	1.9
MON_146	Pito	23/12/14	15	2.8	30.7	6.1	8.2	335	11.7	1.5
MON_145	Playa	23/12/14	40	4.9	55.6	6.0	8.7	295	11.8	5.2

Table A.2: Accession numbers for the most abundant OTUs sequences uploaded to Genbank. Additional accession numbers for Halobacteria and Nanohaloarchaeota are shown in Fig S1.

Archaeal lineage	OTU id	Genbank accession number
Aenigmarchaeota	OTU_714	MH024548
Pacearchaeota	OTU_314	MH024528
	OTU_390	MH024534
	OTU_435	MH024536
	OTU_671	MH024546
	OTU_721	MH024549
	OTU_4358	MH024585
	OTU_9144	MH024587
Thermoplasmata	OTU_274	MH024525
	OTU_316	MH024526
	OTU_772	MH024547
Woesearchaeota	OTU_146	MH024522
	OTU_171	MH024524
	OTU_305	MH024530
	OTU_388	MH024533
	OTU_377	MH024535
	OTU_413	MH024537
	OTU_463	MH024538
	OTU_559	MH024541
	OTU_530	MH024543
	OTU_549	MH024544
	OTU_561	MH024545
	OTU_872	MH024551
	OTU_810	MH024552

A.2 Supplementary Figures

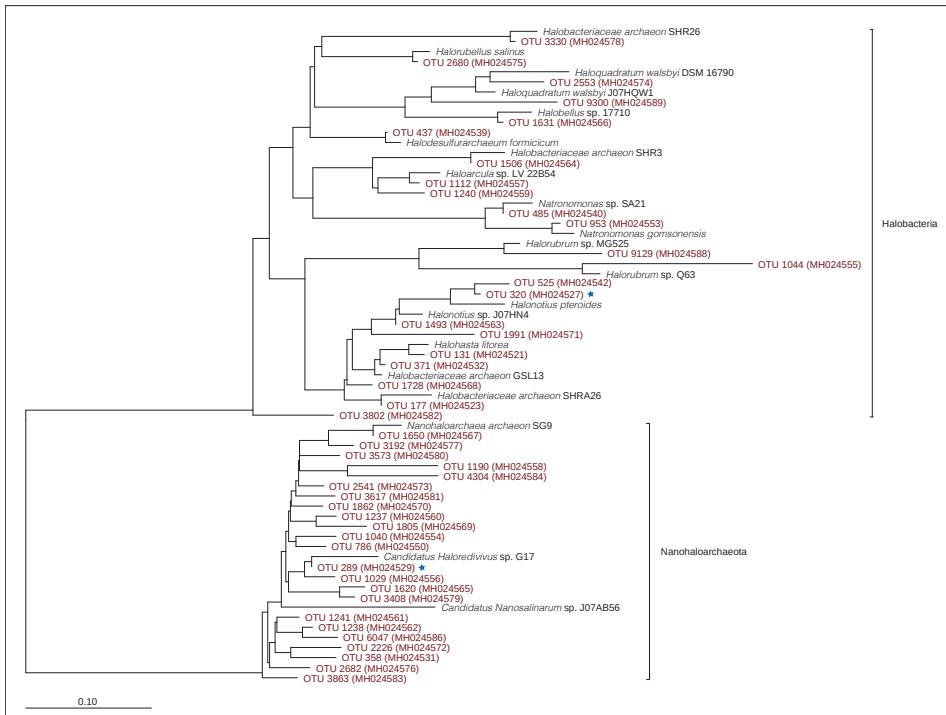


Figure A.1: ARB maximum parsimony tree provide by default with short 16S rRNA gene sequences added using the tool “add species to existed tree / ARB–parsimony tool interactive”. The most abundant OTUs ($>0.01\%$) of Halobacteria and Nanohaloarchaeota from the most hypersaline samples ($>35\%$ salinity) are included. Genbank accessions numbers are shown.

Chapter 3 Supporting Information

B.1 Supplementary Tables

Table B.1: Detailed information on sampling date (dd/mo/yr), sampling site, water height (cm), salinity (% w/v), conductivity (mS/cm), water temperature (°C), pH, redox conditions (mV), dissolved oxygen (mg/l), Chl a (mg/L) and DAPI cells density count (cel/mL) for each sample in Monegros. Community descriptive parameters are also included: bacterial richness and its relative abundance (%) in the 16S rRNA gene pool (including Archaea); total eukaryotic richness and individual richness and relative abundance (%) (in the 18S rRNA gene pool) of photosynthetic and heterotrophic protists and fungi.

Code	Site	Date	Data	H.	Sal.	Cond.	Temp.	pH	Redox	D.O.	Chla	DAPI	Bacteria	Photo. P.	Hetero. P.	Fungi	Euk.				
MON_1	Guallar	12/11/12	Euk.	13	2.9	29.5	9.5	8.2	290.4	11.5	9.3	–	–	33	34.44	66	64.63	19	0.93	118	
MON_14	Guallar	4/12/12	Both	10	4	40	10.5	8.1	350	10.6	3.7	1.72×10 ⁷	562	99.57	29	84.64	63	15.26	8	0.11	100
MON_27	Guallar	9/1/13	Both	8	6	65.3	4.5	8.4	225	12.8	0.9	1.64×10 ⁷	319	99.62	26	44.28	62	55.14	11	0.57	99
MON_40	Guallar	22/2/13	Both	10	10.5	118.9	9.4	8.2	244	10.6	1.6	1.54×10 ⁷	401	99.59	29	64.46	54	34.46	10	1.08	94
MON_47	Guallar	20/3/13	Both	10	6.8	80.6	13.6	8.2	245	10.3	5.9	1.99×10 ⁷	491	98.37	8	80.96	7	19.04	0	0.00	15
MON_60	Guallar	16/4/13	Both	3	16.9	166.4	23.3	7.9	270	8.8	2.9	1.39×10 ⁷	410	98.93	23	75.20	17	20.46	10	4.35	50
MON_69	Guallar	26/5/13	Both	0.5	23	197.6	26.5	7.6	270	8.8	10	1.53×10 ⁷	584	96.06	27	83.42	65	12.29	15	4.29	107
MON_90	Guallar	27/11/13	Both	10	2.8	26.2	2	8.1	250	12.1	1.1	1.24×10 ⁷	417	99.78	53	45.98	100	51.95	29	2.07	182
MON_103	Guallar	18/12/13	Euk.	5	3.5	44.1	12.8	8.6	230	13.4	7	1.91×10 ⁶	–	–	25	29.54	62	70.32	7	0.15	94
MON_116	Guallar	22/1/14	Both	6	5.8	67.4	8.5	8.4	265	11.6	1.7	1.51×10 ⁷	426	98.92	40	71.19	97	27.27	30	1.55	167
MON_136	Guallar	22/12/14	Both	10	3.95	42	7.9	8.35	145.5	10.65	2.5	9.86×10 ⁵	440	97.65	31	72.82	70	26.21	39	0.97	140
MON_2	Piñol	12/11/12	Both	13	1.2	16	13.5	8.6	315	11.3	5.4	2.78×10 ⁶	285	99.99	44	41.63	46	29.42	28	28.95	118
MON_15	Piñol	4/12/12	Both	10	2	21.6	12	8.3	305	10.6	8.4	2.30×10 ⁷	175	99.95	38	33.73	35	65.23	12	1.04	85
MON_28	Piñol	9/1/13	Both	10	4	32.5	5.5	8.7	206	13.3	1.4	1.68×10 ⁷	215	100	13	22.94	12	48.95	3	28.12	28
MON_48	Piñol	20/3/13	Both	8	4	49.6	15	8.6	275	10.5	9.6	7.71×10 ⁶	407	99.05	39	41.41	50	58.14	12	0.45	101
MON_61	Piñol	16/4/13	Both	5	7.8	102.2	24.9	8.8	235	10.6	6.3	1.22×10 ⁷	205	99.46	29	30.34	58	69.40	14	0.26	101
MON_70	Piñol	26/5/13	Both	1	8.4	103	25.1	8.6	220	8.9	27.9	–	327	99.08	27	8.94	38	91.03	1	0.03	66
MON_91	Piñol	27/11/13	Bact.	8	1.6	14.4	5	8	245	11.8	0.9	8.65×10 ⁶	523	99.73	–	–	–	–	–	–	–
MON_104	Piñol	17/12/13	Both	10	2.5	29.7	9	8.9	245	13.9	5.6	4.51×10 ⁶	345	99.90	38	50.84	48	48.81	13	0.35	99
MON_117	Piñol	22/1/14	Bact.	10	4.1	49	9	8.3	272	11.8	5.4	1.01×10 ⁷	557	98.79	–	–	–	–	–	–	–
MON_137	Piñol	22/12/14	Both	8	2.9	31.4	8.1	8.39	187.1	11.9	2.3	1.58×10 ⁶	393	98.69	28	13.79	45	85.49	29	0.72	102
MON_17	Muerte	4/12/12	Both	2	3.4	37.8	12	8.1	312	10.3	5.2	1.99×10 ⁷	402	99.22	23	94.69	32	5.19	4	0.13	59
MON_29	Muerte	9/1/13	Both	5	5.8	65.6	7.8	8.2	260	12.3	0.9	3.41×10 ⁶	417	96.91	41	51.51	89	42.33	40	6.16	170
MON_49	Muerte	20/3/13	Euk.	5	6.2	81.7	15.1	8.1	303	10.1	3.7	8.52×10 ⁶	–	–	46	55.05	85	41.04	27	3.91	158
MON_62	Muerte	16/4/13	Both	1	32	207.4	32	7.3	290	7.4	2.1	5.57×10 ⁶	555	97.39	29	27.39	34	71.54	7	1.07	70
MON_92	Muerte	27/11/13	Both	5	2	20.3	7.5	8.3	270	11.2	1.4	6.52×10 ⁶	405	99.90	63	56.10	116	39.70	41	4.20	220

Table B.1 Continued from previous page

Code	Site	Date	Data	H.	Sal.	Cond.	Temp.	pH	Redox	D.O.	Chla	DAPI	Bacteria	Photo. P.	Hetero. P.	Fungi	Euk.				
													Rich. Abund.	Rich. Abund.	Rich. Abund.	Rich. Abund.	Rich.				
MON_105	Muerte	17/12/13	Euk.	5	4.2	52.5	9.9	8.4	300	15.7	6.1	3.52x10 ⁶	—	23	38.49	64	61.31	13	0.19	100	
MON_118	Muerte	22/1/14	Both	3	5.9	71.1	9.7	8.3	275	11.3	2.9	8.42x10 ⁶	433	98.81	35	74.51	49	25.15	11	0.33	95
MON_138	Muerte	22/12/14	Both	6	4	44.8	8.4	8.61	224	12.8	1.5	3.65x10 ⁶	419	99.61	34	44.72	91	54.40	19	0.88	144
MON_3	Camaron	12/11/12	Euk.	25	1.1	14.1	14.9	8	360	10.3	2.5	—	—	39	70.90	94	22.03	20	51.13	153	
MON_11	Camaron	13/11/12	Both	40	1.2	14.2	11.7	8.3	305	12.2	0.5	3.19x10 ⁷	136	100	77	57.23	111	33.35	17	9.42	205
MON_16	Camaron	4/12/12	Euk.	30	1.8	19.1	12.4	8.4	296	10.2	19.9	—	—	37	85.25	75	14.32	10	0.43	122	
MON_30	Camaron	9/1/13	Both	40	2.5	27	7	8.3	250	12.6	6	2.25x10 ⁷	275	99.86	24	9.84	34	87.51	12	2.65	70
MON_46	Camaron	23/2/13	Bact.	30	3.1	34.3	2.3	8	288.5	12.4	1.8	6.97x10 ⁶	408	99.74	—	—	—	—	—	—	—
MON_50	Camaron	20/3/13	Both	20	3.2	39.8	15.3	8.4	288	10.3	5.7	1.11x10 ⁷	272	99.88	29	33.58	42	65.79	6	0.63	78
MON_63	Camaron	16/4/13	Both	10	4.3	68	29.5	9	220	12.3	4.2	8.61x10 ⁶	265	97.52	33	36.96	45	57.75	5	5.29	83
MON_71	Camaron	26/5/13	Both	5	9.8	117.3	22.5	8.2	220	8.8	3.9	7.91x10 ⁶	371	94.83	43	16.58	56	83.03	11	0.39	111
MON_93	Camaron	27/11/13	Bact.	15	2	22.7	10	8.3	270	10.9	9.5	9.79x10 ⁶	383	99.57	—	—	—	—	—	—	—
MON_106	Camaron	17/12/13	Both	5	3.1	38.5	9.1	8.6	235	14.1	1	6.00x10 ⁶	505	99.37	36	70.21	72	25.47	28	4.32	136
MON_119	Camaron	22/1/14	Both	3	4.6	59.2	11.4	8.3	270	10.8	4.3	7.91x10 ⁶	553	96.12	52	43.11	89	51.17	26	5.72	167
MON_129	Camaron	14/3/14	Both	1	15.4	164.7	10	8	242	9.7	46	3.16x10 ⁷	304	94.99	29	60.84	31	39.11	1	0.05	61
MON_139	Camaron	22/12/14	Both	10	4.3	51.3	8.5	8.26	262	11.6	2.1	4.98x10 ⁶	473	96.64	24	7.51	41	92.17	10	0.32	76
MON_6	Pez	12/11/12	Euk.	5	1.1	11.8	10.9	8.2	343.5	10.5	2.2	—	—	78	40.65	114	35.34	41	24.01	233	
MON_20	Pez	4/12/12	Both	4	2.1	21.7	8.1	8	287	11.1	4.2	2.90x10 ⁶	740	99.82	10	76.90	8	0.24	4	22.86	22
MON_33	Pez	9/1/13	Both	3	3.3	38	8.3	8.2	280	12.7	3.6	1.50x10 ⁷	334	99.97	38	43.76	68	54.34	22	1.90	129
MON_45	Pez	23/2/13	Both	10	2.3	26.6	2	8	285.5	12.5	2.1	8.10x10 ⁶	291	99.58	37	80.84	49	18.52	7	0.64	93
MON_53	Pez	20/3/13	Both	5	4	50.1	14	8.3	303.3	10.1	6.6	1.05x10 ⁷	576	98.32	35	61.25	65	36.89	24	1.86	125
MON_64	Pez	16/4/13	Both	2	13.8	165.5	31	7.8	260	8.4	6.4	5.98x10 ⁶	266	99.69	34	51.25	41	48.25	8	0.49	83
MON_73	Pez	27/5/13	Both	2	6	79.1	21.5	8.1	200	9.1	6.7	5.90x10 ⁶	457	99.60	48	15.70	95	82.67	27	1.63	171
MON_96	Pez	27/11/13	Both	2	1.1	8.5	5.2	8	320	11.6	1.1	6.94x10 ⁶	504	99.64	46	40.56	61	57.79	21	1.64	128
MON_109	Pez	18/12/13	Both	2	1.9	19.8	8.6	9	250	11.7	9.4	2.92x10 ⁶	247	99.98	38	15.17	66	84.29	12	0.54	116
MON_122	Pez	23/1/14	Both	6	2.9	35	11.5	8	300	10.5	4.3	9.45x10 ⁶	407	99.50	45	31.68	89	66.01	18	2.32	152
MON_142	Pez	22/12/14	Both	5	1.6	17.68	8.1	8.42	279	13.1	3.2	1.87x10 ⁶	540	99.69	65	46.58	107	51.46	38	1.96	211
MON_5	Rollico	12/11/12	Both	10	1.1	11.6	12.5	8.6	345	10.5	1.7	2.15x10 ⁶	567	99.97	79	36.81	131	36.95	49	26.24	260
MON_18	Rollico	4/12/12	Both	10	2	20.4	10	8.1	318	10.3	6.7	1.13x10 ⁷	602	99.73	72	55.32	107	42.85	34	1.83	213
MON_31	Rollico	9/1/13	Both	5	3.3	37.8	9.5	8.4	238	12.2	0.7	7.43x10 ⁶	350	99.83	47	57.59	76	38.30	28	4.11	152
MON_51	Rollico	20/3/13	Both	5	3.2	40.3	15.4	8.2	245	10	13.7	8.72x10 ⁶	545	99.85	44	40.08	85	56.84	23	3.08	152
MON_72	Rollico	27/5/13	Both	2	5.8	77.6	21.5	8.1	205	9.1	2.8	1.38x10 ⁷	378	98.31	38	40.07	67	59.62	9	0.30	114
MON_94	Rollico	27/11/13	Both	5	0.8	5.6	9	8.1	305	11	1.9	6.08x10 ⁶	569	99.53	56	60.57	130	32.53	54	6.90	241

Table B.1 Continued from previous page

Code	Site	Date	Data	H.	Sal.	Cond.	Temp.	pH	Redox	D.O.	Chla	DAPI	Bacteria Rich. Abund.	Photo. P. Rich. Abund.	Hetero. P. Rich. Abund.	Fungi Rich. Abund.	Euk. Rich.				
MON_107	Rollico	18/12/13	Both	5	1	13.3	11.3	8.8	230	11.5	2.8	6.53×10 ⁶	280	99.99	41	28.64	52	71.14	10	0.22	103
MON_120	Rollico	23/1/14	Both	6	2.1	28.7	11.3	8	290	10.6	10.6	1.17×10 ⁷	479	99.78	52	59.32	79	39.19	18	1.49	149
MON_140	Rollico	22/12/14	Both	4	1.8	18.36	8.7	8.28	261	12.8	1	5.43×10 ⁶	434	99.50	62	68.09	101	28.90	67	3.02	230
MON_4	Rebollon	12/11/12	Both	5	1	7.8	14.3	8.2	340	10.4	1.3	–	372	100	83	47.34	69	22.78	47	29.88	201
MON_19	Rebollon	4/12/12	Both	5	1.8	15.7	9.1	8.1	280	10.9	1.3	1.91×10 ⁷	357	99.96	96	54.26	149	30.54	55	15.20	300
MON_32	Rebollon	9/1/13	Bact.	3	2.4	27.2	9.5	8.8	245	14.6	1	1.68×10 ⁷	309	99.95	–	–	–	–	–	–	–
MON_52	Rebollon	20/3/13	Both	5	2.2	26.2	14.8	9.1	258	10.5	13.2	1.12×10 ⁷	395	99.75	28	58.04	62	41.32	9	0.64	99
MON_95	Rebollon	27/11/13	Both	15	1	5.1	7.5	8.1	305	11.6	1.8	1.51×10 ⁷	423	99.94	19	35.98	29	63.88	9	0.14	57
MON_108	Rebollon	18/12/13	Bact.	3	1.1	12.4	10	9.1	242	12.6	27.1	6.66×10 ⁶	397	99.96	–	–	–	–	–	–	–
MON_141	Rebollon	22/12/14	Both	3	1.5	14.57	8.9	8.8	225	14.3	1.6	3.10×10 ⁶	466	99.98	44	29.64	99	68.59	29	1.77	172
MON_10	Pito	13/11/12	Both	15	1.5	18.3	13.9	7.8	330	11.7	6.2	1.65×10 ⁷	358	99.87	117	39.38	175	53.82	59	6.80	351
MON_23	Pito	5/12/12	Both	4	2.3	25	9.2	8	302	10.7	2.3	1.91×10 ⁷	793	98.57	26	40.91	15	59.05	3	0.04	44
MON_36	Pito	10/1/13	Both	10	3	36.5	12	8	295	10.5	4.2	1.22×10 ⁷	413	97.30	21	68.43	19	29.90	5	1.66	45
MON_44	Pito	23/2/13	Both	10	4.8	53.4	1.5	7.9	293.3	12.7	3.8	6.68×10 ⁶	470	99.12	36	52.14	39	47.71	3	0.14	78
MON_56	Pito	21/3/13	Both	5	3.8	45.4	13.9	8	240	11	6.1	4.97×10 ⁶	565	98.55	38	43.61	45	56.20	10	0.20	94
MON_65	Pito	17/4/13	Both	3	7	94.1	22.1	8.1	325	9.7	6.6	1.21×10 ⁷	291	96.82	30	55.08	41	44.51	5	0.42	76
MON_78	Pito	12/6/13	Both	0.2	28	–	34.5	7.5	260	7.5	1.8	8.72×10 ⁶	457	97.64	9	43.14	9	54.68	2	2.18	20
MON_99	Pito	28/11/13	Bact.	10	1.2	12	3.5	8	295	12.6	1.1	6.98×10 ⁶	386	99.51	–	–	–	–	–	–	–
MON_112	Pito	19/12/13	Bact.	5	1.7	20	6.7	8.3	275	11.6	1.4	9.65×10 ⁵	366	99.93	–	–	–	–	–	–	–
MON_125	Pito	23/1/14	Both	8	2.8	34.5	11.2	8	304	10.8	2.7	1.45×10 ⁷	345	99.87	59	16.71	74	78.01	16	5.27	149
MON_146	Pito	23/12/14	Both	15	2.8	30.7	6.1	8.21	335	11.75	1.5	2.93×10 ⁶	437	96.91	22	1.71	47	98.11	8	0.18	77
MON_13	Salineta	13/11/12	Both	13	6.9	66.3	13.1	9.2	278	10.6	79.8	2.25×10 ⁷	127	99.92	29	81.50	32	17.35	16	1.15	77
MON_25	Salineta	5/12/12	Both	25	7.8	77.1	8.2	9	286	12.1	237.4	4.67×10 ⁶	316	99.78	38	67.57	56	30.38	7	2.05	101
MON_38	Salineta	10/1/13	Both	50	10	94.3	12	8.7	275	11.8	71.3	9.17×10 ⁷	149	99.94	33	90.18	55	7.94	8	1.88	96
MON_42	Salineta	22/2/13	Both	30	15	128.8	9.4	8.4	246.6	10.4	35.8	9.97×10 ⁶	351	99.07	34	21.86	47	73.27	18	4.87	99
MON_58	Salineta	21/3/13	Both	30	11	107.5	16.1	8.4	270	10.1	40.9	4.16×10 ⁷	199	98.80	23	26.56	32	70.24	15	3.20	70
MON_67	Salineta	17/4/13	Both	15	24	156.8	28	8.1	188.6	9	23.8	4.58×10 ⁷	323	86.40	15	8.62	7	91.37	1	0.01	23
MON_75	Salineta	26/5/13	Both	5	23	152	19.1	8.2	235	12.9	85.7	2.35×10 ⁶	380	92.23	12	0.20	3	99.80	1	0.01	16
MON_80	Salineta	12/6/13	Both	5	35	150.3	34	7.9	220	9.5	54.3	1.02×10 ⁷	339	83.67	10	0.32	4	99.67	0	0.01	14
MON_84	Salineta	8/7/13	Both	1	38	142	32.5	7.9	230	7.9	5.6	1.84×10 ⁷	185	74.97	27	1.08	20	98.60	8	0.32	55
MON_88	Salineta	9/10/13	Both	1	40	143.3	18.8	7.7	265	8.9	2.4	5.91×10 ⁷	270	53.65	18	1.41	5	96.51	14	2.09	37
MON_101	Salineta	28/11/13	Both	10	7.5	67.5	7.4	–	315	11.6	3.7	3.47×10 ⁷	407	98.18	23	72.70	36	26.26	22	1.04	81
MON_114	Salineta	18/12/13	Both	10	9	85.6	11	8.4	245	15.7	54.7	1.18×10 ⁷	234	99.93	24	46.30	28	53.55	6	0.15	58

Table B.1 Continued from previous page

Code	Site	Date	Data	H.	Sal.	Cond.	Temp.	pH	Redox	D.O.	Chla	DAPI	Bacteria		Photo. P.		Hetero. P.		Fungi		Euk.
													Rich.	Abund.	Rich.	Abund.	Rich.	Abund.	Rich.		
MON_127	Salineta	23/1/14	Both	15	11.2	105.2	12.2	8.8	278	11.1	201.5	2.59×10 ⁷	400	98.58	21	37.56	34	61.75	5	0.69	60
MON_131	Salineta	14/3/14	Both	10	24.2	167.9	10	8	235	10.7	95.5	1.46×10 ⁷	363	90.04	21	10.71	24	89.17	7	0.12	52
MON_134	Salineta	30/4/14	Both	1	37	129.3	24	7.8	235	7.8	87.3	—	261	99.75	36	0.28	56	99.72	18	0.00	15
MON_144	Salineta	23/12/14	Both	20	8.8	80.8	6	8.95	270	13	55.7	7.61×10 ⁶	193	84.28	10	54.37	5	44.73	0	0.89	110
MON_12	Playa	13/11/12	Both	10	2.1	26.3	14.4	8.8	306	12	1.2	3.30×10 ⁷	331	98.94	41	53.31	75	46.28	16	0.41	132
MON_24	Playa	5/12/12	Both	15	3	31	8.5	8.2	303	11.5	0.4	1.32×10 ⁷	321	99.72	48	46.84	88	52.26	18	0.90	155
MON_37	Playa	10/1/13	Both	30	4.5	49.7	11.3	8.4	270	11.6	0.5	2.04×10 ⁷	314	99.28	25	55.93	40	43.61	10	0.46	75
MON_41	Playa	22/2/13	Both	40	5.9	67	11	8.5	230.8	11.3	3.3	9.35×10 ⁶	430	99.20	51	37.40	67	61.34	23	1.26	141
MON_57	Playa	21/3/13	Both	30	5.8	67.5	16.3	8.6	250	12.4	5.9	1.28×10 ⁷	254	99.77	40	18.09	82	80.05	23	1.86	145
MON_66	Playa	17/4/13	Both	40	10.1	124.4	23.5	8.1	312	9.4	3.8	1.89×10 ⁷	387	98.85	33	12.52	42	87.18	8	0.30	83
MON_74	Playa	26/5/13	Both	40	10.6	117.6	17.5	8.7	225	9.2	14	—	355	98.85	40	1.98	31	97.84	12	0.18	83
MON_79	Playa	12/6/13	Both	25	14.8	169.3	27.9	8.1	225	8.8	11.7	2.64×10 ⁷	301	98.93	37	16.93	35	82.79	9	0.28	81
MON_83	Playa	8/7/13	Both	30	16.2	168.1	30	8.2	230	7.3	4	4.35×10 ⁷	129	98.61	26	57.49	25	41.90	15	0.61	66
MON_87	Playa	9/10/13	Euk.	10	34.8	184.3	22.6	7.7	250	8.4	6.8	—	—	—	25	13.93	25	82.91	17	3.15	68
MON_100	Playa	28/11/13	Both	40	5	58.1	4	8	280	11.7	0.8	2.26×10 ⁷	315	99.83	38	85.41	95	13.73	26	0.86	159
MON_113	Playa	18/12/13	Both	25	6	78.3	8.5	8.4	144	12	8	—	385	99.07	30	69.72	30	30.19	5	0.09	65
MON_126	Playa	22/1/14	Both	30	6.6	86.4	10.8	8.6	285	11.5	6	1.31×10 ⁷	308	99.54	29	14.66	48	85.21	8	0.13	85
MON_130	Playa	14/3/14	Both	25	11.5	134.9	10	8.4	215	9.9	29.6	3.99×10 ⁷	361	98.48	31	27.70	37	72.20	5	0.10	73
MON_133	Playa	30/4/14	Both	20	21	194.5	25	8	215	10.1	23.6	—	360	99.47	21	24.71	25	75.25	4	0.05	61
MON_145	Playa	23/12/14	Both	40	4.9	55.6	6	8.71	295	11.78	5.2	6.59×10 ⁶	284	94.58	27	5.61	31	94.32	3	0.08	50
MON_9	A. Alta	13/11/12	Bact.	10	1	4.7	8.6	7.9	326	11.8	0.5	1.88×10 ⁷	378	99.97	105	25.63	158	72.02	27	2.35	295
MON_22	A. Alta	5/12/12	Both	5	1.2	8.1	7.8	7.9	290	11.3	0.6	2.93×10 ⁷	279	99.96	89	15.14	89	78.48	29	6.38	209
MON_35	A. Alta	10/1/13	Both	5	1.7	15.6	10	8	285	11.1	3	3.45×10 ⁷	233	100	79	33.68	98	50.38	30	15.94	211
MON_55	A. Alta	21/3/13	Both	2	1	10.4	15	8.4	265	12.5	0.6	1.19×10 ⁷	254	99.99	60	27.17	122	69.56	20	3.27	203
MON_98	A. Alta	28/11/13	Bact.	20	0.9	3.6	3	8.1	320	12.1	3.4	1.61×10 ⁷	357	99.97	—	—	—	—	—	—	—
MON_111	A. Alta	19/12/13	Bact.	5	0.9	5.3	7.5	8.2	275	10.9	6.8	4.78×10 ⁶	348	100	—	—	—	—	—	—	—
MON_124	A. Alta	23/1/14	Bact.	8	1	8.7	8.5	8	248	11.8	0.9	6.91×10 ⁶	549	99.97	—	—	—	—	—	—	—
MON_148	A. Alta	23/12/14	Both	25	1	6.95	6.3	8.25	299	11.06	1.2	2.11×10 ⁶	327	99.95	49	51.33	106	46.46	32	2.22	189
MON_8	A. Baja	13/11/12	Both	5	0.9	4.6	6.3	8.2	339	12	1.2	2.79×10 ⁷	329	99.98	136	34.69	165	62.36	39	2.95	342
MON_21	A. Baja	5/12/12	Both	2	1.2	7.8	5.9	7.9	242	11.9	0.7	2.51×10 ⁷	386	99.88	58	21.95	75	70.58	22	7.47	155
MON_34	A. Baja	10/1/13	Both	3	1.8	15.6	9.5	8	256	11.6	0.8	3.39×10 ⁷	290	99.93	42	47.93	47	45.15	15	6.92	105
MON_54	A. Baja	21/3/13	Both	2	1.5	13.3	11.2	8.6	258	11.8	0.9	2.05×10 ⁷	252	99.96	52	22.90	120	59.71	23	17.39	196
MON_97	A. Baja	28/11/13	Both	4	0.9	4.3	1.5	9.2	280	11.7	5.5	1.39×10 ⁷	361	99.97	41	52.11	65	47.31	10	0.57	116

Table B.1 Continued from previous page

Code	Site	Date	Data	H.	Sal.	Cond.	Temp.	pH	Redox	D.O.	Chla	DAPI	Bacteria		Photo. P.		Hetero. P.		Fungi	Euk.	
													Rich.	Abund.	Rich.	Abund.	Rich.	Abund.			
MON_110	A. Baja	19/12/13	Both	3	0.9	6.4	7.9	8.3	270	10.5	11.4	6.05×10 ⁶	488	99.97	27	3.16	22	96.79	3	0.06	52
MON_123	A. Baja	23/1/14	Both	2	1.1	10.3	8	8	244	11.6	1.4	8.14×10 ⁶	481	99.94	86	17.15	126	80.24	19	2.60	232
MON_147	A. Baja	23/12/14	Both	3	0.9	5.95	7.2	8.34	296	13.1	1.9	3.16×10 ⁶	384	99.95	66	30.38	99	68.70	23	0.92	190
MON_7	Salobral	13/11/12	Both	20	3.3	31.7	5.9	9	341	9.4	2.9	1.06×10 ⁸	234	99.72	58	52.96	79	46.55	17	0.49	154
MON_26	Salobral	5/12/12	Both	30	3.2	31.3	8	8.8	275	13.5	1.7	1.31×10 ⁷	347	98.92	78	21.39	165	76.54	35	2.07	279
MON_39	Salobral	10/1/13	Both	40	3.2	32	11	8.9	280	13.8	2.9	2.88×10 ⁷	347	99.52	66	28.93	135	69.88	25	1.19	227
MON_43	Salobral	22/2/13	Both	30	3.7	35.6	10.1	8.9	210	11.5	3.8	1.60×10 ⁷	361	99.77	44	41.17	81	57.10	21	1.74	146
MON_59	Salobral	21/3/13	Bact.	50	3	31.3	16	8.9	230	10.8	15.5	4.97×10 ⁶	311	98.55	—	—	—	—	—	—	—
MON_68	Salobral	17/4/13	Both	40	2.9	34.3	25.8	9.3	170	8.2	4.1	1.66×10 ⁷	238	99.43	41	29.41	54	70.54	4	0.05	99
MON_76	Salobral	26/5/13	Both	40	3.4	36.2	20	9.5	180	14.2	4.4	—	216	99.49	38	4.32	56	94.76	16	0.92	110
MON_81	Salobral	12/6/13	Both	40	3	42.4	30.3	9.1	190	11.5	2.8	1.64×10 ⁷	332	98.70	51	9.15	85	90.10	16	0.74	154
MON_85	Salobral	8/7/13	Bact.	30	3.8	47.1	29.6	8.8	122.6	8.1	5.4	3.33×10 ⁷	388	96.22	—	—	—	—	—	—	—
MON_77	Balsa M	26/5/13	Bact.	—	0.1	0.9	17.4	8.3	195	9.8	3	1.89×10 ⁷	288	100	—	—	—	—	—	—	—
MON_86	Balsa M	8/7/13	Euk.	40	0.1	1	29.3	8.4	210	8.7	1.3	1.55×10 ⁷	—	88	41.54	62	57.82	16	0.64	166	
MON_89	Balsa M	9/10/13	Euk.	40	0.1	0.9	18.8	9.2	199	9.7	44.2	—	—	94	24.27	135	71.04	33	4.69	263	
MON_102	Balsa M	28/11/13	Euk.	20	0.1	0.7	6.4	8.5	230	12.4	25.4	1.02×10 ⁷	—	82	26.88	93	72.93	9	0.19	184	
MON_115	Balsa M	17/12/13	Euk.	40	0.1	0.7	6.5	8.6	260	13.2	18.6	—	—	107	16.46	79	83.07	17	0.48	203	
MON_128	Balsa M	22/1/14	Bact.	40	0.3	0.7	8	9.1	240	12.1	9.7	6.02×10 ⁶	401	99.98	—	—	—	—	—	—	—
MON_135	Balsa M	30/4/14	Euk.	30	0.1	0.8	21	8.8	200	10.7	2	—	—	50	44.65	34	51.18	21	4.17	105	

Table B.2: Pearson's correlation and significance for richness and diversity values against salinity (log, % w/v). For eukarya, correlation is given considering the whole eukaryal community and also photosynthetic and heterotrophic protists and fungi, independently.

Group	zOTU richness		Shannon-Weaver diversity index	
	Pearson's cor	Significance (pvalue)	Pearson's cor	Significance (pvalue)
Bacteria	-0.144	<i>ns</i>	-0.059	<i>ns</i>
Eukarya	-0.589	<0.001	-0.533	<0.001
Photosynthetic protists	-0.547	<0.001	-0.352	<0.001
Heterotrophic protists	-0.603	<0.001	-0.438	<0.001
Fungi	-0.442	<0.001	-0.218	<0.05

Table B.3: Most abundant zOTUs within the taxa found along the gradient for the different saline ranges and their individual sequence GenBank accession number.

Domain	Salinity range	Taxonomic group	zOTU	Relative abundance inside the taxonomic group (%)	GenBank sequence identifier
Bacteria	Freshwater to slightly-brackish	Flavobacterium	ZOTU_29	39.3	MT129582
		Burkholderiaceae-unclass	ZOTU_162	48.6	MT129606
		Fluviicola	ZOTU_97	44.0	MT129599
		Pedobacter	ZOTU_71	49.2	MT129593
		Hydrogenophaga	ZOTU_95	48.3	MT129598
		Polaromonas	ZOTU_42	93.3	MT129586
		Porphyrobacter	ZOTU_13	64.9	MT129577
		Pseudomonas	ZOTU_31	59.4	MT129583
		Polynucleobacter	ZOTU_60	85.8	MT129591
		Linnohabitans	ZOTU_74	89.9	MT129594
	Brackish	Peredibacter	ZOTU_283	12.0	MT129612
		Seohaecicola	ZOTU_8	81.9	MT129573
		SM2D12-unclass	ZOTU_359	16.5	MT129614
		Rhodobacteriaceae-unclass	ZOTU_3	44.7	MT129569
		Rheinheimera	ZOTU_22	63.0	MT129580
		Algoriphagus	ZOTU_32	64.7	MT129584
		Pseudohongiella	ZOTU_236	36.5	MT129610
		Salinarimonas	ZOTU_160	39.2	MT129605
		Boseongicola	ZOTU_197	100.0	MT129607
	Marine	Methylophaga	ZOTU_10	61.0	MT129575
		Thiomicrospira	ZOTU_41	84.5	MT129585
		Aliidimarina	ZOTU_107	75.0	MT129600
		Fabibacter	ZOTU_108	100.0	MT129601
		Izimaplasmataceae-unclass	ZOTU_48	55.3	MT129589
		Idiomarina	ZOTU_59	32.5	MT129590
		Loktanella	ZOTU_14	47.6	MT129578
		Acholeplasma	ZOTU_79	74.9	MT129595
		Litoricola	ZOTU_91	99.9	MT129597
		NS3a marine group	ZOTU_9	100.0	MT129574
	Hypersaline	Psychroflexus	ZOTU_1	56.6	MT129568
		Roseovarius	ZOTU_11	57.0	MT129576
		Sulfurimonas	ZOTU_351	36.0	MT129613
		Phaeodactylibacter	ZOTU_110	83.1	MT129602
		Pseudoalteromonas	ZOTU_45	58.7	MT129587
		Microbacteriaceae-unclass	ZOTU_4	94.6	MT129570
		Alteromonas	ZOTU_47	65.4	MT129588
		Vibrio	ZOTU_69	41.9	MT129592
		Marivita	ZOTU_5	94.5	MT129571
	Extreme hypersaline	Nitrococcaceae-unclass	ZOTU_7	87.1	MT129572
		Spirochaeta 2	ZOTU_276	38.2	MT129611
		Halanaerobacter	ZOTU_82	53.3	MT129596
		Halobacteroidaceae-unclass	ZOTU_390	27.1	MT129615
		Halomonas	ZOTU_16	62.0	MT129579
		Halothiobacillus	ZOTU_157	51.6	MT129604
		Spiribacter	ZOTU_212	49.4	MT129609
		E6aC02-unclass	ZOTU_154	83.5	MT129603
		Halanaerobium	ZOTU_204	51.4	MT129608
		Salinivibrio	ZOTU_28	95.3	MT129581
Eukarya	Freshwater to slightly-brackish	Ciliophora	ZOTU20	12.2	MT095067
		Fungi	ZOTU47	16.9	MT095069
		Cercospora	ZOTU15	33.6	MT095066
		Discicristata	ZOTU75	38.9	MT095070
		Labyrinthulomycetes	ZOTU168	13.9	MT095071
		Cryptomonas	ZOTU30	98.0	MT095068
		Teleaulax	ZOTU323	98.7	MT095073
	Brackish	Ochrophyta	ZOTU1	43.4	MT095062
	Marine	Metazoa (Animalia)	ZOTU4	50.3	MT095065
	Hypersaline	Dinoflagellata	ZOTU2	95.3	MT095063
		Bicosoecida	ZOTU313	59.5	MT095072
	Extreme hypersaline	Chlorophyta	ZOTU3	33.7	MT095064

B.2 Supplementary Figures

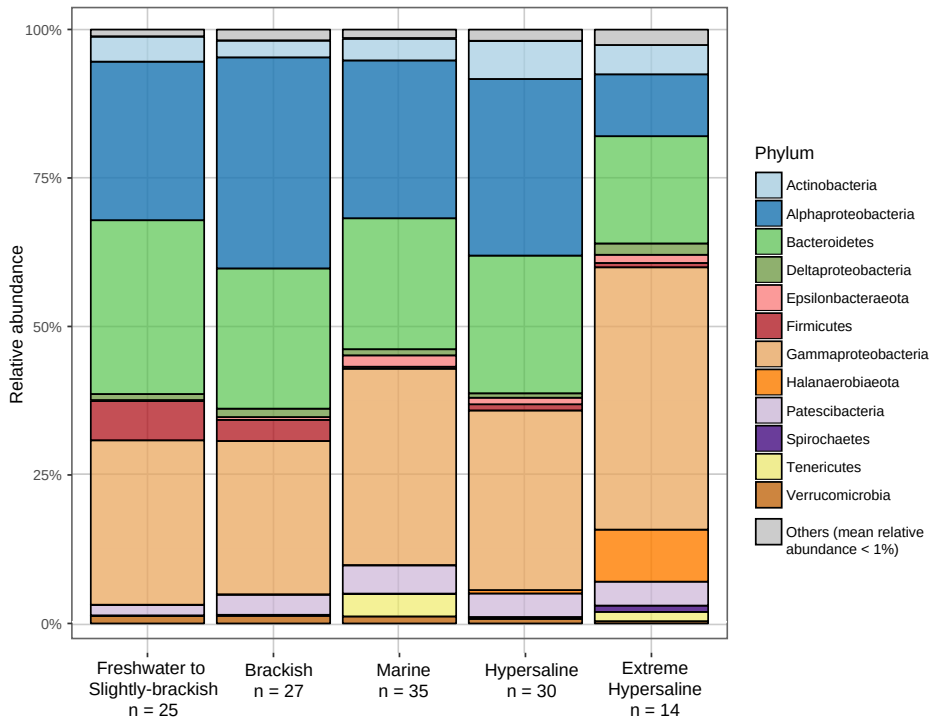


Figure B.1: Taxonomic profile representing the most abundant bacterial phyla along the salinity ranges.

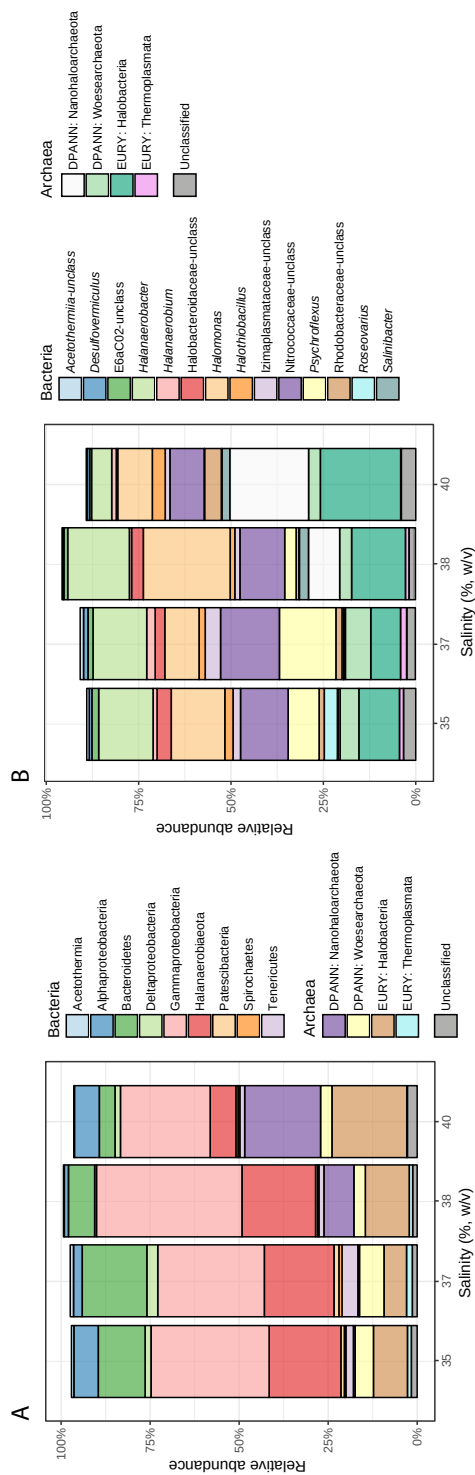


Figure B.2: Detailed taxonomic profile for the extreme saline samples (>35% w/v) including the whole 16S rRNA gene pool (archaea and bacteria), at phylum level (a) and at higher taxonomic resolution (b) (genus level, when possible).

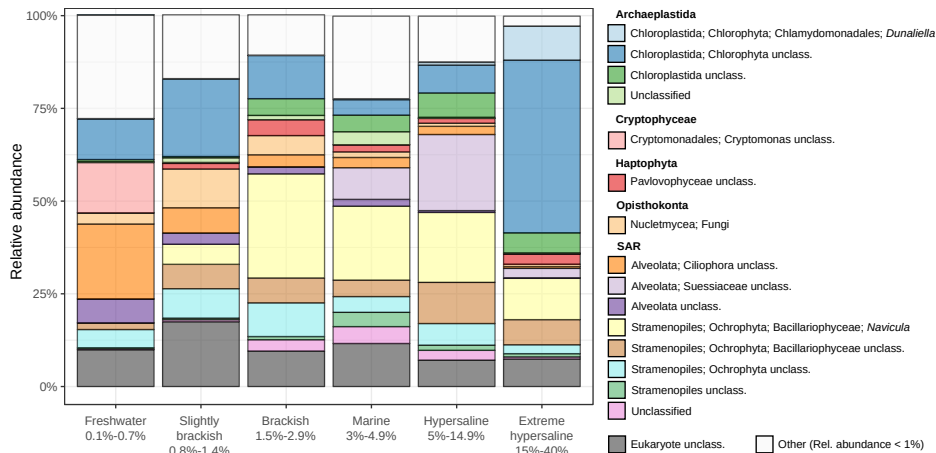


Figure B.4: Taxonomic profile representing the most abundant eukaryotic groups along salinity ranges.

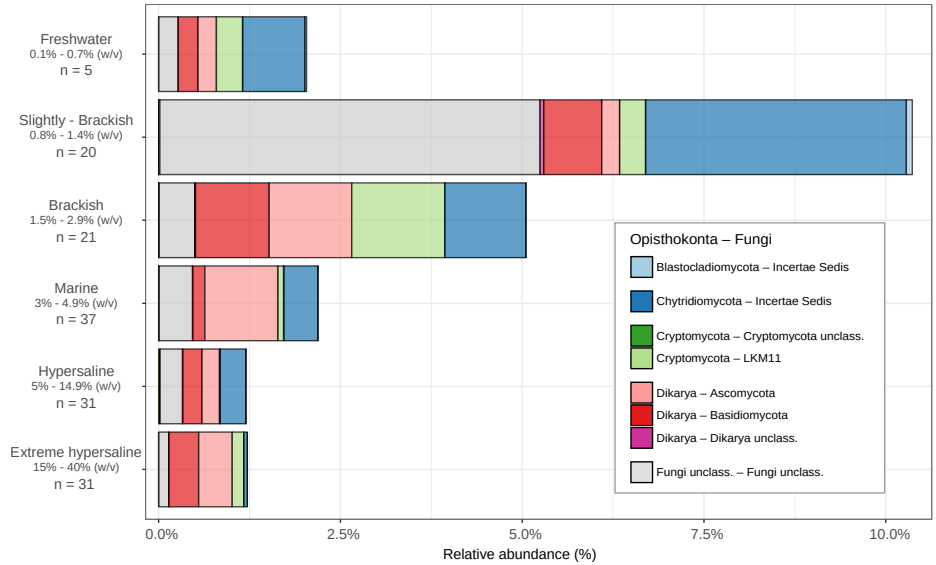


Figure B.5: Taxonomic profile of the Fungi community along salinity ranges, considering their relative abundance within the whole 18S rRNA gene pool.

Chapter 4 Supporting Information

C.1 Supplementary Information

Study site and sampling procedures

Samples were collected from the Monegros Desert area (NE Spain, 41°42'N, 0°20'W) which harbors one of the largest sets of inland saline lakes in Europe. The singularity of these saline wetlands (locally named *saladas*) has been already documented from different viewpoints (e.g. Castañeda et al. 2013). Shallow saline lakes exhibit a wide microbial diversity, which has been recently described (Casamayor et al., 2013; Triadó-Margarit and Casamayor, 2013). This semi-arid region includes an endorheic area (c.a. 400 km²) with 149 scattered saline depressions ranging in size from < 2 to 239 ha. Some of these depressions host hypersaline wetlands and their brines are of Cl-SO₄-Na-(Mg) type.

A total of 148 samples were taken from 14 different lagoons located on the left side of the Ebro river at a mean elevation of 328 m a.s.l. (Figure C.4). From these lagoons, 13 of them were salt sites and another one, named as Balsa Molinos, was a freshwater pond used as a local reservoir. Samples covered two complete hydrological periods and an isolated sample from the beginning of the third period. Conductivity, temperature, pH, dissolved oxygen and redox potential were measured *in-situ* using an HQ40d multi-parameter meter (Hach, Loveland, CO, USA). Salinity was measured by a hand refractometer (Atago S-28E, Japan). Variations in salinity, water column height and temperature were severe both among lagoons and inside lagoons along the hydrological period. Such hydrological changes in playa lakes are conditioned by the high-evaporation/low-rainfall regime, presenting one of the highest mean annual water deficits in Europe (Herrero and Snyder, 1997). Notwithstanding, ephemeral brines occur in periods of hydric deficit, due to the proximity of a permanent saline water table. Then, different salinity conditions were recorded among study sites and sampling times. We explored the spatio-temporal variation of a regional set of local aquatic communities driven by a wide salinity gradient (0.1 to 40% of dissolved salts) (Figure C.5 A). The vast majority of these lagoons are shallow, making them sensitive to changes in the atmospheric temperature and also to strong wind perturbations prevailing in

the region. Temporal series showing wind speed values along the sampling period were available from the Meteorological state agency reporting data of an automatic station in the area (Bujaraloz) (Figure C.5 B).

DNA-based molecular analyses

DNA extraction and sequencing.

For DNA analyses, water samples were prefiltered in situ through a 50 micron-pore-size net, to retain large zooplankton and algae, and 100–500 mL were subsequently filtered on 5-micron and then 0.2-micron pore-size polycarbonate membranes (47 mm diameter). The membranes were stored in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose), enzymatically digested and phenol extracted (phenol–chloroform–isoamyl alcohol; 25 : 24 : 1, v/v/v). The purification and concentration step was carried out with Amicon® Ultra 4 Centrifugal Filter Units – 100 000 NMWL (Millipore). DNA extracts were quantified using Quant-it assays in a Qubit™ fluorometer (Invitrogen). PCR and sequencing were done according to the genomic core facilities and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>). For NGS sequencing step, high-speed multiplexed SSU rRNA gene sequencing with the Illumina MiSeq platform was carried out. A read length of 2x250 bp amplicons was applied. The variable V4 region of the 16S rRNA gene (c.a. 250 nucleotides) were amplified with the primers F515 (5′-GTGCCAGCMGCCGCGGTAA-3′) and R806 (5′-GGACTACHVGGGTWTCTAAT-3′) (Caporaso et al., 2011). The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605, and also upon request to the authors.

Sequence processing

Raw sequences were processed using the UPARSE pipeline (Edgar, 2013). The dataset had an overall % of sequences $\geq$ Q30 greater than 70%. The total number of sequences prior to quality filtering was 14,445,263. After merging of read pairs, filtering by read length (250 bp) and an expected error of 0.25, we kept 6,536,143 (corresponding to 45% of the original reads). The dereplicated reads were sorted by size and singletons were excluded. Then, chimeric reads were detected and filtered out by UCHIME following two strategies: de novo and reference-based (against the ‘Gold’ reference database available for ChimeraSlayer). These reads were clustered to Operational Taxonomic Units (OTUs) by using a cut-off of 0.03% as identity threshold. A total of 12,681 OTUs were obtained. At

this point, sequences pool prior to singletons and chimera discarding were mapped back into the OTUs set, using an identity threshold of 0.97. As a result, 88.9% of those sequences were matched (5,810,631 reads). The taxonomic assignment was done under SINA aligner v.1.2.11 (Pruesse et al., 2012), using SILVA_128 reference database (Quast et al., 2013) with the de-replicated version at 99% sequence similarity (SSURef_NR99_123_SILVA). We filtered out 361 OTUs by their poor alignment quality score (<80) and their percent identity (<75). Then, we excluded those classified as chloroplasts or mitochondria, 229 and 136 OTUs, respectively. A set of 1,927 archaeal and 24 eukaryal OTUs were also discarded from the analysis. In addition, we kept only those samples that at this point conserved library sizes greater than 10,000 reads/sample. As a result, the final OTU table contained 9,993 OTUs, being detected in a pool of 136 samples.

Matching genomes

In order to associate OTUs to genome content, we used a 16S rRNA database (Wattam et al., 2014) linked to the IMG genomic database (Markowitz et al., 2014). OTUs were matched to 16S rRNA gene records available in the PATRIC genomic database (as of January 2016) (Wattam et al., 2014) using the `usearch_global` command (Edgar, 2010). We linked a total of 678 OTUs to 544 closest genomic matches of at least 97%. This represented a total of 3,017,902 sequences (63% of the dataset). The average percentage of abundance matched per sample was 61.84%, with a minimum of 9.38% and a maximum of 91.7%. Functional predictions based on representative genomes allow to infer genomic and metabolic potential of 16S rRNA sequences (Langille et al., 2013; Nemergut et al., 2016). For that purpose we downloaded from IMG Markowitz et al. (2014) available genomic traits of potential interest: Genome size, Gene count, % GC, % Coding base, % CDS, % RNA, rRNA count, %Transporter proteins, % Signal peptide and % Transmembrane proteins. Some of these traits accurately indicate ecological properties of microbial communities, such as genome size or rRNA count (Barberán and Casamayor, 2014; Ortiz-Álvarez et al., 2018b).

Null models of community assembly

The theoretical basis of null models of community assembly developed in this research are presented in full detail in the main manuscript. In addition, the explanations of the methodological framework and a user's guide to carry out the method are presented below (see the section *RTCC Method protocol: Randomized Trait Community Clustering*).

Statistics

Statistical analyses were run in the R environment (<http://www.r-project.org/>). Community ecology related parameters and multivariate statistics were calculated by using the *vegan* package (Oksanen et al., 2016) and figures were drawn with *ggplot2* (Wickham, 2009). For distance-based community analyses, Bray-Curtis dissimilarities after Hellinger standardization Legendre and Gallagher (2001) were calculated and represented by non-metric multidimensional scaling (*metaMDS* function). We used Mantel tests with the Spearman method to determine the correlation between community similarity patterns from bacteria and variables in metadata. Besides, correlation analyses based on Spearman's rank order coefficient were carried out to investigate whether the association between the NMDS axis scores and environmental variables. Significant breakpoints of observed patterns along environmental gradients, either from ordination analyses or clustering indices, were estimated by means of the maximum F statistic derived from sequential Chow-tests (*sctest* function from *strucchange* package) Zeileis et al. (2003)(see Figure C.3).

RTCC Method protocol: Randomized Trait Community Clustering

In this section we provide detailed instructions to guide users. Figure C.1 shows the different methodological steps followed in this work. RTCC method is a trait-based clustering index method which can be applied to either phenotypic or genotypic trait data and to these observed from macro-organismal communities. Then, the steps presented in the *top panel* (in gray) that are related to OTU-based approach and to the trait inference (or collection) have not to be necessarily as those presented here, and are presented as preliminary steps. Related to these preliminary steps, Figure C.2 shows the results used to make a decision about the representativity of the subset matching genomes in our approach. Complementary, a Mantel Test (based on Spearman correlation) was also used to conclude on this decision-making step.

Randomization tests for the RTCC method are implemented in *C* language (file *permut.c*) and then called by several *python* scripts. For that purpose, the python *cypes* library is needed. Other standard python libraries to be imported are *numpy*, *pandas*, *random*, *scipy* and *matplotlib*. We provide a *OSX* compilation of the *C* shared library in the file *permlib.so*. For other operating systems a shared library has to be compiled. The *gcc* command line used to compile a shared library for python scripts to properly work is

```
gcc -shared -o permlib.so -fPIC permut.c -lc -lm
```

For other *C* compilers substitute the *gcc* command by the corresponding one.

In the following subsection, we provide methodological details necessary to complete the different steps indicated in the *lower panel* of the Figure C.1 (from *BOX 1* to *5*). Inset frames provide details of command lines to run the specific scripts. Code sources for our scripts can be found at the end of this text. Examples for the analyses are provided as well.

Methodological framework

BOX 1: Apply randomization test to selected traits based on a random null model —use *gen-Metacom.py* script

For each sample we test the hypothesis that observed average trait dissimilarity (see Eq.1 in the main manuscript) is compatible with a given null model. Non-parametric randomizations performed allow us to assign a p-value for each sample. For all samples in the dataset we represent the p-value distributions obtained by the null model for a given functional trait (see Figure 5.1 in the main manuscript). By using the script *gen-Metacom.py*, different files with p-value distributions are obtain, each one corresponding to a specific trait. Note that the analysis has been only performed by using the *Random assembly* null model, which considers equal probability on choosing species from the pool of all samples —metacommunity. Indeed, if clustering is not observed for this random model, it would be less likely to be observed in the cases of the models *ii*, *iii* and *iv*, because they are nested and, then, they are increasingly restrictive.

STEP 1 for the analysis: Run the script *gen-Metacom.py* to generate p-value distributions for all traits (0–9) using 100 null model realizations to calculate p-values (purely random null model). Required files: Refer to provided example data of traits (*traits_plus_range.csv*) and presence-absence data frame (*table_presence_absence.csv*).

```
python gen-Metacom.py 100
```

BOX 2: Select those traits exhibiting a strongest significant clustering to conduct the following analyses —use *boxplot.py* script

The script allows to obtain clustering distribution patterns for different functional traits into a single figure (see Figure 3 in the main manuscript). Significantly low p-values are related to clustering ($p < 0.05$), and significantly high p-values to over-dispersion. For the whole sample set, which we define as a metacommunity, we calculate the fraction of samples that yield significant clustering, i.e., its clustering index h (see Eq.2 in the main manuscript).

STEP 2: Run the script *boxplot.py* to plot p-value distributions using the values calculated in *STEP 1*. Indicate the type of image file to be drawn —*png* in the example.

```
python boxplot png
```

BOX 3: Sequential exclusion of samples in decreasing (or increasing) order of the environmental variable *AND* Apply different null models to alternative metacommunities —use *hindex-repetitions.py* script

We average several realizations of the null model in order to obtain average clustering index values and error bars (see Figure 1B in the main manuscript). Samples are sequentially removed along decreasing values of the environmental variable to define nested metacommunities; for each of them we calculated the clustering index h (see Figure 5.1 C —left panel— and Figure 5.4 A-B in the ms). For the sequential removal of samples, values of the environmental variable have to be sorted in decreasing order in the metadata file.

Note that to adjust the analysis in decreasing order is necessary to indicate to the variable *ord_flag* in the script a value different than 1 (e.g. set a 0). In this case, the function returns the curves from the h index corresponding to the indicated null model along the gradient of the environmental variable, using the specified trait in the vertical axis variable, and *nrep* repetitions to give an average value of the h index. Set the variable *test* —randomization test indicator— to values 0, 1, 2 or 3 to apply different null models to alternative metacommunities (corresponding to *random*, *ranges*, *weighted by abundances*, and *ranges-weighted* models). If necessary, to analyze data in reverse order, i.e., in increasing order of the gradient, sort the file of metadata from the lesser to the higher value of the variable.

STEP 3: Run the script *hindex-repetitions.py* to generate h index values for the trait 8 (0-9) using 100 realizations for the null model 2 (weighted by abundances) with decreasing order of the environmental variable (`ord_flag = 0`). The h index is averaged using 50 repetitions. Required files: Refer to the provided example data of traits (*traits_plus_range.csv*), presence-absence data frame (*table_presence_absence.csv*), metadata for a specific environmental gradient (*metadata.csv*) (sort the table in decreasing order —alternatively in increasing— of the variable), and the relative abundance (a number between 0-1) of each species in the whole metacommunity —i.e., the entire species pool (*metacomm_sum.csv*).

```
python hindex-repetitions.py 100 8 50 2 0
```

BOX 4: Check that exclusion order matters (through random orderings) —use *hindex-repetitions.py* script

The significance of this curve is assessed by building an ensemble of curves corresponding to removing samples in random orderings, which are used to define the shaded gray area (see Figure 5.1 C —right panel— and Figure 5.4 A-B in the ms). Then, this shaded area corresponds to the confidence interval for the same null model used in the step above but randomizing sequential sample removal. Note that to adjust the analysis at random orderings of the environmental variable is necessary to set the `ord_flag` in the script equal to 1.

STEP 4: Run the script *hindex-repetitions.py* to generate h index for the trait 1 (0-9) using 100 realizations for the null model 3 (weighted by abundances and ranges of the environmental gradient) using random orderings of the environmental variable (`ord_flag = 1`). The h index is averaged over 100 repetitions. Required files: The same as those necessary to proceed with the example in *STEP 3*.

```
python hindex-repetitions.py 100 1 100 3 1
```

A significant correlation of the trait clustering pattern with an environmental gradient exists if clustering index values for a given trait (under ordered sequential removal) lie outside the shaded area, this is, the null model fails to explain the variation of clustering along the gradient. By combining the output data from the *STEP 3* and *4* is possible to decide about the existence of such differences.

BOX 5: Chow-test to infer thresholds —use *Chow-test.R* script

Estimating structural changes in regressions models to infer thresholds by means of Chow test (*sctest* function from the package "strucchange"). The output obtained after running this script corresponds to such presented in Figure C.3 by using the example file.

STEP 5: Run the script *Chow-test.R* in R to estimate breakpoints (i.e., structural changes) in the progress of Test F Statistic —Chow test— along the gradient of the environmental variable. Note that breakpoints are determined from the maximum value of the statistic. Taking the value of the covariate that corresponds to the maximum value of the statistic allows to obtain a threshold value. Required data: clustering index values for a given trait as a function of the environmental variable resulting from *STEP 3*. See the file *clustering_index_vs_env_gradient.csv*.

A threshold in the value of an environmental variable along a gradient is indicating substantial structural changes in community assembly drivers.

Source code for the RTCC Method and sample data files

The following files are available in **GitHub**:

<https://github.com/xtriado/Consistent-thresholds-RTCC>

Script 1: *gen-Metacom.py*

Script 2: *boxplot.py*

Script 3: *hindex-repetitions.py*

Script 4: *Chow-test.R*

RTCC C code: *permut.c*

OSX shared library: *permlib.so*

Sample data 1 - *traits_plus_range.csv* Trait data and environmental ranges of the pool of species in the metacommunity.

Sample data 2 - *table_presence_absence.csv* Presence-absence data.

Sample data 3 - *metadata.csv* Salinity gradient data from 0.1 to 40% of dissolved salts.

Sample data 4 - *metacomm_sum.csv* Relative abundance data (between 0 and 1) for each species, normalized to the total abundance observed in the metacommunity.

Sample data 5 - *clustering_index_vs_env_gradient.csv* Clustering index curve data.

C.2 Supplementary Figures

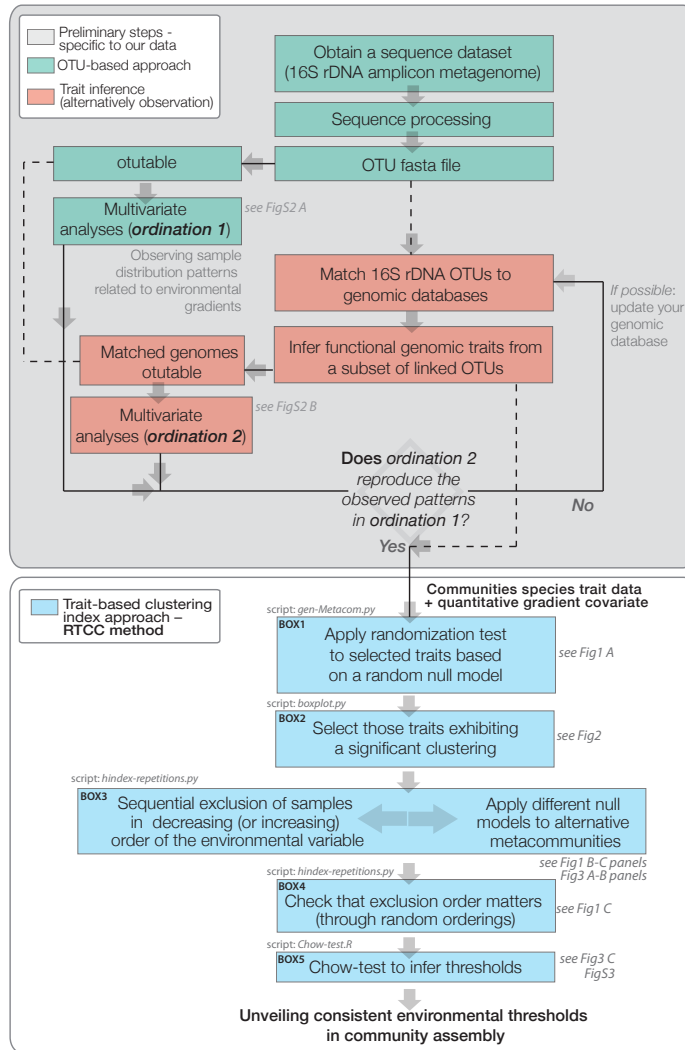


Figure C.1: **Flow diagram** The different steps carried out on the development of the study are summarized in the diagram. Note that the presented method can be applied to either phenotypic or genotypic trait data and to those observed from macro-organismal communities. Then, the steps related to OTU-based approach and those related to the trait inference (or collection) have not to be necessarily as those presented here, and are presented as preliminary steps (*top panel* in gray). The *bottom panel* shows the methodological steps specific to RTCC method. The correspondence between the different methodological steps and their related results, as well as the program scripts used, is indicated in gray labels as side marks.

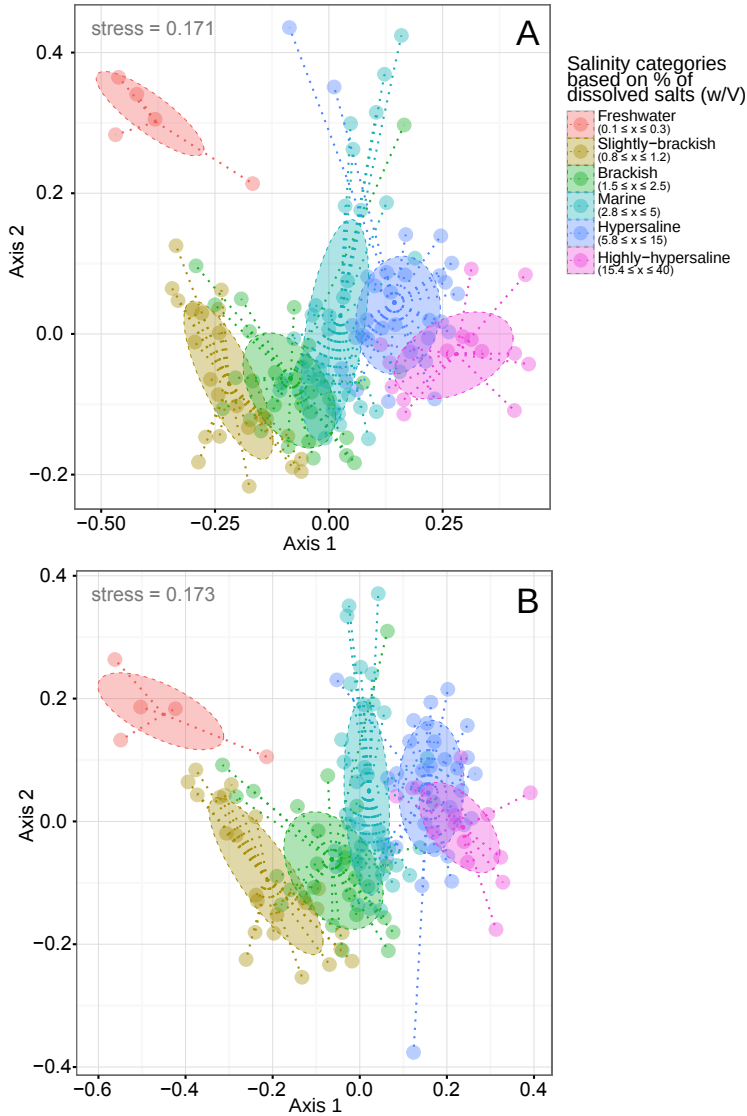


Figure C.2: Whole communities *vs* community subsets based on matching genomes Non-metric multidimensional scaling (NMDS) ordination analysis of community similarities based on Bray-Curtis dissimilarities considering the whole dataset (A) and the subset of OTUs matching bacterial genomes used for our trait-based approach (B). Colours are related to the salinity categories based on % of dissolved salts (w/V). The ellipsoid areas are indicative of distribution (and dispersion) of each group on the ordination space. We used this figure to assess if the ordination pattern of the whole dataset is consistent with the one of the subset (see the flowchart of the figure S1). We compared the community dissimilarity matrices of the subset with the whole community through a mantel test (Mantel r – Spearman: 0.9197; $p = 0.001$) (Fig. S2).

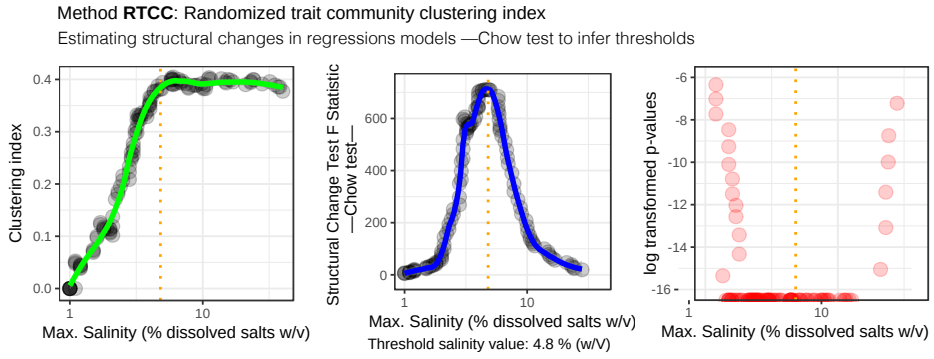


Figure C.3: **Chow-test**. Estimating structural changes in regressions models to infer thresholds by means of Chow test: the final step of the RTCC method. The presented figure corresponds to the output of the script "Chow-test.R" obtained for the example file *clustering-index-vs-env-gradient.csv*. **Left panel** reproduce the results of the clustering index pattern observed for the case of the salinity gradient by using null model *iv* (see figure 4B-C in the main manuscript). **Central panel** shows the structural change Test F statistic —Chow test— for the provided data. **Right panel** shows the p-values related to such statistics. Note that breakpoints (i.e., structural changes) are determined from the maximum value of the statistic. Taking the value of the covariate that corresponds to the maximum value of the statistic allows to obtain a threshold value. Based on our clustering index-based approach, a threshold in the value of an environmental variable along a gradient is indicating substantial structural changes in community assembly drivers.

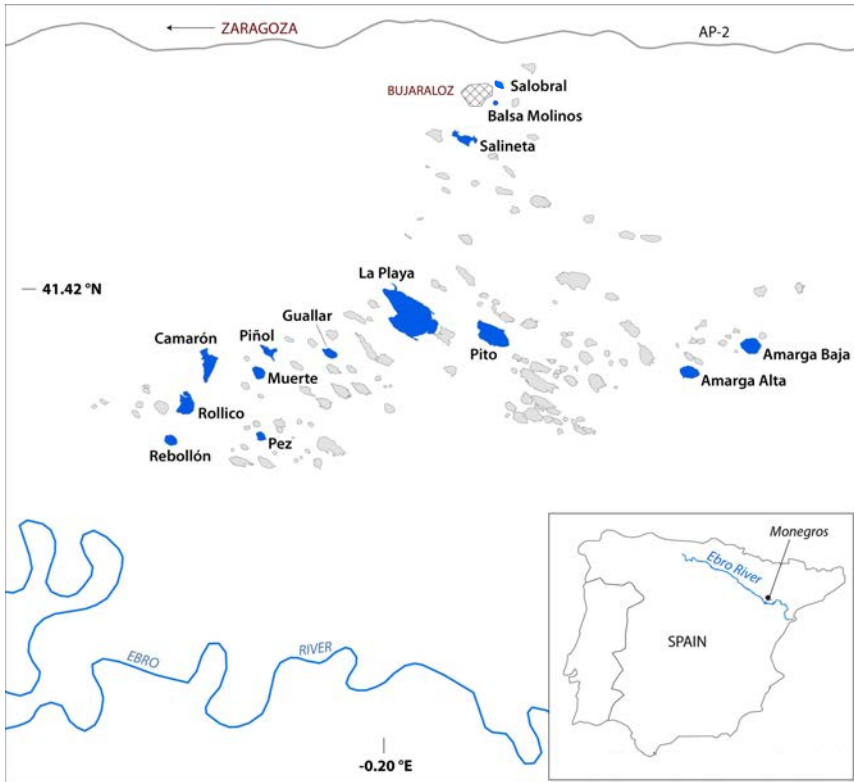


Figure C.4: **Study Sites.** Geographical location of the Monegros Desert endorheic area within the semi-arid Central Ebro Basin, NE Spain. The system includes a landscape of scattered saline depressions excavated on a gypsum-rich bedrock (gray spots on the map). The basins for the investigated lagoons in the present study are highlighted.

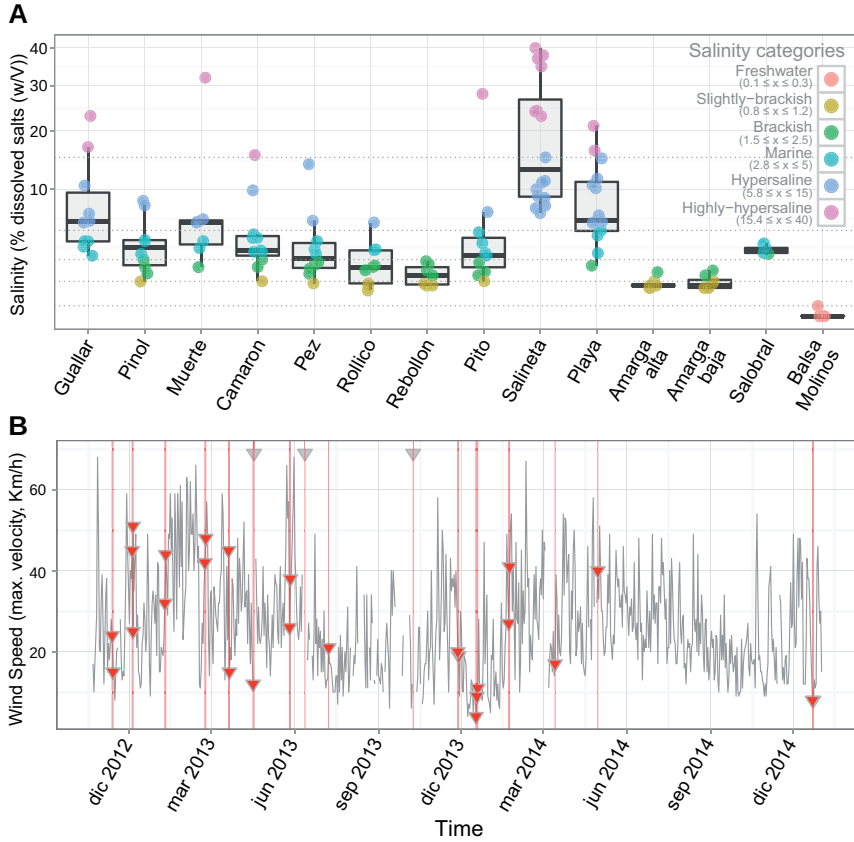


Figure C.5: **Environmental factors.** Box-plots showing the gradient explored for salinity in the 14 study sites (panel A). Temporal series for the sampling period showing wind speed as the value of maximum average wind velocities on a 10 minutes intervals registered during a day, expressed in Km/h (panel B). Sampling days are marked with vertical red lines. Red triangles are indicating the value of wind speed at these times.

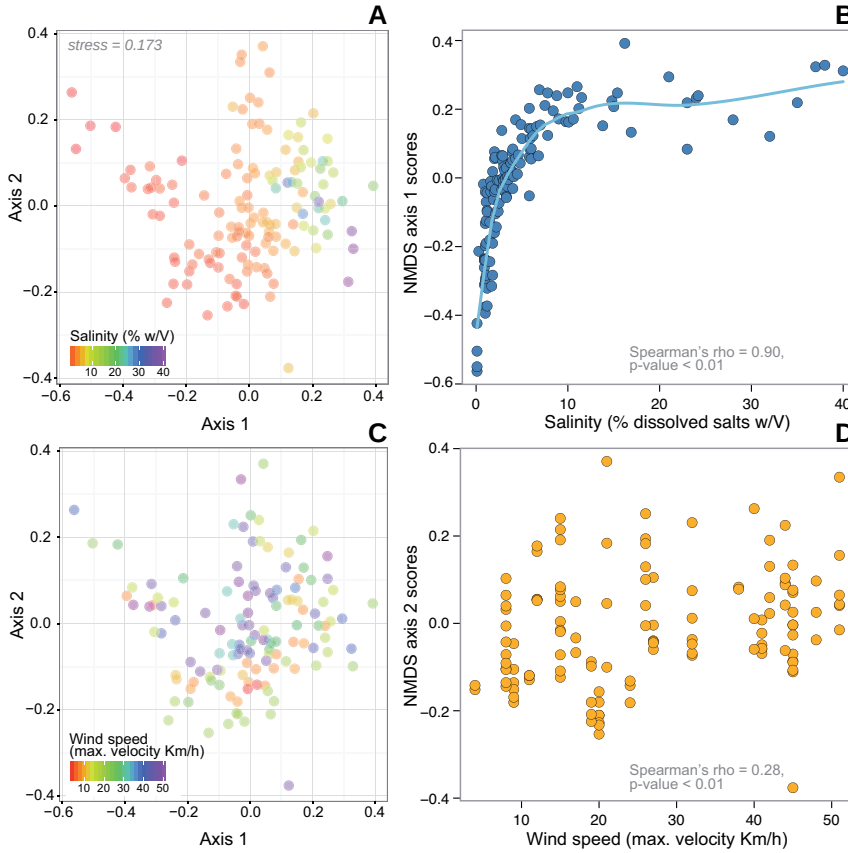


Figure C.6: Ordination analysis Non-metric multidimensional scaling (NMDS) ordination analysis of community similarities based on Bray-Curtis dissimilarities (A and C). Points represent samples and are coloured accordingly to their salinity (panel A) and wind speed (panel C). Relationship between NMDS axis scores (representing community similarity) and environmental parameters: salinity in panel B and wind speed in panel D. A Local Polynomial Regression Fitting (loess function – blue line) was used to represent the monotonic association of NMDS axis scores and salinity (panel B). We found contrasting patterns for two environmental factors: i) a deterministic gradient forced by salinity variation, and ii) a stochastic continuous perturbation due to the strong wind prevailing in the region. Bacterial communities presented significant differences in composition and structure along the salinity gradient (Mantel r – Spearman: 0.3262 ; Significance: 0.001) (panel A). Moreover, NMDS axis 1 scores and salinity showed a strong monotonic variation (Spearman's rank correlation coefficient $\rho = 0.90$, $p < 0.01$), panel B). Wind perturbation did not determine a significant distribution on the ordination analysis (Mantel r – Spearman: 0.03184; $p = 0.055$). In addition, the relationship between NMDS axis 2 scores and values for wind speed was weak ($\rho = 0.28$, $p < 0.01$) in comparison with that observed for salinity (panels C and D). We observed significant breakpoints of the NMDS axis 1 score at 2.1% and 7 – 10% of dissolved salts (w/V) (values were calculated by means of Chow tests, see methods section).

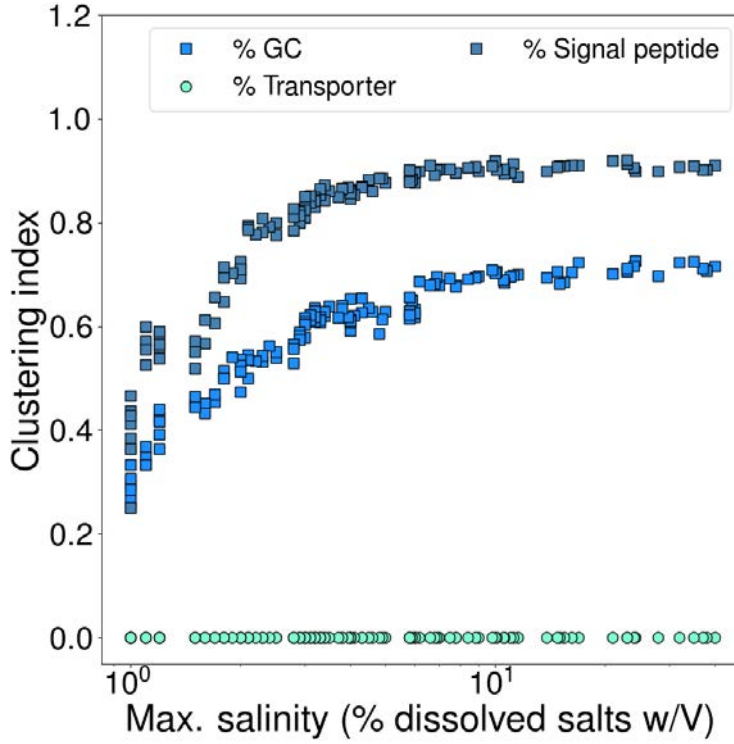


Figure C.7: **Different trait performance in the variation of clustering index *vs* salinity** The percentage of GC pairs in sequences exhibits a similar behavior in the variation of clustering index *vs.* salinity (here we show results for the random assembly null model). Note that for more complex models we expect clustering index to decrease, hence they will determine the threshold poorly. Other traits can exhibit no clustering signal at all (i.e., the percentage of DNA associated to transporter proteins, see Figure 3 in the manuscript).

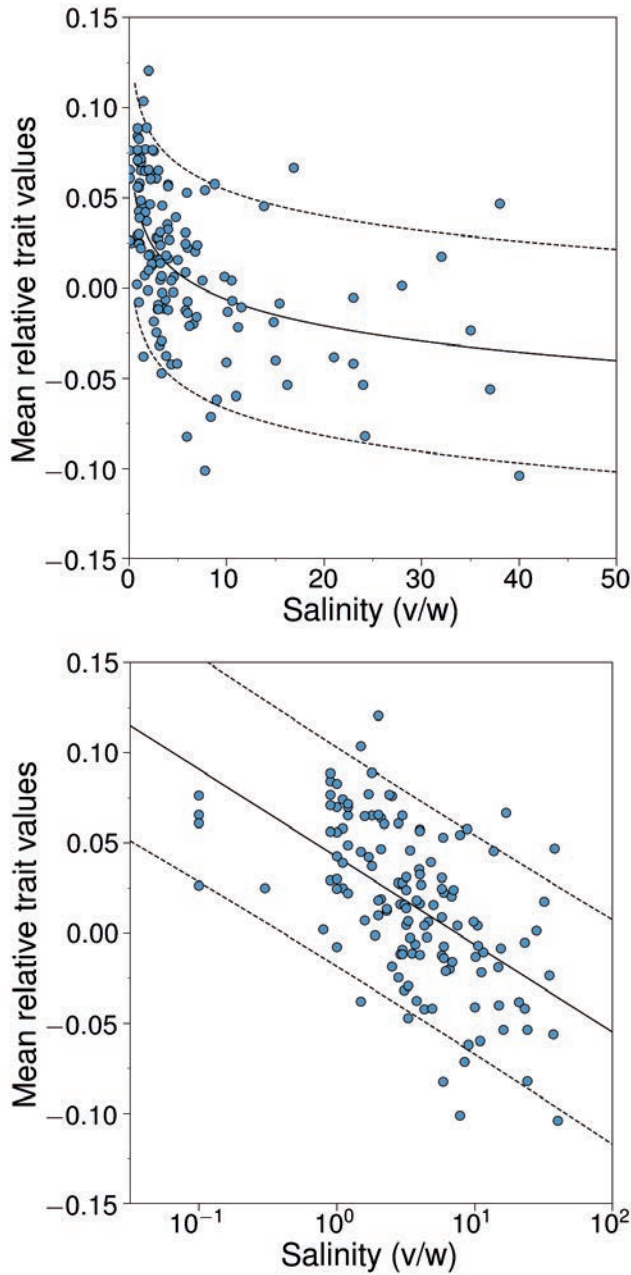


Figure C.8: **Trait variation along salinity gradient** Signal peptide values averaged across genomes detected in each sample (relative to the metacommunity average) along the salinity gradient. Both linear and logarithmic scale representations are used for salinity axis.

Chapter 5 Supporting Information

D.1 Supplementary Tables

Table D.1: Variation partitioning analysis (VPA) results including forward selected environmental variables and number of selected PCNMs.

Microbial group	Salinity range	Selected environmental variables	N° PCNMs	[E S]	[EnS]	[S E]	Residuals
Bacteria	Full metacomunity (0.1 - 40%)	conductivity, salinity, pH, temperature, redox	9	0.1	0.128	0.154	0.619
Bacteria	Low salinity range (0.1 - 4.9%)	salinity, pH, temperature, redox	7	0.078	0.104	0.166	0.651
Bacteria	High salinity range (5 - 40%)	salinity, pH, temperature, conductivity	2	0.163	0.125	0.081	0.631
Eukarya	Full metacomunity (0.1 - 40%)	conductivity, salinity, pH, temperature, redox	9	0.095	0.09	0.143	0.672
Eukarya	Low salinity range (0.1 - 4.9%)	salinity, pH, redox, temperature	8	0.056	0.08	0.137	0.727
Eukarya	High salinity range (5 - 40%)	salinity, pH, temperature, conductivity	3	0.187	0.106	0.096	0.611

Table D.2: Pond salinity variability, Akaike information criterion (AIC) of each model, and model improvement when adding salinity.

Pond	Salinity min. and max. value (%)	Microbial community	AIC without salinity	AIC with salinity	Model improvement with salinity (R ²)
Guallar	2.8 – 23	Bacteria	-10862.7	-10749.0	0.09
		Eukarya	-3435.8	-3386.7	0.29
Pinol	1.2 – 8.4	Bacteria	-12975.9	-12762.9	0.09
		Eukarya	-2377.8	-2366.2	0.14
Camaron	1.1 – 15.4	Bacteria	-13928.9	-13837.4	0.14
		Eukarya	-4283.9	-4197.7	0.57
Pez	1.1 – 13.8	Bacteria	-14068.5	-14017.6	0.04
		Eukarya	-4746.1	-4681.2	0.20
Rollico	0.8 – 5.8	Bacteria	-13796.6	-13668.9	0.04
		Eukarya	-4498.4	-4451.6	0.26
Rebollon	1 – 2.4	Bacteria	-7903.3	-7876.5	0.03
		Eukarya	-1967.0	-1907.0	0.34
Pito	1.2 – 28	Bacteria	-15670.8	-15532.6	0.09
		Eukarya	-2563.7	-2458.9	0.45
Salineta	6.9 – 40	Bacteria	-13526.1	-13422.4	0.13
		Eukarya	-3191.4	-3126.2	0.45
Playa	2.1 – 34.8	Bacteria	-19958.0	-19841.6	0.06
		Eukarya	-6160.4	-6061.7	0.48
Amarga Alta	0.9 – 1.7	Bacteria	-7996.9	-7877.0	0.29
		Eukarya	-2206.9	-2191.2	0.74
Amarga Baja	0.9 – 1.8	Bacteria	-9138.4	-9102.2	0.00
		Eukarya	-3807.1	-3782.0	0.07
Balsa Molinos	0.1 – 0.3	Bacteria	-445.8	-445.8	0.00
		Eukarya	-1988.1	-1988.1	-0.01
Muerte	2 – 32	Bacteria	-7621.6	-7595.4	0.87
		Eukarya	-2934.2	-2792.6	0.68
Salobral	2.9 – 3.8	Bacteria	-8586.2	-8533.1	0.14
		Eukarya	-2958.5	-2949.4	-0.01

D.2 Supplementary Figures

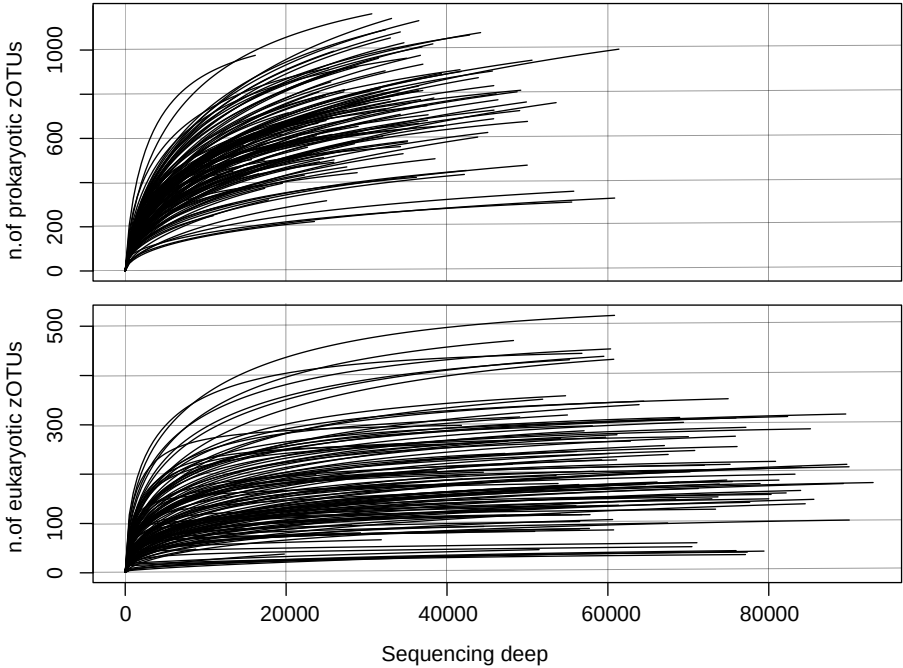


Figure D.1: Rarefaction curves of each sample for both microbial groups.

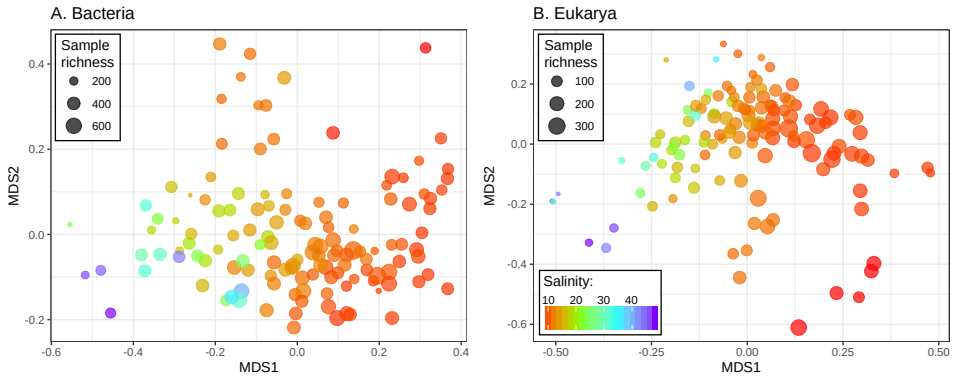


Figure D.2: NMDS representation of the Bray-Curtis distances among samples for both datasets, with salinity (color) and sample richness (size) represented for each sample.

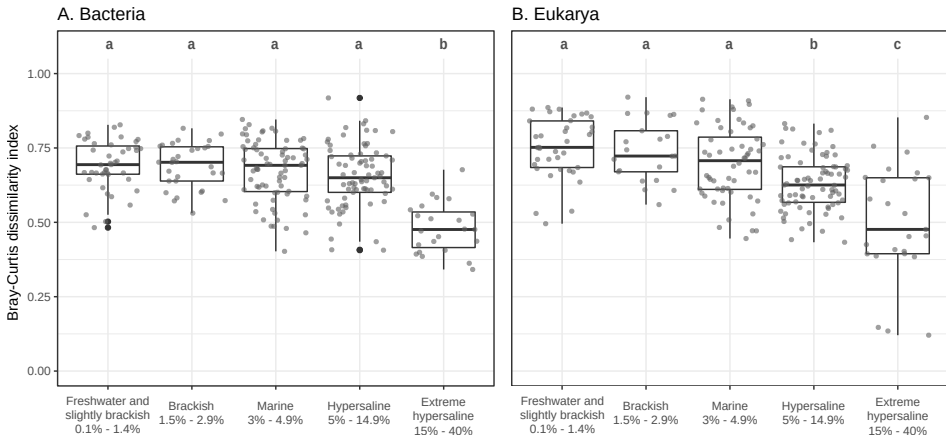


Figure D.3: Boxplot representation of the Bray-Curtis dissimilarity indexes between samples belonging to the same pond and salinity range for both datasets. The analysis was restricted to pairs of samples of the same pond in order to remove the potential effect of dispersal. Likewise, only pairs of samples of the same salinity range were considered in order to reduce the potential effect of wide salinity variations.

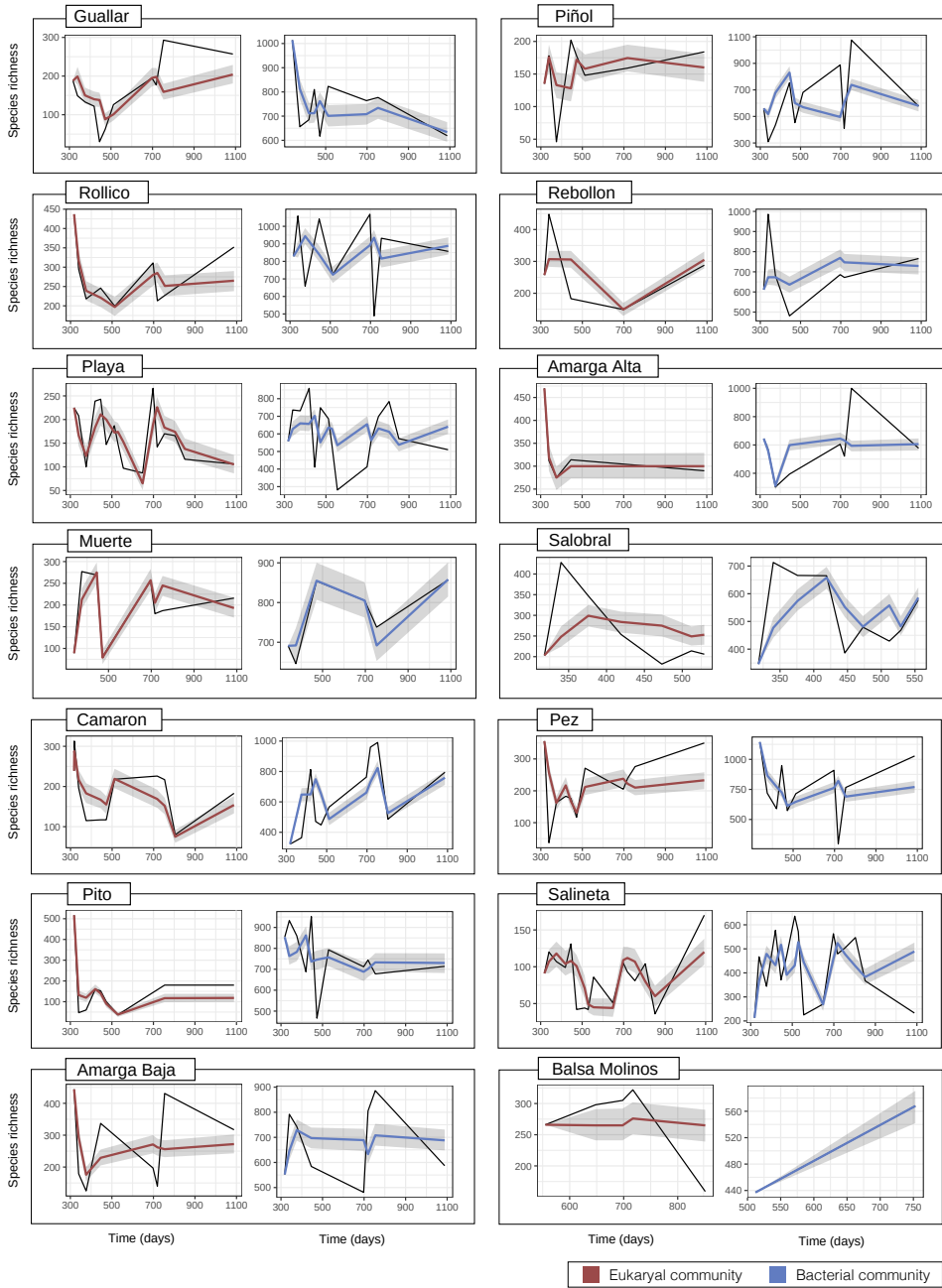


Figure D.4: Bacterial and microeukaryotic richness of sampled ponds along the temporal series and the predicted richness values (mean and 95% confidence interval) based on colonization and extinction rates influenced by salinity oscillations.

Chapter 6 Supporting Information

E.1 Supplementary Information

Sampling scheme

The microbial dataset used in the present study contains a detailed spatio-temporal survey of bacterial, eukaryal and archaeal communities inhabiting the Monegros Desert lacustrine system (NE Spain, 41°42N, 0°20W), which has been already characterized in previous studies (Menéndez-Serra et al., 2020, 2021). Sampling was carried out monthly in 14 different ponds of the Monegros Desert lacustrine system along three hydroperiods. Water samples covered wide salinity variations, from freshwater (0.1 %, w/v) to near saturation (40%), both between and inside ponds due to dilution and concentration processes. Previous studies showed strong effect of salinity fluctuations on the microbial communities (Casamayor et al., 2013; Menéndez-Serra et al., 2020, 2021; Triadó-Margarit et al., 2019).

Sequencing and raw sequences processing

For DNA analyses, water samples were prefiltered in situ through a 50 µm-pore-size net to retain large zooplankton and algae, followed by a sequential filtering on 5 µm (mainly retaining nanoplanktonic and a fraction of macroplanktonic eukaryotes) and 0.2 µm (retaining bacteria and archaea) pore-size polycarbonate membranes (47- mm diameter). Microbial communities were studied by the PCR amplification and Illumina sequencing of the 16S and 18S rRNA genes. For prokaryotes, i.e. bacteria and archaea, PCR amplification of the V4 region of the 16S rRNA gene (primers set 515F-806R, Caporaso et al. 2010) and sequencing with the Illumina 2x250 MiSeq platform were carried out according to the genomic core facilities protocols and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>). PCR amplification of the V9 region of the eukaryal 18S rRNA gene (primers set 1391f-EukrBr, Amaral-Zettler et al. 2009) and sequencing with the Illumina 2x150 MiSeq platform were also carried out following the same protocols. Raw sequences were processed using the UPARSE pipeline (Edgar, 2013). Sequences were merged and quality filtered. After de-noising

and chimera filtering, the UNOISE algorithm (Edgar, 2013) defined Operational Taxonomic Units at 100% identity, i.e., zero-radius OTUs (zOTUs). The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012), using the SILVA 132 reference database (Quast et al., 2013). The resulting data set contained 116 samples with 5,653 prokaryotic (bacteria and archaea) zOTUs and 2,658 eukaryal zOTUs. The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605.

Null network models construction

In order to test the significance of the descriptive parameters of the observed networks, null network models were computed by two complementary approaches: (i) Erdős-Rényi (Erdős and Rényi, 1960) random graph model and (ii) the randomization of the presence-absence observation matrix prior to the network construction process. The Erdős-Rényi random graph model $G(n,p)$ computes graphs with a fixed number of nodes (n) and interaction probability between two nodes (p). Random graphs were created using the *sample_gnp* function on “igraph” R package (Csardi and Nepusz, 2006), based on the number of nodes and the probability of interaction between nodes of the observed networks (Connor et al., 2017). For the randomization of the presence-absence matrices we implemented the “curveball” algorithm (Strona et al., 2014) which randomizes the matrices while maintaining species occurrences and samples species richness. Following matrix randomizations, network construction workflow was applied over the randomized matrices as described in the *Network construction* subsection in the main text. This approach allows us to obtain a parameter distribution of additional networks parameters as node number, link number and mean degree, which remain fixed on the Erdős-Rényi random graphs models construction process. For both null models and each network and subnetwork, 1000 randomized networks were computed in order to create a parameters distribution to determine a p-value for the observed parameters. Observations laying outside the 95% confidence interval of the simulated distribution ($p\text{-value} < 0.05$) denoted a non-random assembly in the network. Null models parameters distributions and the significance of observed parameters are represented in the Supplementary Figures E.4 and E.5.

E.2 Supplementary Figures

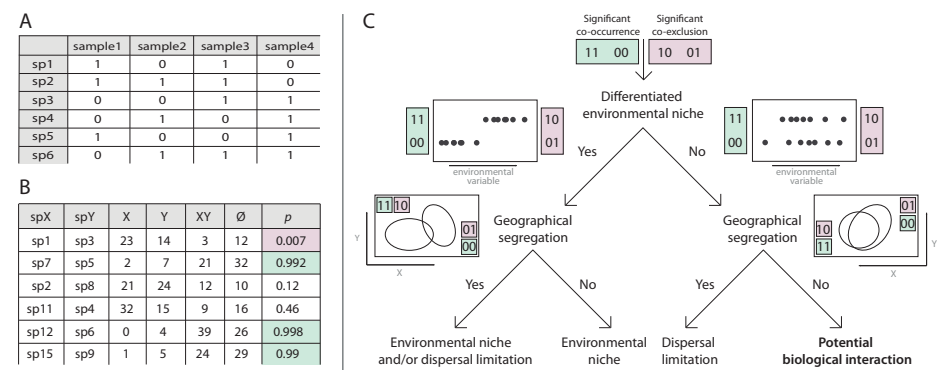


Figure E.1: Network construction workflow. From a presence/absence observation table (A), the probabilistic method proposed by Veech determines the signification of the observed co-occurrences (p) compared to a probability based on the number of samples where only one (X,Y), both (XY) or none (Ø) of the species appears (B). The observed co-occurrence can be significantly higher than expected ($p > 0.99$) evidencing a significant co-occurrence, or lower than expected ($p < 0.01$) evidencing a significant co-exclusion. Following (C), the framework proposed by Blois and collaborators allows to determine the nature of the statistically significant co-occurrences and co-exclusions. This is done after simultaneously test for each pair of species the potential differences for the selected environmental variable and the spatial distribution of samples groups. For co-occurrences analyses, differences are tested between samples where both species had been observed and samples where any of them had been observed (11,00). For co-exclusions, differences are tested between samples where species of the pair had been observed individually (10,01).

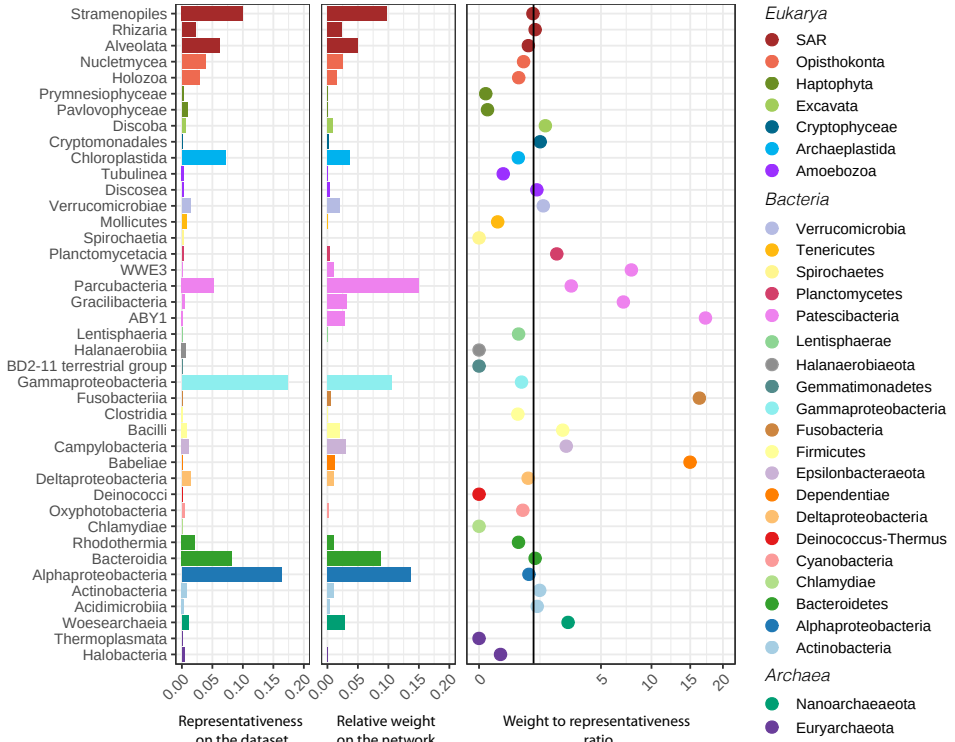


Figure E.2: Representativeness on the dataset, relative weight on the co-occurrence network, and both aspects ratio for the different taxonomic groups in the complete metacommunity network. Bacteria and archaea are grouped at the class level. Eukaryotes are grouped at the L2 level of the SILVA database. Vertical black line indicates 1:1 ratio.

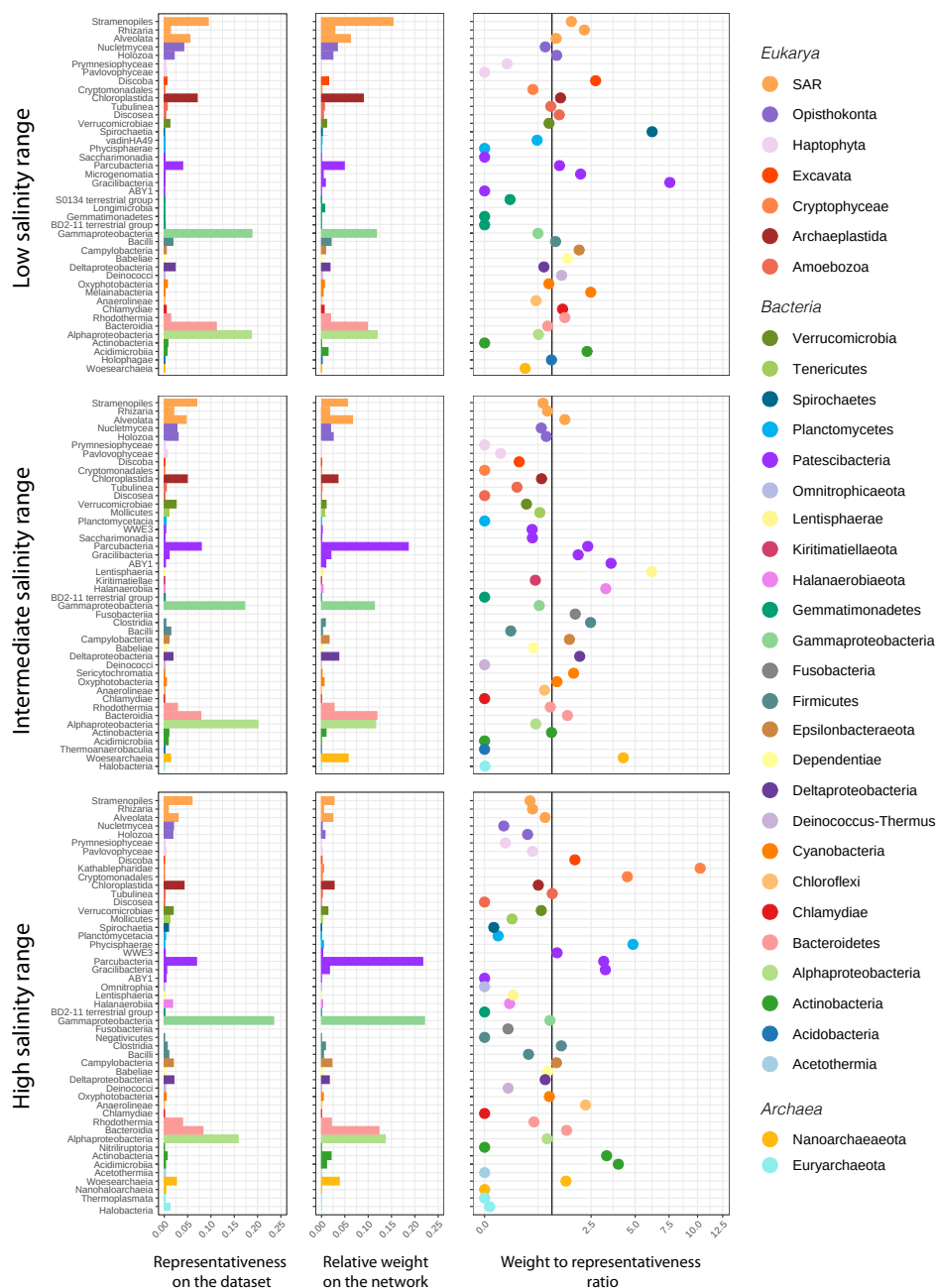


Figure E.3: Representativeness on the dataset, relative weight on the co-occurrence network, and their ratio for the different taxonomic groups at the three salinity ranges. Bacteria and archaea are grouped at the class level. Eukaryotes are grouped at the L2 level of the SILVA database. Vertical black line indicates 1:1 ratio.

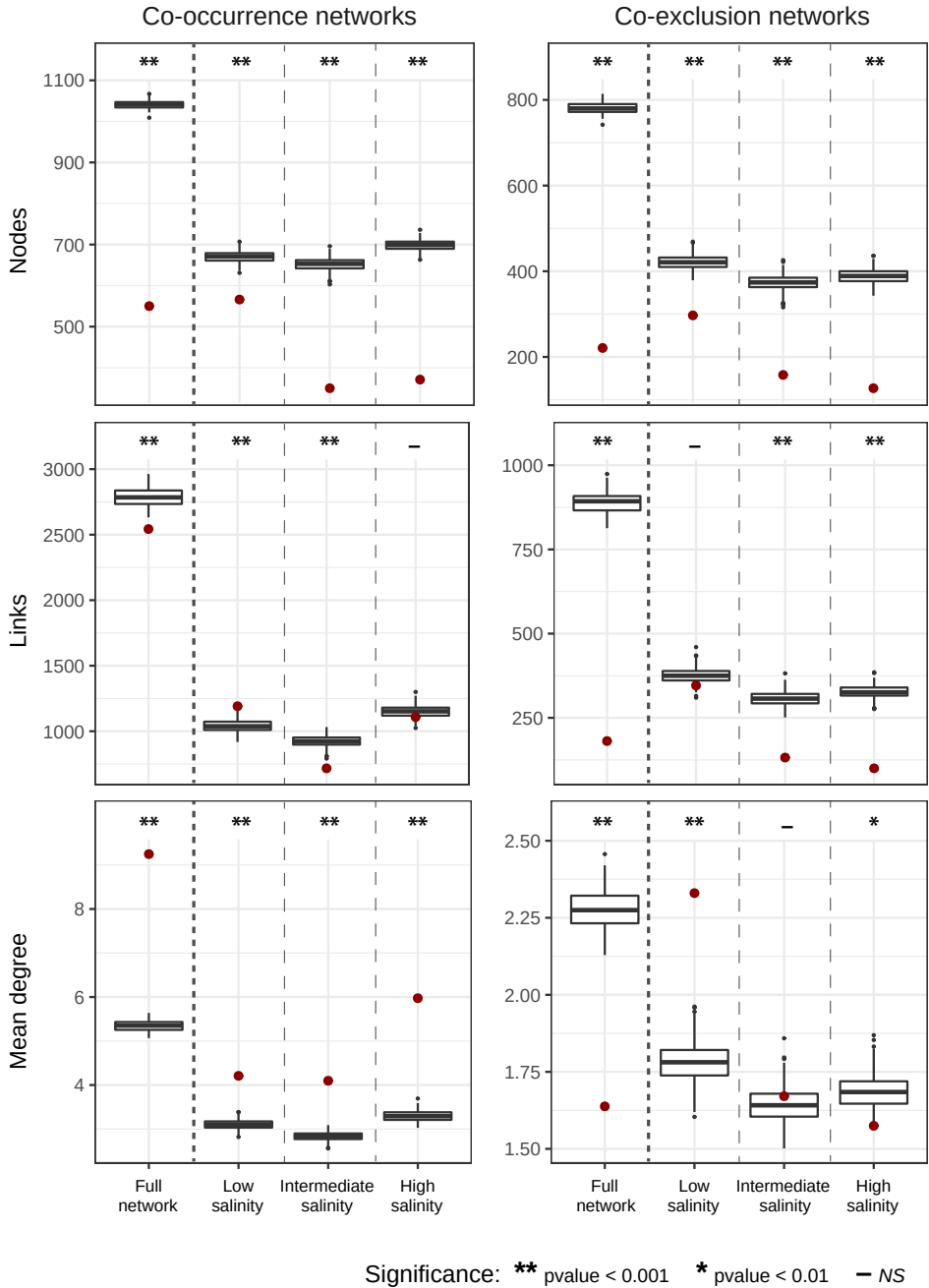


Figure E.4: Boxplot representation of the network parameters exclusively obtained from the matrix randomization null model approach. Red dots show the observed parameter values in the full networks and the consecutive subnetworks.

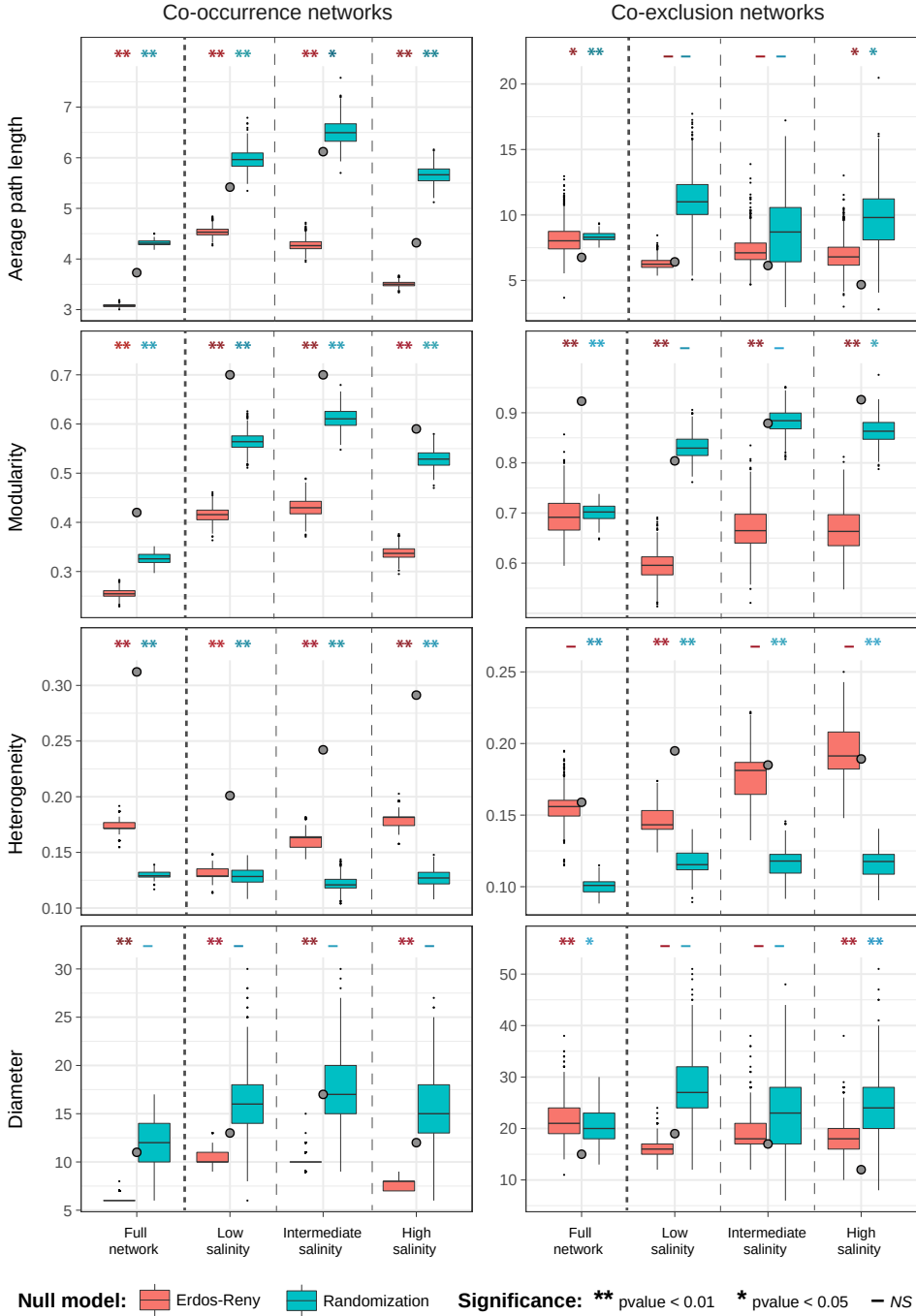


Figure E.5: Boxplot representation of the networks descriptive parameters calculated for both null networks models. Grey dots represent the observed parameter values on the full networks and the consecutive subnetworks.

Chapter 7 Supporting Information

F.1 Supplementary Figures

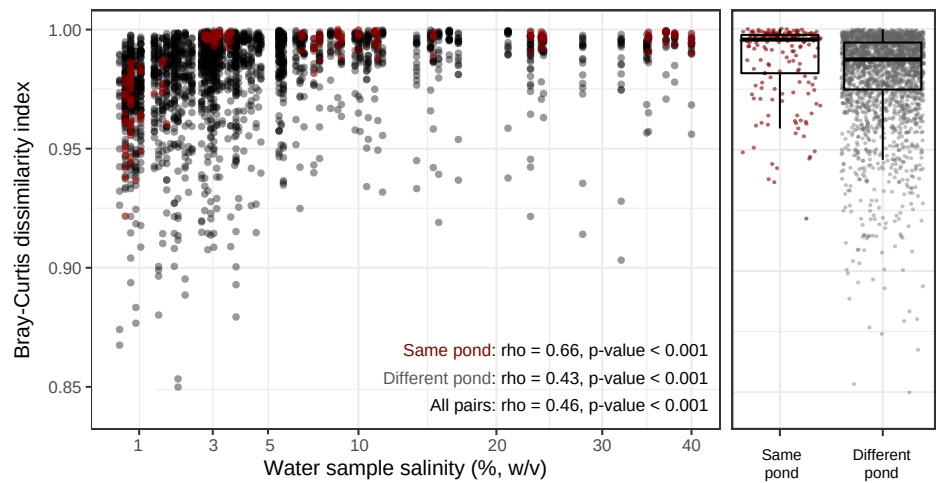


Figure F.1: Pairwise dissimilarity between water-sediment pairs of samples represented along the salinity gradient of water samples (left panel). Boxplot representation of water-sediment pairs of samples Bray-Curtis dissimilarity index, for the same and different pond (right panel).

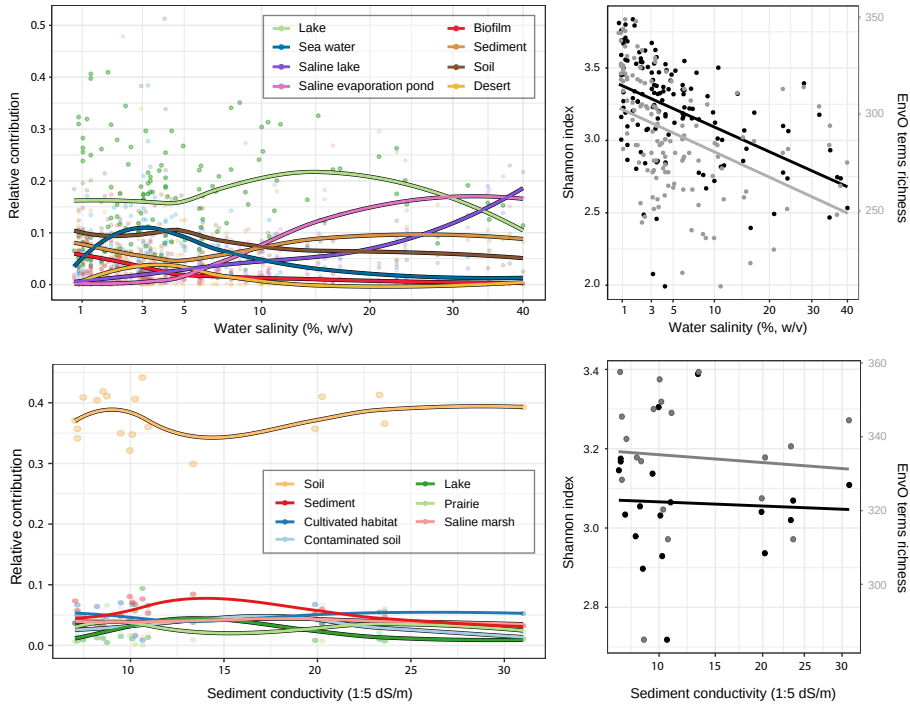


Figure F.2: Evolution of the most abundant sources (EnvO terms) for each water and sediment sample along the salinity and conductivity gradient, respectively (left panels). Diversity of sources in water and sediment samples (richness and Shannon index; grey and black dots, respectively), along the tested gradient (right panel).

Published Chapters

RESEARCH ARTICLE

Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain)

Mateu Menéndez-Serra¹, Vicente J. Ontiveros², Xavier Triadó-Margarit¹, David Alonso² and Emilio O. Casamayor^{1,*}

¹Integrative Freshwater Ecology Group, Centre of Advanced Studies of Blanes (CEAB), Spanish Research Council (CSIC), 17300 Blanes, Catalonia, Spain and ²Theoretical and Computational Ecology Group, Centre of Advanced Studies of Blanes (CEAB), Spanish Research Council (CSIC), 17300 Blanes, Catalonia, Spain

*Corresponding author: CEAB-CSIC, Accés Cala St Francesc, 14. E-17300. Blanes (SPAIN). Tel: +34 972 336 101; Fax: +34 972 337 806; E-mail: casamayor@ceab.csic.es

One sentence summary: Dynamics and idiosyncratic patterns of ecological distributions were unveiled for the Archaea microbiome of inland saline lakes from one of the most arid regions in Europe.

Editor: Martin W. Hahn

ABSTRACT

We characterized the rich Archaea microbiome of shallow inland lakes (Monegros Desert, NE Spain) by 16S rRNA gene tag sequencing covering a wide salinity range (0.1%–40% w/v) along 3 years. Up to 990 operational taxonomic units (OTUs; >97% identity) were detected allocated in 14 major archaeal phyla and heterogeneously distributed along the salt gradient. Dynamics and idiosyncratic ecological distributions were uncovered for the different phyla. A high genetic richness was observed for Woesearchaeota and Pacearchaeota (>370 OTUs each), followed by Halobacteria (105), Nanohaloarchaeota (62) and Thermoplasmata (19). Overall, the distribution of genetic richness was strongly correlated with environmental niche amplitude, but not with occurrence. We unveiled high occurrence for a very rich Woesearchaeota assemblage, and an unexpected positive correlation of Pacearchaeota abundance with salinity at >15% dissolved salt content. The estimated dynamic behaviour (temporal ‘turnover’ rates of presence/absence data) unveiled Thaumarchaeota and Halobacteria as the most dynamic groups, and Aenigmarchaeota and Thermoplasmata as the most stable. The DPANN Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate rates, suggesting higher resilience to environmental perturbations. A rich and dynamic Archaea microbiome was unveiled, including unseen ecological traits for relevant members of the still largely unknown DPANN group, supporting a strong ecological differentiation between Pacearchaeota and Woesearchaeota.

Keywords: archaea; inland saline lakes; community ecology; DPANN Woesearchaeota; Pacearchaeota; halophiles; 16S rRNA gene

INTRODUCTION

Inland lakes in endorheic areas are very dynamic ecosystems globally distributed in semi-arid and arid regions of the world

(Williams 1996). Most shallow inland lakes show wider ranges of salinities, environmental conditions, and salt compositions than coastal areas (Williams 1996). Numerous microbial ecology studies have been carried out on diverse terrestrial saline

Received: 1 October 2019; Accepted: 30 January 2020

© FEMS 2020. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

environments in all the continents (North- and South-America (e.g. Demergasso et al. 2008; Dillon, McMath and Trout 2009), Australia (Narasimarao et al. 2012), European Mediterranean areas (Gasol et al. 2004; Ventosa et al. 2014), Asia (Wu et al. 2006), Africa (Cavalazzi et al. 2019), Antarctica (Williams et al. 2014), just to mention a few among many others). The role of salinity as a physiological stressor has also promoted the study of idiosyncratic ecophysiological adaptations along salinity gradients in simplified ecological communities (Pedrós-Alió, Calderón-Paz and Gasol 2000). Terrestrial saline aquatic environments (both inland and coastal) are excellent natural model systems and have been traditionally associated with highly specialized planktonic organisms. At the most saline ends, dominance of a few groups has been reported, i.e. *Halococcus* spp., *Halobacterium rubrum* and the recently uncovered Nanohaloarchaea (Ghai et al. 2011; Narasimarao et al. 2012; Ventosa et al. 2015). At intermediate salinities, different studies show high levels of heterotrophic activity (Gasol et al. 2004), planktonic algae accumulation (Estrada et al. 2004), and presence of largely unknown organisms (Demergasso et al. 2008; Triadó-Margarit and Casamayor 2013). Among them, inland lakes are prone to contain new groups of Halobacteria and poorly characterized uncultured Archaea (Narasimarao et al. 2012; Casamayor, Triadó-Margarit and Castañeda 2013).

Several studies in the last decade have improved the picture on the Archaea domain, with higher diversity and environmental distribution than historically considered (Auguet, Barberan and Casamayor 2010; Rinke et al. 2013). There are however important gaps still on the current knowledge of the Archaea ecology even in apparently well studied habitats where archaea were predominant organisms, such as high saline systems. Recently, the monophyletic superphylum DPANN has been shown to contain a large number of widespread taxa most of them missing cultured counterparts, such as Nanohaloarchaeota, Pacearchaeota, Woesearchaeota, and Nanoarchaeota (Rinke et al. 2013; Castelle et al. 2015; Ortiz-Alvarez and Casamayor 2016). Because of the difficulties in culturing and the still limited application of both powerful single cell genomic analysis (Dal Molin and Di Camillo 2018) and metagenomics approaches (Narasimarao et al. 2012) in archaea, the biology and ecology of many of these groups are still unknown. In addition, the detailed study of temporal dynamics in archaeal ecological communities has been rarely carried

out (Galand et al. 2010; Hugoni et al. 2013), missing substantial information on the ecology of the third domain of life.

Within this context, we explored community patterns and temporal dynamics of the aquatic archaeal microbiome in the Monegros Desert (NE Spain). The Monegros ecosystem contains one of the largest sets of inland saline lakes in Europe and constitutes a unique landscape of great ecological value, easy to reach and sample (Casamayor, Triadó-Margarit and Castañeda 2013). Physico-chemical parameters strongly change with time in the same aquatic spot offering a wide repertoire of heterogeneous environmental conditions within a very narrow geographic distance of a few kilometers. A preliminary study already evidenced the high microbial novelty present in Monegros, including novel archaea (Casamayor, Triadó-Margarit and Castañeda 2013). In the present study, we aimed to unveil temporal dynamics, ecological patterns and environmental preferences for the most unknown Archaea, covering different inter-annual dry and wet periods. We added a quantitative temporal dimension to the analysis by applying a community ecology method based on temporal changes in the absence/presence of local populations. This method applies a mathematical estimation of 'colonization' and 'extinction' rates (Ontiveros et al. 2019), showing different ecological dynamics for the different groups. The range of habitat gradients occupied and the estimated dynamical behavior, unveiled community level traits and patterns unseen to date for some of the most intriguing uncultured archaeal groups.

EXPERIMENTAL PROCEDURES

The Monegros Desert is an endorheic platform (c. 400 km²) located in the Central Ebro Basin (NE Spain, 41°42'N, 0°20'W, 328 m a.s.l.), with 149 scattered depressions ranging in size from < 2 to 239 ha, some of them hosting hypersaline wetlands. This is one of the most arid regions in Europe (mean annual rainfall 350 mm, mean annual temperature 14.9 °C, and mean reference evapotranspiration 1255 mm yr⁻¹, Faci González and Martínez-Cob 1991), and unique for the combination of geological, edaphic, mineralogical and hydrological characteristics (Pueyo Mur and Inglés Urpinell 1987; Samper-Calvete and García-Vera 1998; Mees, Castañeda del Álamo and Ranst 2011). We carried out monthly sampling in 14 different ponds (Table 1) covering different dry-wet periods with a total

Table 1. Range of environmental parameters measured in the different salt lakes in Monegros.

Lagoon	Salinity (% w/v)	Conductivity (mS/cm)	Temperature (°C)	pH	OD (mg/l)	Redox (mv)	Chla (µg/l)
Guallar	2.8–23	26.2–197.6	2–26.5	7.6–8.4	8.8–12.8	145.5–350	0.9–10
Pinol	1.2–8.4	14.4–103	5–25.1	8–8.8	8.9–13.3	187.1–315	0.9–27.9
Muerte	2–32	20.3–207.4	7.5–32	7.3–8.61	7.4–12.8	224–312	0.9–5.2
Camaron	1.2–15.4	14.2–164.7	2.3–29.5	8.0–9.0	8.8–14.1	220–305	0.5–46
Pez	1.1–13.8	8.5–165.5	2–31	7.8–8.42	8.4–13.1	200–320	1.1–6.7
Rollico	0.8–5.8	5.6–77.6	8.7–21.5	8–8.8	9.1–12.8	205–345	0.7–13.7
Rebollon	1–2.4	5.1–27.2	7.5–14.8	8.1–9.1	10.4–14.6	225–340	1–27.1
Pito	1.2–28	12–94.1	1.5–34.5	7.5–8.3	7.5–12.7	240–335	1.1–6.6
Salineta	6.9–40	66.3–167.9	7.4–34	7.7–9.2	7.8–15.7	188.6–315	2.4–237.4
Playa	2.1–21	26.3–194.5	4–3	8–8.8	7.3–12.4	144–312	0.4–29.6
Amarga Alta	0.9–1.7	3.6–15.6	3–15	7.9–8.4	10.9–12.5	248–326	0.5–6.8
Amarga Baja	0.9–1.8	4.3–15.6	1.5–11.2	7.9–9.2	10.5–13.1	242–339	0.7–11.4
Salobral	2.9–3.8	31.3–47.1	5.9–30.3	8.8–9.5	8.1–14.2	122.6–341	1.7–15.5
Balsa Molinos	0.1–0.3	0.7–0.9	8–17.4	8.3–9.1	9.8–12.1	195–240	3–9.7

of 128 water samples analysed (see details in Table S1, Supporting Information). Conductivity, temperature, pH, dissolved oxygen and redox potential were measured *in situ* using an HQ40d multiparameter meter (Hach, Loveland, CO, USA). For chlorophyll *a* (Chl *a*) analysis 1 L water was filtered through 47-mm-diameter Whatman GF/F filters (0.7 µm nominal particle retention), stored in the dark, and kept frozen. Chl *a* concentration was determined in acetone extracts by spectrophotometry (UV-2401PC, ultraviolet-visible Spectrometer; Shimadzu) as previously reported (Bartrons, Catalan and Casamayor 2012). Salinity was measured *in situ* by a hand salinity refractometer (Atago S-28E, Japan). The water column height was measured at the deepest part of the ponds. Substantial changes in water column height, temperature, salinity and other environmental variables were observed both among ponds and inside the same pond along the hydrological period. Detailed information about environmental parameters and sampling date and sites is available in Table S1 (Supporting Information).

DNA extraction and sequencing

For DNA analyses, water samples were pre-filtered *in situ* through a 50 micron-pore-size net, to retain large zooplankton and algae, and then 100–500 mL were subsequently filtered on 5 and 0.2-micron pore-size polycarbonate membranes (47 mm diameter), respectively. The 0.2 µm PC membranes were stored in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose), enzymatically digested and phenol extracted (phenol-chloroform-isoamyl alcohol; 25:24:1, v/v/v) as previously reported (Ferrera et al. 2004). DNA purification and concentration was carried out with Amicon® Ultra 4 Centrifugal Filter Units–100 000 NMWL (Millipore). DNA extracts were quantified using Quant-it assays in a Qubit™ fluorometer (Invitrogen). Sequencing with the Illumina 2 × 250 MiSeq platform was carried out on PCR amplicons of the 16S rRNA gene (V4 region, primers set F515–R806, Caporaso et al. 2010) according to genomic core facilities, protocols and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>).

Bioinformatic processing and genetic analyses

Raw sequences were processed using UPARSE (Edgar 2013). The total number of sequences prior to quality filtering was 14 445 263 and 70.6% of them had ≥ Q30. After merging of read pairs, the dataset was filtered by setting a maximum number of expected errors of 0.25 (Edgar and Flyvbjerg 2015), and chimeric reads were detected and filtered out by UCHIME following both *de novo* and reference-based analyses (against the ‘Gold’ reference database available for ChimeraSlayer). Filtered sequences were clustered into operational taxonomic units (OTUs) at 97% identity. The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse, Peplies and Glöckner 2012), using the SILVA 128 database (Gurevich et al. 2013). OTUs were filtered by their alignment quality score and sequences classified as mitochondria and chloroplasts were excluded. A final filtering by library size keeping only samples with > 10,000 reads was carried out. The resulting OTU table consisted of 7241 prokaryotes OTUs (6251 bacterial OTUs and 990 archaeal OTUs), corresponding to a total of 3374 654 reads (of which 54 807 were assigned to Archaea) for 128 samples. Data was normalized by rarefying the reads of all samples to a minimum threshold of 10 000 reads/sample. For community related analysis, rarefaction at the same depth was repeated 100 times in order to avoid the

loss of less abundant OTUs, and the resulting rarefactions were unified on an average rarefied OTU table.

Genetic distances between OTUs were calculated using the ‘*calc.distmx*’ command in the USEARCH pipeline. Reference sequences were included to properly differentiate Candidatus Pacearchaeota (i.e. GenBank accession numbers MFWP01000008, MNVW01000021, MNVX01000029 and MNVY01000002, Anantharaman et al. 2016) from Candidatus Woesearchaeota (accession number MNWW01000056, Nawrocki, Jones and Eddy 2018). The resulting distance matrix was visualized by Non-metric multi-dimensional scaling (NMDS) analysis run with MetaMDS functions in Vegan package. We estimated genetic diversity by calculating the nucleotide diversity using the R package ‘Pegas’ (<https://cran.r-project.org/web/packages/pegas/index.html>). The phylogenetic tree shown in Fig. S1 (Supporting Information), is an ARB maximum parsimony tree provide by default (<http://www.arb-home.de/>) with 16S rRNA gene amplicon sequences added using the tool ‘add species to existed tree / ARB-parsimony tool interactive’, providing only very rough information because of the short read length.

The complete raw data set is available in GenBank under Bio-Project record ID PRJNA429605, and GenBank accession numbers for highlighted OTUs are shown in Table S2 and Fig. S1 (Supporting Information).

Statistical analysis and calculation of colonization and extinction rates

Statistical analyses were run in R (<http://www.r-project.org>). Species richness, evenness and diversity indexes were calculated using ‘Vegan’ package version 2.4–5 (Oksanen et al. 2016). The environmental amplitude (range of environmental conditions occupied) for the different archaeal lineages, was calculated according to salinity, pH, redox potential, dissolved oxygen and Chl *a* data, weighting its variance with the relative abundance of the lineage. We avoided the use of variables that co-varied with salinity such as conductivity (Spearman, rho 0.98 and *P* value < 0.001) and temperature (Spearman, rho 0.50 and *P* value < 0.001). Spearman’s Rank order correlation coefficient was used to test monotonic associations between numeric variables.

The estimated temporal ‘turnover’ rates were calculated for each archaeal lineage based on the absence/presence of local populations for consecutive sampling times (‘extinction / colonization’ rates, respectively). ‘Colonization’ and ‘extinction’ rates were calculated using the ‘island’ package on R (<https://CRAN.R-project.org/package=island>; Ontiveros et al. 2019) and provide an effective view on the dynamic behavior of the different archaeal populations in time (‘turnover’ rate). The ‘turnover’ rates were expressed as day^{−1}.

RESULTS

Archaeal abundance and diversity

The relative abundance of Archaea in the 16S rRNA gene amplicon mixture ranged between 0.01 and 50% of total prokaryotes. Archaea abundance consistently increased along the salinity gradient, mainly at salinities > 20% (w/v), reaching the highest relative abundances at the highest salinity end (Fig. 1). For salts concentrations <20%, relative abundances were < 5% in all cases.

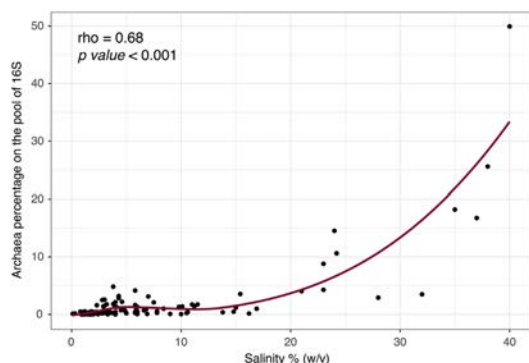


Figure 1. Relationship between archaeal percentage in the prokaryotic 16S rRNA amplicon mixture and salinity (% w/v) for the complete dataset analyzed.

The genetic richness of the Archaea planktonic microbiome was notorious. We observed the presence of 14 different lineages, with occurrences (percentage of samples in which the group was detected) between 96% and 0.8% (Table 2). As a whole, the OTUs richness was higher for Woesearchaeota (402 OTUs) and Pacearchaeota (376 OTUs), followed by Halobacteria (105 OTUs) and Nanohaloarchaeota (62 OTUs). The highest values of abundance in the amplicon mixture were found for both Nanohaloarchaeota (24%) and Halobacteria (22%). Woesearchaeota and Pacearchaeota did not reach relative abundances >5% for any of the samples, although they were widely distributed being present in >87% of the total dataset. Halobacteria were also found in a large percentage of samples (82%), followed by Nanohaloarchaeota (49% occurrence). Minor groups were Thermoplasmata, Thaumarchaeota and Aenigmarchaeota, among others (Table 2). Only six reads (1 OTU) matched methanogens (i.e. Methanomicrobia, genus *Methanohalophilus*).

Archaeal assembly along the salinity gradient

We selected the seven most abundant lineages and with occurrences >15%, for a more detailed downstream analysis. These groups showed a distinctive distribution along the salinity gradient (Fig. 2). At salinities below 15‰, the most abundant groups were mainly Woesearchaeota, and in one case we detected noticeable concentrations of both Thermoplasmata (3% relative abundance in Lake Muerde on January 2013, salinity 5.8‰, temperature 7.8°C; OTU 316 in Table S2, Supporting Information) and Pacearchaeota (2% relative abundance in Lake Salobral on July 2013, salinity 3.8‰, temperature 29.6°C). Conversely, at > 15‰ salinity the archaeal assemblages were mostly dominated by Nanohaloarchaeota (up to 24% in Lake Salineta, October 2013, salinity 40‰, temperature 18.8°C), Halobacteria, and Pacearchaeota. Finally, Thaumarchaeota (ammonia-oxidizing archaea, AOA) showed a maximum abundance of 0.23% at marine like salinities (close to 3‰ w/v) and were mostly detected at salinities up to 10‰ (8 OTUs, accession numbers MH028008 to MH028015) indicating a strong selection by local environmental conditions on this functional group.

Overall, salinity was the environmental variable which explained most of the variations in the relative abundances of the different archaeal populations and not major effects were observed for the additional environmental/biological parameters measured ($P > 0.05$ for pH, redox potential, dissolved oxygen, Chl *a* concentration, and bacterial diversity). We found

a strong correlation between salinity and relative abundances of Pacearchaeota (Spearman, $\rho = 0.59$, P value < 0.001) and Woesearchaeota (Spearman, $\rho = 0.65$, P value < 0.001). This correlation was also significant for Nanohaloarchaeota (Spearman, $\rho = 0.36$, P value < 0.01) and Halobacteria (Spearman, $\rho = 0.41$, P value < 0.001). The robustness of the correlations increased when the whole salinity gradient was divided into the three main salinity ranges based on both physicochemical and biological parameters (Por 1972; Kushner 1978; Imhoff 2001), i.e. <5‰ (saline), 5–15‰ (hypersaline), and >15‰ (extreme hypersaline), and some unseen trends were subsequently unveiled (Fig. 3). Nanohaloarchaeota and Halobacteria showed stronger correlations (c. $\rho = 0.70$) at the most saline end, whereas the relationship was non-significant at <15‰ salinity. Woesearchaeota only showed a significant positive relationship at salinities <5‰. Conversely, relative abundances of Pacearchaeota showed an unexpected positive correlation with salinity in the highly hypersaline range (>15‰ salinity, Fig. 3). These results unveiled that these two highly unknown taxa tend to behave differently in front of salinity. Candidatus Pacearchaeota and Woesearchaeota were the richest Archaea taxa in Monegros and showed the highest genetic variation along the salinity gradient in comparison with the remaining Archaea populations (Fig. 4). Although both had been grouped within the Deep-sea Hydrothermal Vent Euryarchaeota Group 6 (DHVEG-6) in former taxonomic arrangements, Pacearchaeota and Woesearchaeota appeared clearly separated from each other in the NMDS ordination according to its nucleotide sequence composition (Fig. 4).

A closer look to the most saline end of the gradient (>35‰ salinity) showed that Halobacteria at these salinities were spread in 12 already known genera, whereas most of the Nanohaloarchaeota were only distantly related to any previously reported counterpart (Fig. S1, Supporting Information, with available individual accession numbers shown). Interestingly, two of the most abundant taxa in the most saline sample (Lake Salineta, October 2013, 40‰ salinity) were the Halobacteria OTU 320 (97% identity to *Halonotius pteroides*) and the Nanohaloarchaeota OTU 289 (100% identical to Candidatus *Haloredivivus* G17 AGNT01000024) previously found in Santa Pola solar salterns in the Spanish Mediterranean coast (Ghai et al. 2011), approx. 450 km away from Monegros.

Archaeal ecology and dynamic behaviour

Community ecology analyses provided new insights and unveiled unseen ecological trends for the different Archaea groups. As a whole, group occurrence was strongly correlated with OTUs richness (Spearman, $\rho = 0.96$, P value < 0.005), i.e. the highest occurrence was found for the richest groups. In addition, the distribution of genetic richness (i.e. nucleotide diversity) for the seven major archaeal phyla was strongly correlated with environmental amplitude (Spearman, $\rho = 0.89$, P value < 0.05, Fig. 5, panel B) but not with lake occurrence (Fig. 5, panel A). Aenigmarchaeota had the highest environmental amplitude (i.e. larger range of habitat gradients occupied) despite the fact that only 5 OTUs were detected for this phylum, and showed the lowest occurrence (19%). Conversely, Nanohaloarchaeota, Halobacteria and Thaumarchaeota had the bulk of their abundances within a restricted environmental niche (Fig. 5, panel B), despite having a medium-high occurrence (between 35% and 85%). The analysis of OTU richness vs. evenness also showed interesting trends (Fig. 5,

Table 2. OTU richness, occurrence and abundances for the archaeal lineages found in Monegros.

Archaeal lineage	OTUs	Occurrence (%)	Max. abundance (%)*
EURY: Halobacteria	105	105 (82,03%)	22.08
DPANN: Woesearchaeota	402	123 (96,09%)	3.34
DPANN: Nanohaloarchaeota	62	63 (49,22%)	23.94
DPANN: Pacearchaeota	376	112 (87,5%)	4.98
EURY: Thermoplasmata	19	30 (23,44%)	2.99
DPANN: Aenigmarchaeota	5	24 (18,75%)	0.35
TACK: Thaumarchaeota	8	44 (34,38%)	0.23
Archaea unclass.	6	14 (10,94%)	0.27
Asgard: Lokiarchaeota	1	5 (3,91%)	0.04
EURY: Methanomicrobia	1	3 (2,34%)	0.04
EURY: Unclass.	2	5 (3,91%)	0.01
EURY: WSA2	1	2 (1,56%)	0.02
EURY: Archaeoglobi	1	3 (2,34%)	0.01
EURY: Miscellaneous Group (MEG)	1	1 (0,78%)	0.01

*Relative abundances in the whole prokaryotic amplicon mixture.

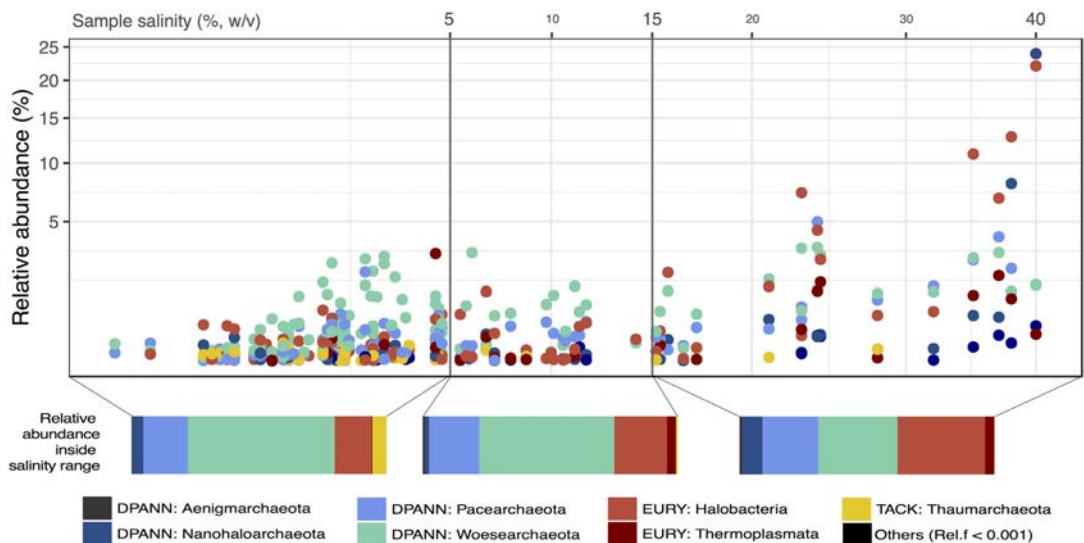


Figure 2. Relative abundances of the main archaeal lineages along the salinity gradient, and averaged values for the three traditional saline ranges, i.e. <5‰ (saline), 5–15‰ (hypersaline), and >15‰ (extreme hypersaline), respectively.

panel C); Pacearchaeota and Woesearchaeota showed the highest richness and evenness, whereas at the other end, Aenigmarchaeota, Thermoplasmata, Thaumarchaeota, and Nanohaloarchaeota were low richness-low evenness groups. Halobacteria was the only group with low richness-high evenness structure (i.e. relatively low number of OTUs and with similar population density). Finally, the estimated dynamic behaviour ('turnover' rates) for the main archaeal lineages also showed important differences among groups (Fig. 5, panel D). The most dynamic groups (the highest 'colonization/extinction' turnover) were Thaumarchaeota and Halobacteria. The most stable taxa (the lowest turnover) were Aenigmarchaeota and Thermoplasmata. The DPANN groups Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate turnover rates between the two ends.

DISCUSSION

Comprehensive studies carried out in the last years have provided a better picture of the Archaea ecology and diversity (e.g. Auguet, Barberan and Casamayor 2010; Rinke et al. 2013). There are however important gaps still in the current knowledge of the environmental Archaea microbiome. For instance, the widespread DPANN group has been recently arranged (Castelle et al. 2015; Spang, Caceres and Ettema 2017), but little is known on the ecology, dynamics and environmental preferences of these archaea. Our study added relevant information analysing the ecological niche, range of environmental preferences, and temporal dynamics of uncultured DPANN and other relevant archaea, using a detailed temporal survey in a high dynamic set of inland salt lakes.

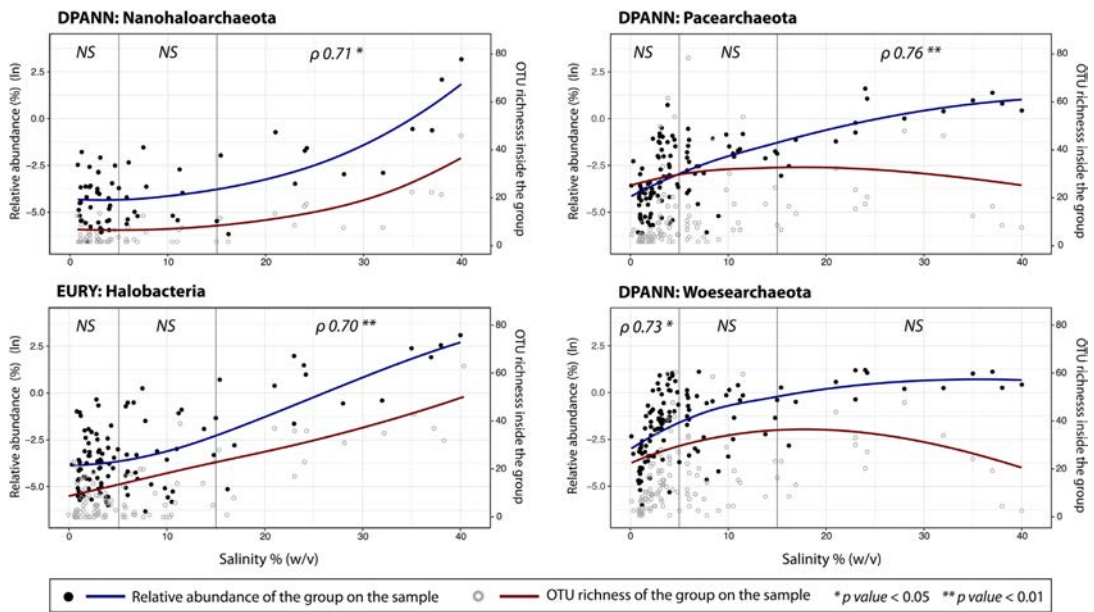


Figure 3. Relative abundances and OTU richness for the four most abundant groups (Woesearchaeota, Pacearchaeota, Nanohaloarchaeota, Halobacteria) along the salinity gradient. Correlations shown for the three traditional saline ranges, i.e. <5% (saline), 5%–15% (hypersaline), and >15% (extreme hypersaline), respectively. Spearman's rho correlation and p value are shown when significant; NS = non-significant. The smooth fitted curves from these parameters were obtained by means of Local Regression (loess) approach.

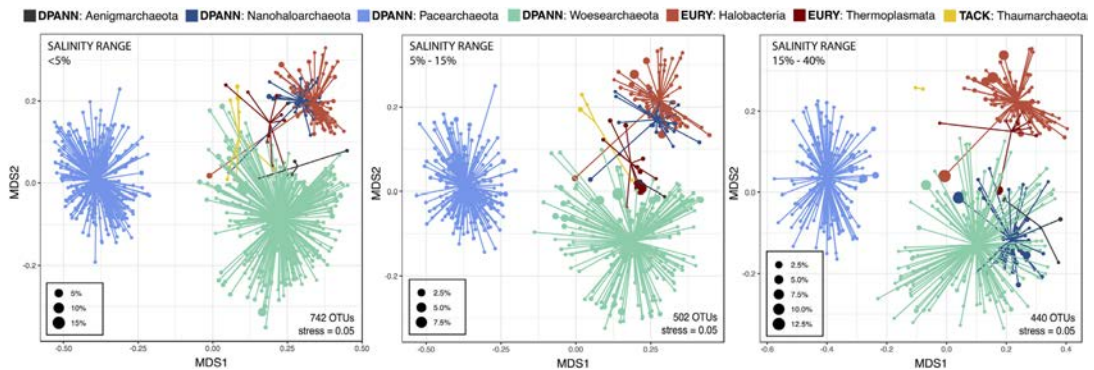


Figure 4. NMDS representation based on the genetic distances among OTUs in the three main salinity ranges (<5% saline, 5%–15% hypersaline, and >15% extreme hypersaline). Each OTU from the same lineage is connected in the two-dimensional space through the centroid. Point sizes show relative abundances.

We observed higher archaeal relative abundance at the most extreme salinity end, mostly of Halobacteria (superphylum Euryarchaeota) but also of the DPANN Nanohaloarchaeota. The importance of Halobacteria (haloarchaea) at extreme salinities have been recurrently reported (Oren 2002). However, the role and distribution of Nanohaloarchaeota have been unseen for a long time (Narasimarao et al. 2012). Its recent detection in diverse hypersaline environments, has arisen a high interest for both plankton and benthic Nanohaloarchaeota populations (Podel et al. 2014; Andrei et al. 2015; Cowan et al. 2015; Ventosa et al. 2015). In our study, we observed similar relative abundances of Nanohaloarchaeota and Halobacteria at the most saline end, although with substantial differences in community structure

and community ecology dynamics. Nanohaloarchaeota showed a more limited environmental breadth, being less abundant than Halobacteria at lower salinities, and lower colonization rates. More studies are needed to confirm whether or not this observation indicates a higher specialization of the DPANN Nanohaloarchaeota to the most extreme saline conditions compared to a more versatile behavior of Halobacteria. Recent discoveries on a symbiotic lifestyle for Nanohaloarchaeota (Hamm et al. 2019) indicate that it is necessary to apply an integrated view on the whole microbial community (including bacteria and eukarya) to understand Nanohaloarchaeota ecology better.

A symbiotic/parasitic lifestyle has also been proposed for the other two important DPANN groups (in terms of relative

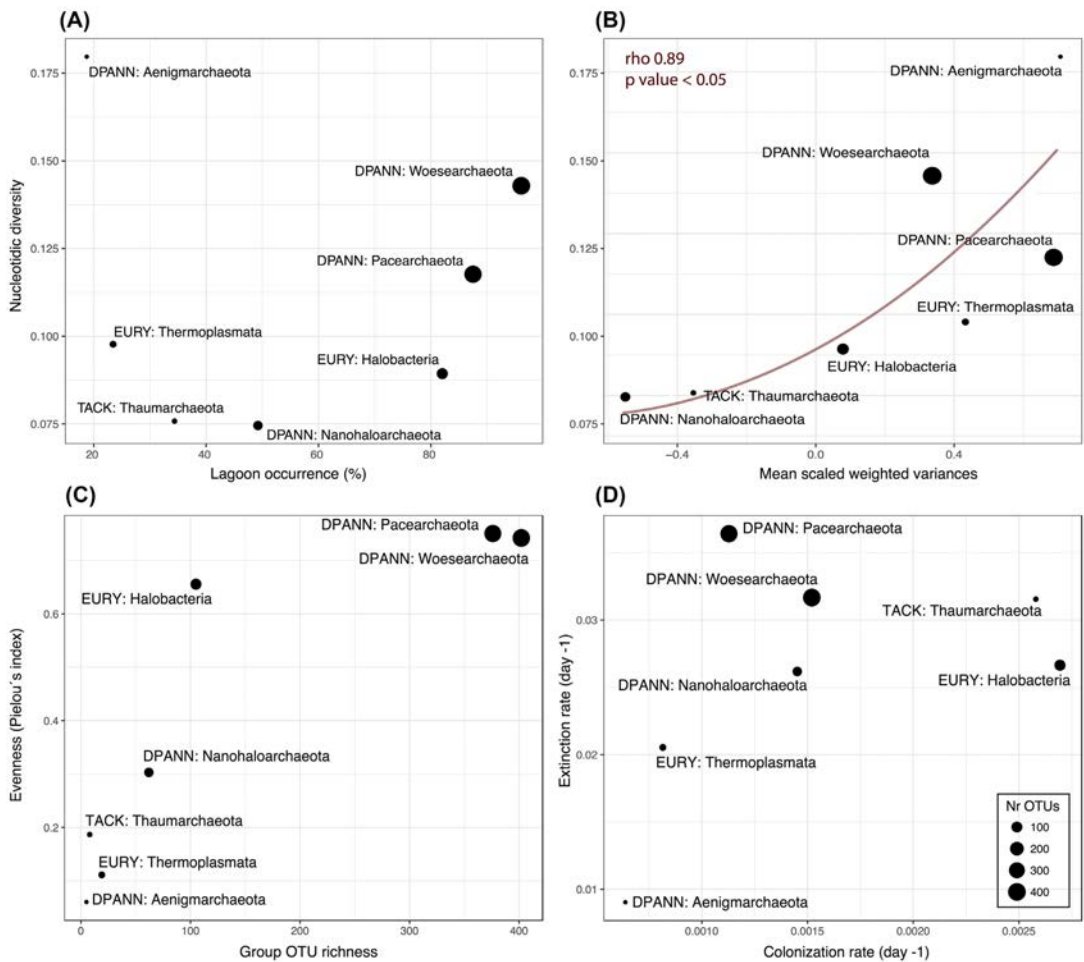


Figure 5. Community ecology descriptors for the main phyla of the Archaea assemblage in Monegros saline lakes. Nucleotidic diversity related to both lagoon occurrence (A) and environmental amplitude (B). Spearman's rho correlation shown when significant. Evenness (Pielou's index) based on abundance distribution related to group OTUs richness (C). 'Colonization and extinction' turnover rates (day⁻¹) based on the absence/presence of the local communities in consecutive sampling times (D). Point sizes show OTU richness.

abundance and occurrence) in the Monegros Archaea microbiome: Woesearchaeota and Pacearchaeota. Woesearchaeota and Pacearchaeota (both formerly named Deep-sea Hydrothermal Vent Euryarchaeota Group 6, DHVEG-6) have been recently proposed as candidate phyla belonging to the DPANN superphylum (Castelle et al. 2015). Since then, they have been uncovered as major components of the archaeal assemblages in a large repertoire of environments, and the need of a better understanding of their biology, ecology and distribution has stimulated research in recent years (Castelle et al. 2015; Ortiz-Alvarez and Casamayor 2016; Liu et al. 2018). The lack of an unequivocal phylogenetic split between Woesearchaeota and Pacearchaeota has led in some cases to overlapping analysis (Lipsewiers et al. 2018; Liu et al. 2018) and blurred ecological differentiation between them. The environmental niches of these two phyla are known to be large and they are widely distributed in saline aquatic

environments, mainly in sediments, microbial mats and plankton, but also in hot springs and freshwater (Robertson et al. 2009; Kan et al. 2011; Pachiadaki et al. 2011). Pacearchaeota has only been rarely detected in marine plankton (Castelle and Banfield 2018), which matches the observed trend in Monegros where higher relative abundances of Pacearchaeota were found at higher salinities. In the studied saline lakes, these two phyla showed the highest OTU richness, evenness, site occurrence and extinction rates among Archaea. All these traits suggest an ecological strategy based on fast colonization but low persistence along the whole salinity gradient for the different OTUs shaping these two groups. As previously mentioned, new insights in the metabolisms of DPANN archaea brought by metagenomics and single-cell genomics, suggest that small genomes, small cell sizes, and limited metabolic variability are common characteristics. Thus, the lack of core biosynthetic pathways

has led to conclude that most of the DPANN archaea depend on other microorganisms—as a symbionts or episymbionts—to carry out basic metabolic functions (Castelle and Banfield 2018). In a recent study in alpine lakes, relative abundances of Pacearchaeota and Woesearchaeota were observed to be positively correlated with phylogenetic diversity of bacterial communities, respectively (Ortiz-Alvarez and Casamayor 2016). This correlation was however not observed in Monegros. If this correlation is expected to be consistent and somehow related to potential bacterial symbionts, probably their high occurrence along the salinity gradient and their large environmental amplitude may have masked the detection of this potential signature in Monegros. Certainly, this point deserves further and more detailed studies considering more integrated analyses of the whole microbial community.

For the first time, our study quantifies the *in situ* dynamic behavior of uncultured archaea. This new perspective needs of experimental tests to be further substantiated. The more 'dynamic' taxa were Thaumarchaeota (limited to salinities below 10‰, in agreement with previous studies, Fernandez-Guerra and Casamayor 2012) and Halobacteria (with the bulk of its abundance at extreme salinities), and showed their optimal within a restricted environmental niche suggesting these are groups with a higher cell sensitivity to salinity changes. Conversely, the DPANN groups Pacearchaeota, Woesearchaeota, and Nanohaloarchaeota showed intermediate turnover rates. Pacearchaeota and Woesearchaeota showed the widest environmental breadth, the highest richness, and the highest genetic variability. These results, when contextualized in a highly dynamic system with quick changes in community composition mainly driven by changes in salinity (Triadó-Margarit et al. 2019), suggest a higher resilience to environmental perturbations.

Overall, the study carried out in inland salt lakes of Monegros showed a rich and dynamic Archaea microbiome including relevant members of the still largely unknown DPANN group. We unveiled ecological properties for some of the most enigmatic DPANN members as Nanohaloarchaeota, Pacearchaeota, Woesearchaeota, and Aenigmarchaeota, and provided further evidence to support strong ecological differentiation between Pacearchaeota and Woesearchaeota. Further studies both *in situ* and in combination with traditional microbiological approaches, should help to improve our still very limited understanding of the biology and ecology of the monophyletic superphylum DPANN, and on the ecology of inland saline systems in general.

DATA ACCESSIBILITY

DNA sequences. The complete raw data set is available in GenBank under BioProject record ID PRJNA429605.

SUPPLEMENTARY DATA

Supplementary data are available at [FEMSEC](#) online.

ACKNOWLEDGEMENTS

Funding was provided by grant BRIDGES, CGL2015-69043-P from the Spanish Office of Science (MINECO) and European funding (ERDF) to EOC and DA. MMS was funded by the Spanish FPI PhD scholarships program (BRIDGES-MINECO). VJO was funded by the Spanish FPI PhD scholarships program (SITES-MINECO).

Conflicts of interest. None declared.

REFERENCES

- Anantharaman K, Brown CT, Hug LA et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* 2016;7:13219.
- Andrei AS, Robeson MS, Baricz A et al. Contrasting taxonomic stratification of microbial communities in two hypersaline meromictic lakes. *ISME J* 2015;9:2642–56.
- Auguet JC, Barberan A, Casamayor EO. Global ecological patterns in uncultured Archaea. *ISME J* 2010;4:182–90.
- Bartrons M, Catalan J, Casamayor EO. High bacterial diversity in epilithic biofilms of oligotrophic mountain lakes. *Microb Ecol* 2012;64:860–9.
- Caporaso JG, Kuczynski J, Stombaugh J et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* 2010;7:335–6.
- Casamayor EO, Triadó-Margarit X, Castañeda C. Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. *FEMS Microbiol Ecol* 2013;85:503–18.
- Castelle CJ, Banfield JF. Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell* 2018;172:1181–97.
- Castelle CJ, Wrighton KC, Thomas BC et al. Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* 2015;25:690–701.
- Cavalazzi B, Barbieri R, Gómez F et al. The dallol geothermal area, northern afar (Ethiopia)—An Exceptional planetary field analog on earth. *Astrobiology* 2019;19:553–78.
- Cowan D, Ramond J-B, Makhalanyane T et al. Metagenomics of extreme environments. *Curr Opin Microbiol* 2015;25:97–102.
- Dal Molin A, Di Camillo B. How to design a single-cell RNA-sequencing experiment: pitfalls, challenges and perspectives. *Brief Bioinform* 2018;20: bby007:1–11.
- Demergasso C, Escudero L, Casamayor EO et al. Novelty and spatio-temporal heterogeneity in the bacterial diversity of hypersaline Lake Tebenquiche (Salar de Atacama). *Extremophiles* 2008;12:491–504.
- Dillon J, McMath LM, Trout AL. Seasonal changes in bacterial diversity in the Salton Sea. *Hydrobiologia* 2009;632:49–64.
- Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Meth* 2013;10:996–8.
- Edgar RC, Flyvbjerg H. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 2015;31:3476–82.
- Estrada M, Henriksen P, Gasol JM et al. Diversity of planktonic photoautotrophic microorganisms along a salinity gradient as depicted by microscopy, flow cytometry, pigment analysis and DNA-based methods. *FEMS Microbiol Ecol* 2004;49:281–93.
- Faci González JM, Martínez-Cob A. Cálculo de la evapotranspiración de referencia en Aragón. *Diputacion General de Aragon*. 1991, 115.
- Ferrera I, Massana R, Casamayor EO et al. High-diversity biofilm for the oxidation of sulfide-containing effluents. *Appl Microbiol Biotechnol* 2004;64:726–34.
- Fernández-Guerra A, Casamayor EO. Habitat-Associated phylogenetic community patterns of microbial ammonia oxidizers. *PLoS One* 2012;7:e47330.
- Galand P, Gutiérrez-Provecho C, Massana R et al. Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnol Oceanogr* 2010;55:2117–25.
- Gasol J, Casamayor E, Joint I et al. Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat Microb Ecol* 2004;34:193–206.

- Ghai R, Pašić L, Fernández AB et al. New abundant microbial groups in aquatic hypersaline environments. *Sci Rep* 2011;1:135.
- Gurevich A, Saveliev V, Vyahhi N et al. QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–5.
- Hamm JN, Erdmann S, Eloe-Fadrosh EA et al. Unexpected host dependency of antarctic nanohaloarchaeota. *Proc Natl Acad Sci* 2019;116:14661 LP–14670.
- Hugoni M, Taib N, Debroas D et al. Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci* 2013;110:6004 LP–6009.
- Imhoff JF. True marine and halophilic anoxygenic phototrophic bacteria. *Arch Microbiol* 2001;176:243–54.
- Kan J, Clingenpeel S, Macur RE et al. Archaea in yellowstone lake. *ISME J* 2011;5:1784.
- Kushner DJ. Life in high salt and solute concentrations. In: Kushner DJ (ed.). *Microbial Life in Extreme Environments*. London: Academic Press, 1978, 317–68.
- Lipsewiers YA, Hopmans EC, Sinninghe Damsté JS et al. Potential recycling of thaumarchaeotal lipids by DPANN Archaea in seasonally hypoxic surface marine sediments. *Org Geochem* 2018;119:101–9.
- Liu X, Li M, Castello CJ et al. Insights into the ecology, evolution, and metabolism of the widespread Woese archaeal lineages. *Microbiome* 2018;6:102.
- Mees F, Castañeda del Álamo C, van Ranst E. Sedimentary and diagenetic features in saline lake deposits of the Monegros region, northern Spain. *Catena* 2011;85:245–52.
- Narasimgarao P, Podell S, Ugalde JA et al. De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J* 2012;6:81–93.
- Nawrocki EP, Jones TA, Eddy SR. Group I introns are widespread in archaea. *Nucleic Acids Res* 2018;46:7970–6.
- Oksanen J, Kindt R, Legendre P et al. *vegan* Community Ecology Package version 2.4-1 September 2016. 2016.
- Ontiveros VJ, Capitan JA, Arthur R et al. Colonization and extinction rates estimated from temporal dynamics of ecological communities: The island R Package. *Meth Ecol Evol* 2019;10:1108–17.
- Oren A. Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. *J Ind Microbiol Biotechnol* 2002;28:56–63.
- Ortiz-Alvarez R, Casamayor EO. High occurrence of Pacearchaeota and Woese archaeota (Archaea superphylum DPANN) in the surface waters of oligotrophic high-altitude lakes. *Environ Microbiol Rep* 2016;8:210–7.
- Pachiadaki MG, Kallionaki A, Dählmann A et al. Diversity and spatial distribution of prokaryotic communities along a sediment vertical profile of a Deep-Sea mud volcano. *Microb Ecol* 2011;62:655–68.
- Pedrés-Alió C, Calderón-Paz JI, Gasol JM. Comparative analysis shows that bacterivory, not viral lysis, controls the abundance of heterotrophic prokaryotic plankton. *FEMS Microbiol Ecol* 2000;32:157–65.
- Podell S, Emerson JB, Jones CM et al. Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. *ISME J* 2014;8:979–90.
- Por FD. Hydrobiological notes on the high salinity waters of the Sinai Peninsula. *Mar Biol* 1972;14:111–9.
- Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;28:1823–9.
- Pueyo Mur JJ, Inglés Urpinell M. Magnesite formation in recent playa lakes, Los Monegros, Spain. *Geol Soc London, Spec Publ* 1987;36:119 LP–122.
- Rinke C, Schwientek P, Sczyrba A et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 2013;499:431–7.
- Robertson CE, Spear JR, Harris JK et al. Diversity and stratification of archaea in a hypersaline microbial Mat. *Appl Environ Microbiol* 2009;75:1801–10.
- Samper-Calvete FJ, García-Vera MA. Inverse modeling of groundwater flow in the semiarid evaporitic closed basin of Los Monegros, Spain. *Hydrogeol J* 1998;6:33–49.
- Spang A, Caceres EF, Ettema TJG. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* 2017;357:563.
- Triadó-Margarit X, Casamayor EO. High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. *FEMS Microbiol Ecol* 2013;85: 27–36.
- Triadó-Margarit X, Capitan JA, Menéndez-Serra M et al. A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly. *ISME J* 2019;13:2681–9.
- Ventosa A, Fernández AB, León MJ et al. The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles* 2014;18:811–24.
- Ventosa A, de la Haba RR, Sánchez-Porro C et al. Microbial diversity of hypersaline environments: A metagenomic approach. *Curr Opin Microbiol* 2015;25:80–87.
- Wu QL, Zwart G, Schauer M et al. Bacterioplankton Community Composition along a Salinity Gradient of Sixteen High-Mountain Lakes Located on the Tibetan Plateau, China. *Appl Environ Microbiol* 2006;72:5478–85.
- Williams WD. The largest, highest and lowest lakes of the world: Saline lakes. *SIL Proceedings, 1922–2010* 1996;26:61–79.
- Williams TJ, Allen MA, DeMaere MZ et al. Microbial ecology of an Antarctic hypersaline lake: genomic assessment of ecophysiology among dominant haloarchaea. *ISME J* 2014;8: 1645–58.



Ecological and Metabolic Thresholds in the Bacterial, Protist, and Fungal Microbiome of Ephemeral Saline Lakes (Monegros Desert, Spain)

Mateu Menéndez-Serra¹ · Xavier Triadó-Margarit¹ · Emilio O. Casamayor¹

Received: 31 January 2021 / Accepted: 8 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

We studied the 16S and 18S rRNA genes of the bacterial, protist, and fungal microbiomes of 131 samples collected in 14 ephemeral small inland lakes located in the endorheic area of the Monegros Desert (NE Spain). The sampling covered different temporal flooding/desiccation cycles that created natural salinity gradients between 0.1% (w/v) and salt saturation. We aimed to test the hypothesis of a lack of competitive advantage for microorganisms using the “salt-in” strategy in highly fluctuating hypersaline environments where temperature and salinity transitions widely vary within short time periods, as in ephemeral inland lakes. Overall, 5653 bacterial zOTUs and 2658 eukaryal zOTUs were detected heterogeneously distributed with significant variations on taxonomy and general energy-yielding metabolisms and trophic strategies along the gradient. We observed a more diverse bacterial assembly than initially expected at extreme salinities and a lack of dominance of a few “salt-in” organisms. Microbial thresholds were unveiled for these highly fluctuating hypersaline environments with high selective pressures. We conclude that the extremely high dynamism observed in the ephemeral lakes of Monegros may have given a competitive advantage for more versatile (“salt-out”) organisms compared to those better adapted to stable high salinities usually more common in solar salterns. Ephemeral inland saline lakes offered a well-suited natural framework for highly detailed evolutionary and ecological studies.

Keywords Bacteria · Eukaryotes · Inland saline lakes · Community ecology · Halophiles · 16S rRNA gene · 18S rRNA gene · Salt gradient · Monegros desert

Introduction

Life in saline aquatic environments has adapted to cope with high osmotic stress, and salinity has been repeatedly described as one of the main ecological factors affecting microbial community composition at the global scale [1, 2]. The progressive decrease of biodiversity as salinity increases is also usually accompanied by a substantial loss in metabolic functions [3], which shapes important ecological changes along the salt gradient [4]. Osmotic regulation becomes crucial to halophile and halotolerant microorganisms to avoid water deficit and

plasmolysis, and a wide array of strategies are adopted to cope with changes in salt concentrations [5, 6]. At salts concentrations > 0.3% freshwaters turn saline [7], and marine-like salinities can be usually found ranging between 3% and 5% salts w/v in inland waters [8, 9]. At higher salts concentrations, water becomes hypersaline [10], and an additional threshold is set at 15% based on the biological differentiation between moderate and more extreme halophiles [9, 11, 12].

One of the strategies used by cells to reach osmotic equilibrium between inside and outside environments is the accumulation of inorganic salts in the cytoplasm, actively “pumped in” from the saline surroundings [13]. This salt-in strategy is adopted by several of the organisms living at the most extreme hypersaline conditions (e.g., *Halobacteria*, *Halanaerobiales*, and *Salinibacter*), and requires of permanent adaptation in the intracellular machinery, which, in turn, limits further adaptation back to low salinities [14]. A second strategy is the accumulation of compatible solutes, small

✉ Emilio O. Casamayor
casamayor@ceab.csic.es

¹ Integrative Freshwater Ecology Group, Centro de Estudios Avanzados de Blanes (CEAB-CSIC), Accés Cala Sant Francesc 14, 17300 Blanes, Spain

organic molecules with neutral charge and low toxicity at high concentrations, usually combined with active “pumping out” of inorganic salts [15]. In general, salt-out microorganisms are better adapted to wider ranges of salinities but they have limitations to colonize extreme salinities due to the high energetic cost to produce compatible solutes [3]. Indeed, the bioenergetic requirement for maintaining these necessary osmoadaptation mechanisms is one of the main limitations microbes have for colonizing saline environments, and therefore some energetically limited metabolic pathways cannot be present beyond a certain salinity threshold [3]. This strategy is followed by most halophilic bacteria and most of the currently known halophilic microeukaryotes [14, 16, 17].

Osmoadaptation mechanisms are better known in halophilic prokaryotes than in microeukaryotes. This lack of knowledge remains in the higher methodological complexity and the more recent awareness on the ecological relevance of microeukaryotes in hypersaline environments [18]. Most of the knowledge regarding osmoadaptation in microeukaryotes had been gained from the study of the green algae *Dunaliella* and halophilic fungi, e.g., *Hortaea werneckii* [19], *Wallemia ichthyophaga* [20], and *Eurotium rubrum* [21]. However, more recent works have added additional information from phagotrophic protists based on both transcriptomic analysis (*Halocafeteria seosinensis* [22]) and experimental studies (*Schmidingerothrix salinarum* [23]). Although the accumulation of compatible solutes has been described in all the cases as the main osmoadaptation strategy for microeukaryotes, the wide genetic repertoire found related to specific adaptations suggests a larger variety of adaptive strategies [21]. Thus, some of the main differences observed at the genome level include acidic residues in predicted proteins and genes linked to membrane cation transporters and synthesis/transport of compatible solutes. For instance, an elevated presence of genes linked to cation transporters have been associated to low intracellular concentrations of K^+ and Na^+ in *H. werneckii*. Conversely, a limited number of those genes in *W. ichthyophaga* suggests a higher resistance to high intracellular Na^+ concentrations [23]. The expected higher adaptability of “salt-out” organisms may potentially provide competitive advantage in highly dynamic environments where frequent water level fluctuations and high evapotranspiration rates led to abrupt salinity changes, with potential dominance of organisms using the salt-out strategy even at the most extreme salinities [14, 24].

Ephemeral inland shallow lakes in arid and saline areas are very convenient natural model systems to test for selective microbial thresholds in highly fluctuating natural environments with high selective pressures. These ephemeral water masses are very dynamic and extremely responsive to changes in climate, with fluctuations in water volume, temperature, and salinity [25] larger than those reported in more stable hypersaline environments (e.g., solar salterns and large

hypersaline lakes). The endorheic saline basin of Monegros is rich in ephemeral shallow lakes with severe spatio-temporal changes in salinity, easily ranging from freshwater to salt saturation and vice versa [26]. We evaluated the composition and ecological changes of the planktonic microbiota (bacteria, protists, and fungi) along the salt gradient of these highly dynamic environments with frequent and bidirectional salt fluctuations, and the results were contextualized according to the previous knowledge on the ecology of (hyper)saline environments.

Materials and Methods

Study Site and Sampling

The Monegros Desert area (Central Ebro Basin, NE Spain, 41°42'N, 0°20'W) is one of the most arid regions in Europe with a large set of inland ephemeral saline lakes, and constitutes a unique landscape of great geological, edaphic, mineralogical, and hydrological interest [27, 28]. The area is an endorheic platform hosting hypersaline brines of $Cl-SO_4-Na-(Mg)$ type, where a rich microbial community develops [26, 29, 30]. In this endorheic area, the water balance between precipitation (mean annual rainfall = 337 mm/year) and evaporation and evapotranspiration rates (mean = 1408 mm/year) is negative for most of the months and for the whole year, producing one of the highest mean annual water deficits in Europe [31]. Field sampling was carried out in 14 ponds in different time periods and covered large salinity fluctuations (Supplementary Table S1). Conductivity, temperature, pH, dissolved oxygen, and redox potential were measured in situ using an HQ40d multiparameter meter (Hach, Loveland, CO, USA). For chlorophyll *a* (Chl *a*) analysis, 1-L water was filtered through 47-mm-diameter Whatman GF/F filters (0.7- μ m nominal particle retention), stored in the dark, and kept frozen. Chl *a* concentration was determined in acetone extracts by spectrophotometry (UV-2401PC, ultraviolet-visible Spectrometer; Shimadzu) as previously reported [32]. Salinity was measured in situ by a hand salinity refractometer (Atago S-28E, Japan). Five salinity categories were established for comparative purposes as follows: freshwater to slightly-brackish (0.1–1.4%, w/v), brackish (1.5–2.9%), marine (3–4.9%), hypersaline (5–14.9%), and extreme hypersaline (15–40%).

DNA Extraction and Sequencing

For DNA analyses, water samples were prefiltered in situ through a 50- μ m-pore-size net to retain large zooplankton and algae, and then between 100 and 500 mL (depending on the initial cells concentration) were successively filtered on 5- μ m and 0.2- μ m pore-size polycarbonate membranes (47-

mm diameter). The analyzed eukaryal size range (5–50 μm) mainly covered the nanoplanktonic eukaryotes (2–20 μm) and a fraction of microplanktonic eukaryotes (20–200 μm). The 5 μm (for 18S rRNA gene analyses) and 0.2 μm (for 16S rRNA gene analyses) membranes were separately stored in lysis buffer (40 mM EDTA, 50 mM Tris, pH 8.3, 0.75 M sucrose), enzymatically digested and phenol extracted (phenol–chloroform–isoamyl alcohol; 25:24:1, v/v/v) as previously reported [33]. The purification and concentration steps were carried out with Amicon® Ultra 4 Centrifugal Filter Units–100 000 NMWL (Millipore). DNA extracts were quantified using Quant-it assays in a Qubit™ fluorometer (Invitrogen). PCR amplification of the V4 region of the 16S rRNA gene (primers set 515F–806R [34]) and further sequencing with the Illumina 2×250 MiSeq platform were carried out according to the genomic core facilities protocols and methods of the RTSF-MSU (Michigan State University, USA) (<https://rtsf.natsci.msu.edu/>). The V9 region of the 18S rRNA gene was PCR amplified with the primers set 1391f–EukrBr [35] and sequenced with the Illumina 2×150 MiSeq platform (RTSF-MSU).

16S rRNA and 18S rRNA Genes Sequence Analysis

Raw rRNA gene sequences were processed with UPARSE [36]. The total number of sequences prior to quality filtering was 14,445,263 (mean of 79,060 per sample) for the 16S (70.6% had $\geq$ Q30) and 14,374,602 (mean of 83,489 per sample) for the 18S (87.1% had $\geq$ Q30) rRNA genes dataset. After merging of read pairs, each dataset was filtered by setting a minimum sequence length (250 bp for the 16S gene and 150 bp for the 18S gene) and a maximum number of expected errors of 0.25 [37]. After chimera filtering using default parameters, the UNOISE algorithm [38] clustered filtered sequences into Operational Taxonomic Units at 100% identity, i.e., zero-radius OTUs (zOTUs). The taxonomic assignment was carried out with SINA aligner v.1.2.11 [39], using the SILVA 132 reference database [40]. Specific BLAST searches were carried out to obtain a more detailed taxonomic classification of relevant zOTUs when needed. zOTUs were filtered by their alignment quality score ($>85\%$), and 16S rRNA gene sequences classified as archaea, mitochondria, and chloroplasts were excluded. Archaea showed high richness and very low abundances below 20‰ salinities, and were treated separately in a different study covering temporal colonization–extinction turnover rates along the gradient [30]. Only samples with > 6800 reads for bacteria and > 8100 reads for eukaryote were retained. The resulting zOTU tables consisted of 5653 Bacteria (131 samples) and 2658 Eukarya (129 samples). Data was normalized by rarefying the reads of all samples to the minimum number of reads in each dataset. For community related analysis, rarefaction at the same depth was repeated 100 times in order to avoid the

loss of less abundant zOTUs, and the resulting rarefactions were unified on an average rarefied zOTU table. Phylogenetic diversity was calculated in QIIME [41] using *alpha_diversity.py* after building a maximum-likelihood phylogenetic tree using *make_phylogeny.py* [42]. General metabolic and trophic traits for bacteria and eukarya were obtained from the closest-match in SILVA database and after literature surveys [43–46]. Energetic metabolism was focused on (1) photoautotrophy, (2) sulfur chemolithotrophs, (3) sulfur chemoorganotrophs, (4) methanotrophs, and (5) ammonia-oxidizing chemolithotrophs. For the trophic traits of protists and fungi, we established five groups as follows: (1) osmotrophs (which includes photosynthetic organisms), (2) phagotrophs, (3) mixotrophs, (4) saprotrophic, and (5) parasitic.

The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605.

Statistical Analysis

Statistical analyses were run in R environment (<http://www.r-project.org/>). Species richness, and the Shannon–Weaver diversity index were calculated using “vegan” package version 2.4-5 [47]. Spearman’s Rank order correlation coefficient was carried out to test monotonic associations between numeric variables (environmental data, relative abundances, diversity indexes). Pearson’s correlation coefficient was used to test lineal correlation between numeric variables. Differences among group means were tested using Tukey’s range test, function *HSD.test* within “agricolae” package version 1.2-8, after ANOVA test. The Duffrene-Legendre Indicator Species Analysis, that identifies indicator taxa based on fidelity and relative abundance, was carried out at the genus level with the *indval* function on the “labdsv” package version 1.8, in order to show the genera significantly associated (p value < 0.05) to each salinity category. The ten genera with the highest abundance in each salinity category were further analyzed. The *titan2* function from “TITAN2” package version 2.1 was used to identify significant variations on the taxonomic composition along the environmental gradient [48]. This is done after fractioning the environmental gradient in two at each midpoint of the gradient, and calculating taxa significantly associated to each side of that midpoint, which is expressed as a Z score. Then, Z scores are based on the Indicator Species Analysis at each midpoint of the gradient. High values of negative Z scores (Z^-) indicated high association of taxa to the left side of the midpoint, the one with lower salinity, so represent points of the gradient where the environmental variable starts affecting negatively some taxa. Conversely, positive Z scores (Z^+) indicated the opposite pattern, where taxa are significantly associated to the most saline side of the midpoint.

Results

Alpha Diversity Descriptors Along the Salinity Gradient

Bacteria and microbial eukarya showed differences in alpha diversity community descriptors (Richness, Shannon, and Faith's Phylogenetic Diversity) along the gradient (Fig. 1, and see additional information in Supplementary Table S2). Bacterial richness (R) was higher (average 372 ± 116 zOTUs, range 127–793 zOTUs) than microbial eukaryotes R (average 124 ± 69 zOTUs, range 14–351 zOTUs). The Shannon–Weaver diversity index (H) also reached higher values for bacteria (from 1.9 to 5.1, mean 3.47) than for eukarya (from 0.09 to 4.15, mean 2.21). The plots of richness and alpha diversity indicators vs. salinity showed a significant and negative trend for eukaryotes (Fig. 1d and e, respectively), but was not significant for bacteria (Fig. 1a and b). The negative trend was observed for both photosynthetic and heterotrophic protists, and also for fungi (Supplementary Table S2). At any rate, the phylogenetic diversity of both bacteria and eukarya showed a significant reduction as salinity increased that became stronger for microbial eukaryotes (Fig. 1c and f).

Taxa Replacement Along the Salinity Gradient

We identified significant variations on the taxonomic composition along the salinity gradient based on the Z scores (Fig. 2),

with significant ecological thresholds detected. Similar values for Z+ and Z– scores were observed in bacteria whereas, interestingly, eukaryal Z+ scores were significantly lower than Z–. The Z– for both bacterial and eukarya followed a similar pattern, with its maximum and most of the contribution to the cumulative frequency observed between 2% and 3% salinity. The higher Z+ bacterial scores at salinities 5 to 20%, indicated a partial replacement of the bacterial community by better adapted taxa that was not observed for eukarya.

Indicator Bacterial Taxa (IndVal) Along the Salt Gradient

As a whole, bacteria averaged 98% (relative abundances range 54–100%) of the total prokaryotes present in the PCR mixture (Supplementary Table S1). The most abundant phyla were *Gamma*proteobacteria (25–44%, mean relative abundances), *Alphaproteobacteria* (10–35%), *Bacteroidetes* (18–29%), and *Actinobacteria* (c. 5%) (Supplementary Fig. S1, and see a zoom for salinities >35% in Supplementary Fig. S2). *Halanaerobiaeota* (a recently proposed phylum in the SILVA database which includes the single order *Halanaerobiales*) were only noticeable at the most extreme end. Conversely, *Firmicutes* reached higher abundances at low salinities. The phylum *Spirochaetes* showed a positive correlation between relative abundance and salinity (Spearman's rho 0.60, p value < 0.001), with a highly diverse

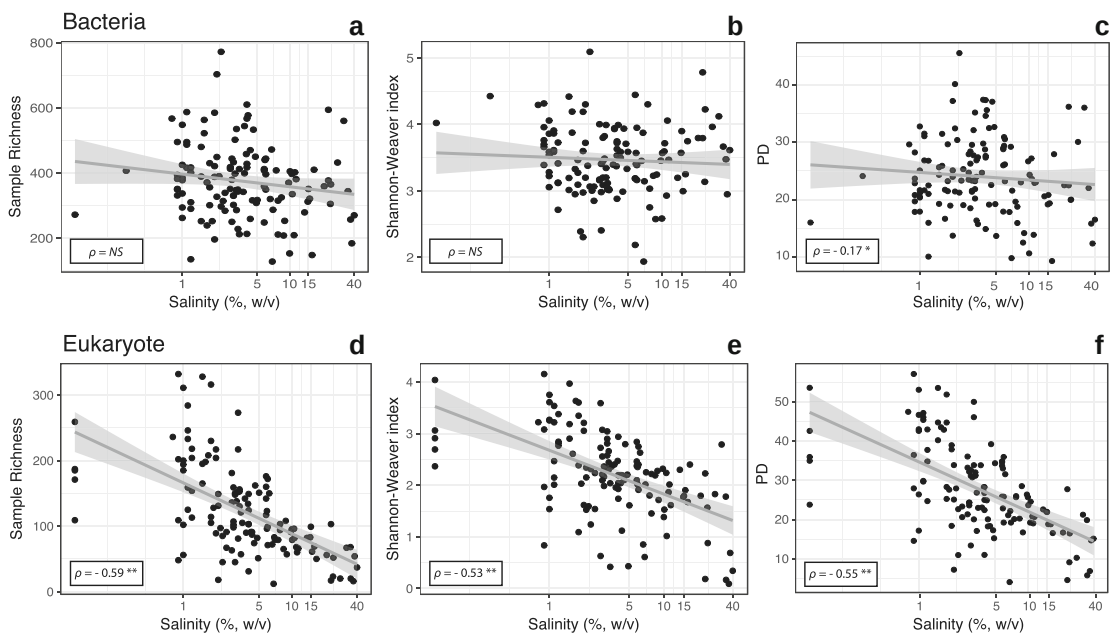
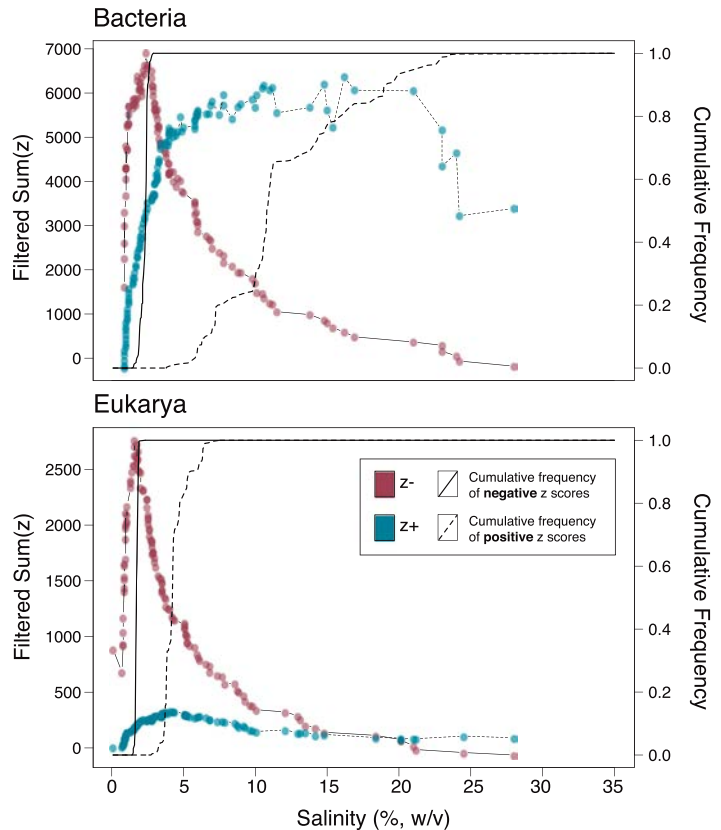


Fig. 1 Variations on ZOTU richness, the Shannon–Weaver diversity index, and Faith's phylogenetic diversity along the salinity gradient, both for bacteria (a, b, c) and eukaryote (d, e, f). Pearson's correlation is given

when significant p value: * p value < 0.05; ** p value < 0.01; NS, nonsignificant correlation

Fig. 2 Representation of the sums of Z scores (positive and negative) calculated for each midpoint of the salinity gradient using the *titan2* function from the “TITAN2” R package, both for bacterial and eukaryal communities. Sums of Z scores are based on the *Indval* analysis performed after segregating all samples along the salinity gradient as less/more saline than the tested point. Continuous and dashed lines represent the accumulated frequency of negative and positive Z scores, respectively, and its pronounced increase in certain points represents potential thresholds on the community composition variation. As the *Indval* analysis requires of a minimum number of samples at each side of the gradient midpoint, this could not be performed at the salt gradient tails



composition, and reached the 3% of bacterial reads at 20–25% salinity in the ephemeral lake Salineta (Supplementary Fig. S3).

The *Indval* analysis unveiled the genera with the highest fidelity for each salinity category (Fig. 3). The genera *Polaromonas*, *Polynucleobacter*, *Limnohabitans*, *Hydrogenophaga*, *Bacillus*, *Pedobacter*, and *Flavobacterium* reached significant abundances at low salinity, whereas *Salinivibrio*, unclassified *Nitrococcaceae* (94.4% of 16S rRNA gene sequence identity with *Nitrococcaceae* members included in SILVA database), *Halothiobacillus*, and genera belonging to the order *Halanaerobiales* became more important in terms of abundance at high salinities. Interestingly, the high abundance of the unclassified *Nitrococcaceae* at salinities >15% was mostly related to a single population (zOTU 7, NCBI acc. number MT129572, Supplementary Table S3) 100% identical to the recently described *Spiribacter salinus* (*Ectothiorhodospiraceae*) [49]. Conversely, *Salinibacter* sp. was always present at relative abundances <0.3% even at salinities >30%, and maximal abundances reached 3–4% only in two samples for the whole dataset (Salineta, samples MON 88 and 84). Finally,

Psychroflexus and unclassified *Rhodobacteraceae*, despite being significantly associated to a given salinity category, were substantially abundant along the whole salinity gradient.

Indicator Eukaryal Taxa (IndVal) Along the Salt Gradient

Despite the significant decrease of eukaryotic diversity observed along the salinity gradient, the whole eukaryal microbiome showed an outstanding diversity in the ephemeral saline lakes of Monegros, with up to 2658 zOTUs belonging to Amoebozoa, Archaeplastida, Centrohelida, Cryptophyceae, Excavata, Haptophyta, Opisthokonta and SAR (Supplementary Fig. S4). Unclassified Ciliophora and *Cryptomonas* showed the highest affinity for low salinities, whereas *Dunaliella* (Chlorophyta) prevailed at extreme salinities (Fig. 3). Most of the reads classified by SINA aligner as unclassified Chlorophyta, belonged to zOTU 3 (100% identity to *Dunaliella salina*, NCBI acc. number MG022673), which reached 80% of the total eukaryote microbiota for most of the extreme-hypersaline samples. It was also remarkable the high abundance of diatoms of the genus *Navicula* (Ochrophyta,

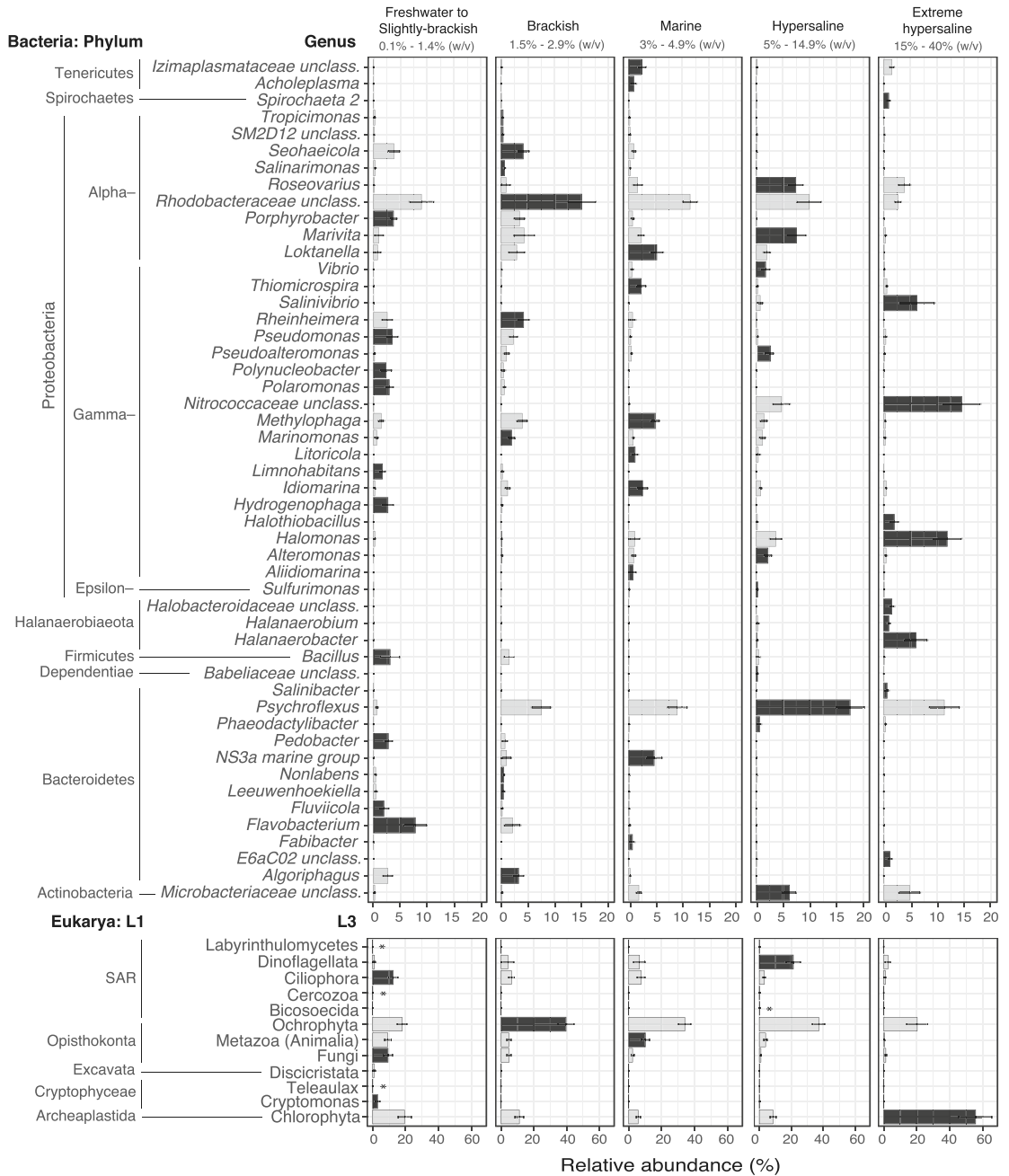


Fig. 3 Relative abundances (%) of the genera obtained from the *Indval* analysis along the different salinity ranges. Highlighted abundances represent the indicator species for each salinity range

Bacillariophyceae) along most of the salinity gradient, reaching > 25% of the relative abundance of eukaryotes in the brackish range. The dinoflagellate *Suessiacea*

(unclassified) represented by a single zOTU (zOTU 2), reached high abundances and fidelity at salinities <15% and showed 100% identity to the halotolerant dinoflagellate

Biecheleria sp. (NCBI acc. number KF463288 isolated from the saline lake Tirez, Spain).

Finally, fungi prevailed at the lower salinity end (Fig. 3 and Supplementary Fig. S5) and were less abundant as salt content increased. The most abundant taxa were Chytridiomycota, and two unknown fungi that deserve further research (zOTU 57, up to 23% of the eukaryotic reads in sample MON 2, salinity 1.2%, NCBI acc. number MW040517; and zOTU 102, up to 15.5% of the eukaryotic reads in sample MON 6, salinity 1.1%, NCBI acc. number MW040518). These unknown fungi showed <91% identity to any previously known sequences and were putatively allocated within the Basidiomycota in SINA.

Metabolic and Trophic Thresholds Along the Salt Gradient in Ephemeral Lakes

Changes in the general energy metabolism processes and trophic strategies were closely linked to shifts in microbial composition. Bacterial photoautotrophy (cyanobacteria and anoxygenic photoautotrophs) was a minor component and with no significant differences along the salt gradient categories (Fig. 4). Conversely, sulfur bacteria both chemolithotrophs (sulfur-oxidizing bacteria, SOB) and chemoorganotrophs (sulfate-reducing bacteria, SRB) increased along the gradient. Methanotrophs showed low abundances and were undetectable at salinities >15%. Ammonia

oxidizers were also poorly represented in the Monegros microbiome, with AOB undetected and AOA (*Thaumarchaeota*) mostly detected at salinities up to 10% [30].

Osmotrophy by photosynthetic microeukaryotes was the dominant eukaryal trophic strategy along the whole salt gradient, reaching almost 100% dominance at salinities >15% (Fig. 5) in agreement with the highest relative abundances of *Dunaliella* and *Navicula*. Conversely, phagotrophic eukaryotes (Ciliophora and Cercozoa) averaged c. 25% of the abundance at salinities <15% and became minor component at salinities > 20%, whereas the abundance of mixotrophic (mostly *Cryptomonas* and Dinoflagellata) reached its maximum at salinities <1.5% and unseen at salinities >3%. Finally, saprotrophic and parasitic microeukaryotes, related to the presence of Chytridiomycota, were observed mostly at <5% salinity.

Discussion

The dynamics and succession of planktonic microbial populations in ephemeral saline lakes have been poorly studied [50]. In hypersaline inland environments of arid regions, it is challenging to collect enough relevant information over temporal and spatial scales to properly study natural microbial processes, and successional studies are scarce [51]. The

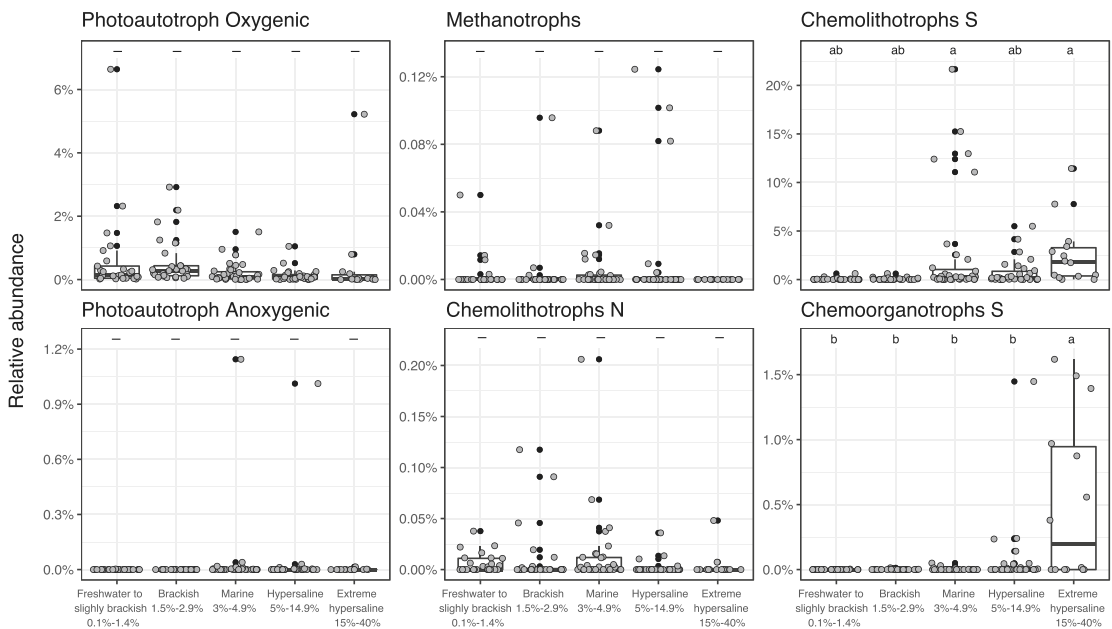


Fig. 4 Relative abundance of the bacterial taxa involved on diverse metabolic processes. Chemoorganotrophs S: as sulfate-reducing bacteria; Chemolithotrophs S: as sulfur-oxidizing bacteria. Gray dots represent group abundance on each sample. Black dots represent boxplots outliers

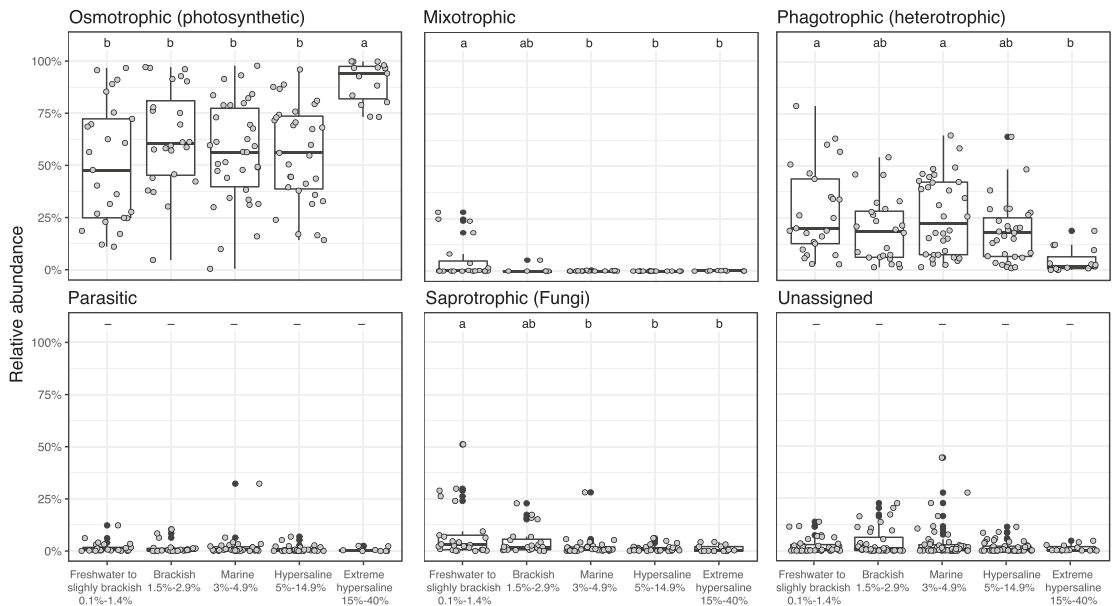


Fig. 5 Relative abundance of the eukaryal taxa based on its trophic strategy. Gray dots represent group abundance on each sample. Black dots represent boxplot outliers

Monegros desert contains one of the largest sets of inland saline lakes in Europe offering a wide repertoire of accessible heterogeneous and small ephemeral lakes within a short distance, easily covering different interannual dry and wet periods and ecological conditions [26, 30]. In the present study, we observed consistent changes in ecological and metabolic thresholds along the saline gradient that selectively affected different microbial populations according to their expected physiological and metabolic traits. The best known ecological thresholds with clearly segregated species distributions are usually found at the interfaces between land and water, oxic and anoxic environments, and marine waters and freshwaters. However, ecological thresholds may also exist when environmental changes are progressive [52]. As a whole, ecological thresholds are very helpful to understand ecosystems structures, to guide conservation policies, and to establish quantitative predictions for the ecological impact of environmental changes [32, 52]. In the intermittent saline lakes of Monegros, the strong environmental selection pressures (low rain precipitation, high solar irradiation, strong and persistent winds, high evaporation rates, and salts concentration and accumulation) provide an easy to access natural model system with progressive changes in salinity where ecological and metabolic thresholds appeared finely allocated.

Microorganisms inhabiting saline and hypersaline environments have been described in detail in the literature, first using culture dependent approaches [53, 54], and more extensively after the use of molecular biology techniques. Generally,

halophilic bacteria and archaea dominate in saline environments, with lower contribution of microeukaryotes to the total diversity. Most of the halophile and halotolerant bacteria are considered to be moderate rather than extreme halophiles [55]. The Bacteria domain contains many halophiles

from diverse phylogenetic groups such as *Proteobacteria*, *Cyanobacteria*, *Spirochaetes*, *Actinomycetes* and *Firmicutes*. The genera *Bacillus* [56], *Salinibacter* [57] and members of the order *Halanaerobiales* [57] are some of the most commonly found taxa expected in these environments. Archaea become more abundant as salinity increases [58], mostly with *Nanohaloarchaeota* and *Halobacteria* populations [30, 58] being *Haloquadratum* and *Halorubrum* predominant at the highest salinities [59–61]. Microeukaryal diversity tends to be lower at high salinities [24] and are mainly represented by Chlorophyta, Dinophyceae and Bacillariophyta, with the photosynthetic algae *Dunaliella* as one of the most common taxa at extreme salinities [62–64]. Together with other reported halophilic algae [65], both halophilic and halotolerant fungi (e.g., *Hortaea werneckii* and *Wallemia ichthyophaga*) and halophilic heterotrophic protists [14, 62] are also usually reported in saline environments.

Therefore, in principle, a similar assembly should be present in the saline lakes of Monegros. However, the taxonomic composition found, specially at the most saline end, significantly differed from the expectations following previous reports. Although specialist groups as *Salinibacter* and *Halanaerobiales* were detected at extreme salinities, they only

represented a very low fraction of the total bacterial assembly. Conversely, a very rich assembly of *Gammaproteobacteria*, *Bacteroidetes* and *Alphaproteobacteria* using “salt-out” strategies dominated the bacterial community at extreme salinities. Most of the populations found matched *Halomonas* [66, 67], *Salinivibrio* [61], *Spiribacter salinus* [49], and members of the order *Halanaerobiales* [55], among others, with high relative abundances at salinities close to NaCl saturation (Supplementary Fig. S2). This fact led to a higher bacterial diversity at extreme salinities than expected from studies carried out in more stable environments, where usually “salt-in” archaea, *Salinibacter*, and *Halanaerobiales* dominate [68, 69]. As previously reported, this strategy is more effective at the most extreme salinities and reduces adaptation to lower salinities [13, 70] and, therefore, should be less competitive for the most dynamic natural environments [14, 71]. Conversely, “salt-out” mechanisms provide higher plasticity at the expense of poorer adaptation at the most saline end [72]. Microeukaryotic assembly was mostly dominated by *Dunaliella* and *Navicula*, both accumulating compatible solutes as the main osmoadaptation strategy [63, 73], with progressively lower contribution of halophilic fungi and heterotrophic protists as salinity increased. Overall, the extremely high dynamism observed in the ephemeral lakes of Monegros selectively enriched more versatile (“salt-out”) organisms rather than those better adapted to stable high salinities, usually more common in solar salterns.

Changes in metabolic and physiological performance were also consistent along the salinity gradient. Autotrophic ammonia oxidation, autotrophic nitrite oxidation, and acetate-based methanogenesis, among others, are usually restricted to low salinity environments [3, 43, 74]. The low abundance of methanotrophs observed in the ephemeral lakes of Monegros may be explained by constraints for aerobic methane oxidation in high-salt environments [43] but also for limitations in methanogenesis. Methanogenic activity based on H_2+CO_2 (Group 1) and acetate (Group 3) have never been detected above c. 12.5% salinity [43, 75]. Methanogenesis based on methylated amines (Group 2) may potentially occur up to 25% of salts (w/v), under anaerobic conditions [43, 58]. However, Monegros is a gypsum-rich area with a rich presence of sulfur bacteria and a potentially active sulfur cycle in most of the sediments [29] that do not favor methanogenesis.

We also observed consistent changes of the trophic strategy for microbial eukaryotes in the ephemeral lakes. The dominance of osmotrophy (which includes photosynthetic organisms) was linked to the high abundance of *Dunaliella* and *Navicula* covering all the gradient. The low abundance of Cyanobacteria at high salinities agrees with previous reports [76] being *Dunaliella* the main primary producer at the most extreme salinities. The progressive decrease of phagotrophic organisms agrees with previous studies in coastal solar salterns that showed lack of bacterivory above 25% salinity

[18]. Groups potentially including halophile and halotolerant melanized fungi (“black yeasts”) usually found in hypersaline environments [77], were detected at low abundances along the gradient in Monegros. Saprotrophic fungi were mostly found at salinities <5%. However, more detailed studies and refined taxonomic assignments are needed for the aquatic fungi found in Monegros. Finally, parasitic microeukaryotes were present all over the gradient suggesting potential key missing pathways in ephemeral lakes dynamics. Taken all together, these unseen groups offered new perspectives for improving the understanding of planktonic food webs [45] and should be more carefully studied in the future.

Overall, we have reported a very dynamic and rich microbial assemblages in the Monegros wetlands system, even under the most extreme conditions. The particular characteristics and high environmental variability of this area, allow us to observe natural transitions in detail both at the taxonomic and at the functional level. Most of the high saline waters traditionally studied are quite stable environments with rather constant salinities [51, 62, 78] or where salt production practices guides salinity increases unidirectionally [61]. Ephemeral saline lakes with high turnover and frequent salinity transitions follow closely environmental perturbations and flooding and desiccation dynamics, offering a well-suited natural framework for the study of evolutionary and ecological microbial responses [32] under highly dynamic fluctuations and strong selective pressures.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00248-021-01732-9>.

Author Contribution EOC designed the study, EOC and XTM carried out the field work, and MMS carried out the analyses. MMS and XTM wrote a first draft of the manuscript. EOC led the final version of the manuscript.

Funding Funding was provided by grant BRIDGES CGL2015-69043-P and INTERACTOMA RTI2018-101205-B-I00 from the Spanish Agency of Research (AEI-MICINN) and European funding (ERDF) to EOC. MMS was funded by the Spanish FPI PhD scholarships program (BRIDGES-MINECO).

Data Availability The whole gene sequence datasets were deposited to the NCBI Sequence Read Archive and are available through BioProject record ID PRJNA429605.

Code Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Conflict of Interest The authors declare no conflict of interest.

References

- Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. *Proc Natl Acad Sci* 104:11436–11440. <https://doi.org/10.1073/pnas.0611525104>
- Auguet JC, Barberán A, Casamayor E (2010) Global ecological patterns in uncultured Archaea. *ISME J* 4:182–190. <https://doi.org/10.1038/ismej.2009.109>
- Oren A (1999) Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* 63:334–348. <https://doi.org/10.1128/MMBR.63.2.334-348.1999>
- Gasol JM, Casamayor EO, Joint I, Garde K, Gustavson K, Benlloch S, Díez B, Schauer M, Massana R, Pedrós-Alió C (2004) Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat Microb Ecol* 34:193–206. <https://doi.org/10.3354/ame034193>
- Booth IR, Louis P (1999) Managing hypoosmotic stress: aquaporins and medianosensitive channels in *Escherichia coli*. *Curr Opin Microbiol* 2:166–169. [https://doi.org/10.1016/S1369-5274\(99\)80029-0](https://doi.org/10.1016/S1369-5274(99)80029-0)
- Wood CM (2011) Rapid regulation of Na⁺ and Cl[−] flux rates in killifish after acute salinity challenge. *J Exp Mar Biol Ecol* 409:62–69. <https://doi.org/10.1016/j.jembe.2011.08.009>
- Rodríguez-Valera F (1993) Introduction to saline environments. In: Vreeland RH, Hochstein LJ (eds) *The biology of halophilic bacteria*. CRC, Boca Raton FL, pp 1–25
- Por FD (1972) Hydrobiological notes on the high-salinity waters of the Sinai Peninsula. *Mar Biol* 14:111–119. <https://doi.org/10.1007/BF00373210>
- Imhoff JF (2001) True marine and halophilic anoxygenic phototrophic bacteria. *Arch Microbiol* 176:243–254. <https://doi.org/10.1007/s002030100326>
- Edgerton ME, Brimblecombe P (1981) Thermodynamics of halobacterial environments. *Can J Microbiol* 27:899–909. <https://doi.org/10.1139/m81-142>
- Kushner DJ (1978) Life in high salt and solute concentrations: halophilic bacteria. In: Kushner DJ (ed) *Microbial life in extreme environments*. Academic Press, London, pp 317–368
- Kushner DJ (1985) *The Halobacteriaceae*. In: Woese CR, Wolfe RS (eds) *The bacteria. A treatise on structure and function*. Academic Press, Orlando, FL, pp 171–214
- Galinski EA, Trüper HG (1994) Microbial behaviour in salt-stressed ecosystems. *FEMS Microbiol Rev* 15:95–108. <https://doi.org/10.1111/j.1574-6976.1994.tb00128.x>
- Gunde-Cimerman N, Plemenitaš A, Oren A (2018) Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. *FEMS Microbiol Rev* 42:353–375. <https://doi.org/10.1093/femsec/fuy009>
- Lang F (2007) Mechanisms and significance of cell volume regulation. *J Am Coll Nutr* 26:613S–623S. <https://doi.org/10.1080/07315724.2007.10719667>
- Trüper HG, Galinski EA (1990) Biosynthesis and fate of compatible solutes in extremely halophilic phototrophic eubacteria. *FEMS Microbiol Rev* 6:247–254. <https://doi.org/10.1111/j.1574-6968.1990.tb04098.x>
- Hagemann M (2016) Coping with high and variable salinity: molecular aspects of compatible solute accumulation BT—the physiology of microalgae. In: Borowitzka MA, Beardall J, Raven JA (eds). Springer International Publishing, Cham, pp 359–372
- Pedrós-Alió C, Calderón-Paz JJ, MacLean MH, Medina G, Marrasé C, Gasol JM, Guixa-Boixereu N (2000) The microbial food web along salinity gradients. *FEMS Microbiol Ecol* 32:143–155. <https://doi.org/10.1111/j.1574-6941.2000.tb00708.x>
- Plemenitas A, Vaupotic T, Lenassi M et al (2008) Adaptation of extremely halotolerant black yeast *Hortaea werneckii* to increased osmolarity: a molecular perspective at a glance. *Stud Mycol* 61:67–75. <https://doi.org/10.3114/sim.2008.61.06>
- Zajc J, Kogej T, Galinski EA et al (2014) Osmoadaptation strategy of the most halophilic fungus, *Wallemia ichthyophaga*, growing optimally at salinities above 15% NaCl. *Appl Environ Microbiol* 80:247 LP–247256. <https://doi.org/10.1128/AEM.02702-13>
- Kis-Papo T, Weig AR, Riley R, Peršoh D, Salamov A, Sun H, Lipzen A, Wasser SP, Rambold G, Grigoriev IV, Nevo E (2014) Genomic adaptations of the halophilic Dead Sea filamentous fungus *Eurotium rubrum*. *Nat Commun* 5:3745. <https://doi.org/10.1038/ncomms4745>
- Harding T, Roger AJ, Simpson AGB (2017) Adaptations to high salt in a halophilic protist: differential expression and gene acquisitions through duplications and gene transfers. *Front Microbiol* 8:944. <https://doi.org/10.3389/fmicb.2017.00944>
- Weinisch L, Kühner S, Roth R, Grimm M, Roth T, Netz DJA, Pierik AJ, Filker S (2018) Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate *Schmidingeriothrix salinarum*. *PLoS Biol* 16:e2003892. <https://doi.org/10.1371/journal.pbio.2003892>
- Filker S, Forster D, Weinisch L, Mora-Ruiz M, González B, Fariás ME, Rosselló-Móra R, Stoeck T (2017) Transition boundaries for protistan species turnover in hypersaline waters of different biogeographic regions. *Environ Microbiol* 19:3186–3200. <https://doi.org/10.1111/1462-2920.13805>
- Waiser MJ, Roberts RD (2009) Saline inland waters. In: Likens GE (ed) *Encyclopedia of inland waters*. Elsevier, Amsterdam
- Casamayor EO, Triadó-Margarit X, Castañeda C (2013) Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. *FEMS Microbiol Ecol* 85:503–518. <https://doi.org/10.1111/1574-6941.12139>
- Pueyo Mur JJ, Inglés Urpinell M (1987) Magnesite formation in recent playa lakes, Los Monegros, Spain. *Geol Soc Lond Spec Publ* 36:119–122. <https://doi.org/10.1144/GSL.SP.1987.036.01.10>
- Samper-Calvete FJ, García-Vera MA (1998) Inverse modeling of groundwater flow in the semiarid evaporitic closed basin of Los Monegros, Spain. *Hydrogeol J* 6:33–49. <https://doi.org/10.1007/s100400050132>
- Menéndez-Serra M, Triadó-Margarit X, Castañeda C, Herrero J, Casamayor EO (2019) Microbial composition, potential functional roles and genetic novelty in gypsum-rich and hypersaline soils of Monegros and Gallocanta (Spain). *Sci Total Environ* 650:343–353. <https://doi.org/10.1016/j.scitotenv.2018.09.050>
- Menéndez-Serra M, Ontiveros V, Triadó-Margarit X et al (2020) Dynamics and ecological distributions of the Archaea microbiome from inland saline lakes (Monegros Desert, Spain). *FEMS Microbiol Ecol* 96:fiaa019. <https://doi.org/10.1093/femsec/fiaa019>
- Herrero J, Snyder RL (1997) Aridity and irrigation in Aragon, Spain. *J Arid Environ* 35:535–547. <https://doi.org/10.1006/jare.1996.0222>
- Triadó-Margarit X, Capitán JA, Menéndez-Serra M, Ortiz-Álvarez R, Ontiveros VJ, Casamayor EO, Alonso D (2019) A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly. *ISME J* 13:2681–2689. <https://doi.org/10.1038/s41396-019-0454-4>
- Ferrera I, Massana R, Casamayor EO et al (2004) High-diversity biofilm for the oxidation of sulfide-containing effluents. *Appl Microbiol Biotechnol* 64:726–734. <https://doi.org/10.1007/s00253-004-1582-x>
- Caporaso JG, Lauber CL, Walters WA et al (2010) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. <https://doi.org/10.1073/pnas.1000080107>
- Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit

- ribosomal RNA genes. PLoS One 4:e6372. <https://doi.org/10.1371/journal.pone.0006372>
36. Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–998. <https://doi.org/10.1038/nmeth.2604>
 37. Edgar RC, Flyvbjerg H (2015) Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31:3476–3482
 38. Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv. <https://doi.org/10.1101/081257>
 39. Pruesse E, Peplies J, Glöckner FO (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829. <https://doi.org/10.1093/bioinformatics/btv401>
 40. Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUASt: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
 41. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JJ, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336. <https://doi.org/10.1038/nmeth.f.303>
 42. Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. PLoS One 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
 43. Oren A (2011) Thermodynamic limits to microbial life at high salt concentrations. Environ Microbiol 13:1908–1923. <https://doi.org/10.1111/j.1462-2920.2010.02365.x>
 44. Llorens-Marés T, Yooseph S, Goll J, Hoffman J, Vila-Costa M, Borrego CM, Dupont CL, Casamayor EO (2015) Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. ISME J 9:1648–1661. <https://doi.org/10.1038/ismej.2014.254>
 45. Ortiz-Álvarez R, Triadó-Margarit X, Camarero L, Casamayor EO, Catalan J (2018) High planktonic diversity in mountain lakes contains similar contributions of autotrophic, heterotrophic and parasitic eukaryotic life forms. Sci Rep 8:4457. <https://doi.org/10.1038/s41598-018-22835-3>
 46. Louca S, Parfrey LW, Doebeli M (2016) Decoupling function and taxonomy in the global ocean microbiome. Science 353:1272 LP–1271277. <https://doi.org/10.1126/science.aaf4507>
 47. Oksanen J, Kindt R, Legendre P, et al (2016) vegan: Community Ecology Package version 2.4-1 September 2016. <https://CRAN.R-project.org/package=vegan>
 48. Baker ME, King RS (2013) Of TITAN and straw men: an appeal for greater understanding of community data. Freshw Sci 32:489–506. <https://doi.org/10.1899/12-142.1>
 49. León MJ, Fernández AB, Ghai R, Sánchez-Porro C, Rodríguez-Valera F, Ventosa A (2014) From metagenomics to pure culture: Isolation and characterization of the moderately halophilic bacterium *Spiribacter salinus* gen. nov., sp. nov. Appl Environ Microbiol 80:3850–3857. <https://doi.org/10.1128/AEM.00430-14>
 50. Navarro JB, Moser DP, Flores A, Ross C, Rosen MR, Dong H, Zhang G, Hedlund BP (2009) Bacterial succession within an ephemeral hypereutrophic Mojave Desert Playa Lake. Microb Ecol 57:307–320. <https://doi.org/10.1007/s00248-008-9426-3>
 51. Podell S, Emerson JB, Jones CM, Ugalde JA, Welch S, Heidelberg KB, Banfield JF, Allen EE (2014) Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. ISME J 8:979–990. <https://doi.org/10.1038/ismej.2013.221>
 52. Catalan J, Barbieri MG, Bartumeus F et al (2009) Ecological thresholds in European alpine lakes. Freshw Biol 54:2494–2517. <https://doi.org/10.1111/j.1365-2427.2009.02286.x>
 53. Ventosa A, Rodríguez-Valera F, Poindexter JS, Reznikoff WS (1984) Selection for moderately halophilic bacteria by gradual salinity increases. Can J Microbiol 30:1279–1282. <https://doi.org/10.1139/m84-202>
 54. Rodríguez-Valera F, Ventosa A, Juez G, Imhoff JF (1985) Variation of environmental features and microbial populations with salt concentrations in a multi-pond saltern. Microb Ecol 11:107–115. <https://doi.org/10.1007/BF02010483>
 55. Oren A (2002) Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. J Ind Microbiol Biotechnol 28:56–63. <https://doi.org/10.1038/sj/jim/7000176>
 56. Kim J-S, Makama M, Petit J, Park NH, Cohan FM, Dungan RS (2012) Diversity of Bacteria and Archaea in hypersaline sediment from Death Valley National Park, California. MicrobiologyOpen 1:135–148. <https://doi.org/10.1002/mbo3.20>
 57. Antón J, Oren A, Benlloch S, Rodríguez-Valera F, Amann R, Rosselló-Mora R (2002) *Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the Bacteria from saltern crystallizer ponds. Int J Syst Evol Microbiol 52:485–491. <https://doi.org/10.1099/00207713-52-2-485>
 58. Andrei A-Ş, Banciu HL, Oren A (2012) Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. FEMS Microbiol Lett 330:1–9. <https://doi.org/10.1111/j.1574-6968.2012.02526.x>
 59. Oh D, Porter K, Russ B, Burns D, Dyall-Smith M (2010) Diversity of *Haloquadratum* and other haloarchaea in three, geographically distant, Australian saltern crystallizer ponds. Extremophiles 14:161–169. <https://doi.org/10.1007/s00792-009-0295-6>
 60. Dillon J, Carlin M, Gutierrez A et al (2013) Patterns of microbial diversity along a salinity gradient in the Guerrero Negro solar saltern, Baja CA Sur, Mexico. Front Microbiol 4:399. <https://doi.org/10.3389/fmicb.2013.00399>
 61. Ventosa A, Fernández AB, León MJ, Sánchez-Porro C, Rodríguez-Valera F (2014) The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. Extremophiles 18:811–824. <https://doi.org/10.1007/s00792-014-0681-6>
 62. Heidelberg K, Nelson W, Holm J et al (2013) Characterization of eukaryotic microbial diversity in hypersaline Lake Tyrrell, Australia. Front Microbiol 4:115. <https://doi.org/10.3389/fmicb.2013.00115>
 63. Oren A (2014) The ecology of *Dunaliella* in high-salt environments. J Biol Res 21:1–8. <https://doi.org/10.1186/s40709-014-0023-y>
 64. Triadó-Margarit X, Casamayor EO (2013) High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. FEMS Microbiol Ecol 85:27–36. <https://doi.org/10.1111/1574-6941.12095>
 65. Javor B (1989) *Dunaliella* and other halophilic, eukaryotic algae BT. In: Javor B (ed) Hypersaline environments: microbiology and biogeochemistry. Springer, Berlin, Heidelberg, pp 147–158
 66. Vreeland RH, Litchfield CD, Martin EL, Elliot E (1980) *Halomonas elongata*, a new genus and species of extremely salt-tolerant bacteria. Int J Syst Bacteriol 30:485–495. <https://doi.org/10.1099/00207713-30-2-485>
 67. El Hidri D, Guesmi A, Najjari A et al (2013) Cultivation-dependent assessment, diversity, and ecology of haloalkaliphilic bacteria in arid saline systems of southern Tunisia. Biomed Res Int 2013:648141–648115. <https://doi.org/10.1155/2013/648141>
 68. Fernández AB, Ghai R, Martín-Cuadrado AB, Sánchez-Porro C, Rodríguez-Valera F, Ventosa A (2014) Prokaryotic taxonomic and metabolic diversity of an intermediate salinity hypersaline habitat assessed by metagenomics. FEMS Microbiol Ecol 88:623–635. <https://doi.org/10.1111/1574-6941.12329>

69. Ghai R, Pašić L, Fernández AB, Martín-Cuadrado AB, Mizuno CM, McMahon KD, Papke RT, Stepanauskas R, Rodríguez-Brito B, Rohwer F, Sánchez-Porro C, Ventosa A, Rodríguez-Valera F (2011) New abundant microbial groups in aquatic hypersaline environments. *Sci Rep* 1:135. <https://doi.org/10.1038/srep00135>
70. Oren A (2008) Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline Systems* 4:2. <https://doi.org/10.1186/1746-1448-4-2>
71. Sleator RD, Hill C (2002) Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol Rev* 26:49–71. <https://doi.org/10.1111/j.1574-6976.2002.tb00598.x>
72. Kempf B, Bremer E (1998) Stress responses of *Bacillus subtilis* to high osmolarity environments: uptake and synthesis of osmoprotectants. *J Biosci* 23:447–455. <https://doi.org/10.1007/BF02936138>
73. Scholz B, Liebezeit G (2012) Compatible solutes in three marine intertidal microphytobenthic Wadden Sea diatoms exposed to different salinities. *Eur J Phycol* 47:393–407. <https://doi.org/10.1080/09670262.2012.720714>
74. Fernández-Guerra A, Casamayor EO (2012) Habitat-associated phylogenetic community patterns of microbial ammonia oxidizers. *PLoS One* 7:e47330. <https://doi.org/10.1371/journal.pone.0047330>
75. Pattnaik P, Mishra SR, Bharati K, Mohanty SR, Sethunathan N, Adhya TK (2000) Influence of salinity on methanogenesis and associated microflora in tropical rice soils. *Microbiol Res* 155: 215–220. [https://doi.org/10.1016/S0944-5013\(00\)80035-X](https://doi.org/10.1016/S0944-5013(00)80035-X)
76. Oren A (2015) Halophilic microbial communities and their environments. *Curr Opin Biotechnol* 33:119–124. <https://doi.org/10.1016/j.copbio.2015.02.005>
77. Gunde-Cimerman N, Zalar P, de Hoog S, Plemenitas A (2000) Hypersaline waters in salterns—natural ecological niches for halophilic black yeasts. *FEMS Microbiol Ecol* 32(3):235–240. <https://doi.org/10.1111/j.1574-6941.2000.tb00716.x>
78. Jacob JH, Hussein EI, Shakhathreh MAK, Cornelison CT (2017) Microbial community analysis of the hypersaline water of the Dead Sea using high-throughput amplicon sequencing. *MicrobiologyOpen* 6:e00500. <https://doi.org/10.1002/mbo3.500>



A Randomized Trait Community Clustering approach to unveil consistent environmental thresholds in community assembly

Xavier Triadó-Margarit ¹ · José A. Capitán ^{2,3} · Mateu Menéndez-Serra ^{1,3} · Rüdiger Ortiz-Álvarez ¹ · Vicente J. Ontiveros ³ · Emilio O. Casamayor ¹ · David Alonso ³

Received: 12 November 2018 / Revised: 21 May 2019 / Accepted: 1 June 2019

© The Author(s), under exclusive licence to International Society for Microbial Ecology 2019

Abstract

Similarities and differences of phenotypes within local co-occurring species hold the key to inferring the contribution of stochastic or deterministic processes in community assembly. Developing both phylogenetic-based and trait-based quantitative methods to unravel these processes is a major aim in community ecology. We developed a trait-based approach that: (i) assesses if a community trait clustering pattern is related to increasing environmental constraints along a gradient; and (ii) determines quantitative thresholds for an environmental variable along a gradient to interpret changes in prevailing community assembly drivers. We used a regional set of natural shallow saline ponds covering a wide salinity gradient (0.1–40‰ w/v). We identify a consistent discrete salinity threshold (ca. 5‰) for microbial community assembly drivers. Above 5‰ salinity a strong environmental filtering prevailed as an assembly force, whereas a combination of biotic and abiotic factors dominated at lower salinities. This method provides a conceptual approach to identify consistent environmental thresholds in community assembly and enables quantitative predictions for the ecological impact of environmental changes.

Introduction

The main goal of community ecology is to understand the underlying forces that determine species numbers and

identities, and their relative abundances, across spatio-temporal scales. Many recent efforts are aimed at developing robust methods to estimate and understand the relative importance of different community assembly processes [1]. Ecologists generally accept that community assembly must necessarily be driven by both dispersal-assembly and niche-assembly mechanisms [2–4]. Dispersal-assembly theories emphasize (1) stochastic colonization (driven largely by species pools) and (2) local extinctions (driven by random events affecting small populations sizes) [5, 6], while niche-assembly highlights (1) the importance of species differences and (2) their interactions in predicting community composition [7–9]. Niche-assembly mechanisms can affect either extinction or colonization. From this perspective, local extinctions result in general from negative interactions. They are often caused by competitive exclusion of sub-dominant species disappearing from the system due to a worse overall performance [10, 11], but strong environmental filtering (cf. species sorting [12]) can also prevent colonization and establishment of poorly adapted species.

In contrast, positive interactions (facilitation) decrease the probability of local extinctions, enhancing persistence and biodiversity [13–15]. Assessing the link between trait

These authors contributed equally: Xavier Triadó-Margarit, José A. Capitán

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-019-0454-4>) contains supplementary material, which is available to authorized users.

✉ Emilio O. Casamayor
casamayor@ceab.csic.es

✉ David Alonso
dalonso@ceab.csic.es

¹ Integrative Freshwater Ecology Group Centre for Advanced Studies of Blanes (CEAB-CSIC), Spanish Council for Scientific Research, Accés Cala St. Francesc 14, E-17300 Blanes, Spain

² Departamento de Matemática Aplicada, Universidad Politécnica de Madrid, Av. Juan de Herrera, 6, E-28040 Madrid, Spain

³ Theoretical and Computational Ecology, Centre for Advanced Studies of Blanes (CEAB-CSIC), Spanish Council for Scientific Research, Accés Cala St. Francesc 14, E-17300 Blanes, Spain

variation, the environment, and biotic interactions (either positive or negative) is not trivial [1, 16–18]. Further, there is an increasing recognition of the need for methods to make robust predictions that help understand responses of species assemblages to environmental shifts [19, 20].

In studying community assembly of macroorganisms, spatio-temporal scales are commonly huge, making long-term ecological research challenging. Microbes, in contrast, thrive in highly diverse communities and are characterized by short time scales. This makes microbial communities ideal study systems, providing robust insights into assembly processes [21, 22]. Species and individuals within assemblages respond to the environment according to their functional characteristics. In particular, microbes with stress tolerance traits are selected in specific environmental conditions [23]. Genomic traits allow the linkage of bacterial functions to their environmental preferences at different scales, from local habitats to biogeographical regions [24]. This approach could be especially helpful across habitats with strong environmental constraints, where a variety of functional adaptations can emerge along a gradient.

Moreover, there is an increasing recognition that variations in the distribution of functional traits along environmental gradients can be quantified [18, 25, 26]. Whenever niche-assembly mechanisms are overridden by strong random processes, the local community composition is expected to be akin to a random sample from the regional pool [27].

When this happens, local trait distributions should not differ significantly from the regional trait distribution. Conversely, differences between local and regional trait distributions are a fingerprint of the existence of specific drivers of community assembly. The classical framework by Webb [28] considers clustering in trait values as a signature of environmental filtering, while Mayfield and Levine [10] remarked that such clustering can result also from competitive exclusion.

Here, we use shifts in the distribution of functional traits along an environmental gradient to assess the ecological impact of a changing environment on community trait structure. Our methods are inspired by the data randomization/resampling null model approach [29]. Our trait-based approach consistently finds thresholds in an environmental variable along a gradient with the aim of interpreting meaningful changes in the type of drivers responsible for community assembly. We call this index-based method RTCC—Randomized Trait Community Clustering.

Materials and methods

Details related to sampling sites, molecular methods, bioinformatics processing, and additional statistics are

provided in SI Materials and Methods, while null models of community assembly are fully described in this section. See also supplementary text and Fig. S1 for a summarized work-flow of the methods carried out in this study. In addition, source code for the RTCC Method and sample data files are available in *GitHub*. Briefly, samples were collected from the Monegros Desert area (NE Spain, 41°42'N, 0°20'W) which harbors among the largest number of inland saline lakes in Europe [30]. A total of 148 samples were taken from 14 different shallow lagoons (Fig. S4). We explored the spatio-temporal variation of a regional set of these local aquatic communities driven by a wide salinity gradient (0.1–40% of dissolved salts) (Fig. S5A). Temporal data of wind speeds were obtained from the Meteorological state agency reporting data of an automatic station in the area (Bujaraloz) (Fig. S5B). For DNA analyses, water samples were pre-filtered in situ through a 50 µm-pore-size net, to retain large zooplankton and algae, and 100–500 mL were subsequently filtered on 5 µm and then 0.2 µm pore-size polycarbonate membranes (47 mm diameter).

The membranes were enzymatically digested and phenol–chloroform extracted [30]. To obtain community profiles, we performed a NGS sequencing step targeting the V4 region of the bacterial 16S rRNA gene by means of a high-speed multiplexed Illumina MiSeq platform. Raw sequences were processed using the UPARSE pipeline [31]. After the processing steps, we only kept those samples that presented library sizes larger than 10,000 reads/sample. The final pool consisted of 136 samples with 9993 operational taxonomic units (OTUs). In order to associate OTUs to genome content, we used a 16S rRNA database [32] linked to the IMG genomic database [33]. OTUs were matched to 16S rRNA gene records available in the PATRIC genomic database (as of January 2016) [32]. The average percentage of abundance matched per sample was ca. 62%. We tested the representativeness of the subset matching genomes in our approach (see Figs. S1 and S2), i.e., we explored if the ordination pattern of the whole dataset was consistent with the ordination of the subset. In addition, a Mantel test (based on Spearman correlation) was also used to conclude on this decision-making step. See also the RTCC Method protocol in SI document. Functional predictions based on representative genomes allow for an inferring of genomic and metabolic potential of 16S rRNA sequences [34, 35]. For this we downloaded available genomic traits of potential interest from IMG [33]: genome size, gene count, % GC, % coding base, % CDS, % RNA, rRNA count, % transporter proteins, % signal peptide, and % transmembrane proteins. We selected 10 traits that summarize genomic structural/functional variability, some of them known to respond to ecological adaptations.

Null models of community assembly

In order to quantify the importance of environmental filtering in the assembly of ecological communities, we developed a measure of the degree of clustering of a set of samples (or sites). The pool of species was formed by all the species observed in the complete set of samples taken from the lagoons. First we calculated a measure of average dissimilarity in the trait distribution of a sample as the pairwise mean trait difference between every species pair in the sample, the functional mean pairwise distance (MPD) [28]

$$\Delta x = \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{j=i+1}^n |x_i - x_j| \quad (1)$$

where x_i is the trait value of species i in the sample (with n reported species). Next we tested if the observed measure Δx significantly differed from the expectation of the same measure across realizations of consistent null models with the same number of species as appearing in the empirical sample. We assumed that any species from the pool can take part in a null-model synthetic community. We build the null-model dissimilarity distribution over replicates to obtain a p -value, i.e., the quantile defined by the empirical difference within the distribution of simulated differences across realizations. We tested whether the empirical sample dissimilarity value was significantly large or small, indicating over-dispersion or clustering, respectively. If the observed difference of our empirical sample was under the lower 5% of the simulated distribution trait differences, we considered that sample as showing significant clustering (Fig. 1a, left). We repeated this test for every sample in our data set, which yielded the p -value distributions for each target trait (Fig. 1a, right). Let N be the number of samples in the data set, and N_c the number of samples that showed significant clustering when tested against the null model. We defined a clustering index,

$$h = \frac{N_c}{N}, \quad (2)$$

as the proportion of samples showing significant clustering over all the samples considered. Note that the clustering index is a global measure for the whole set of samples. Error bars for the clustering index were calculated by averaging this quantity over 100 repetitions of the distribution of p -values according to the null model (Fig. 1b).

The null models developed differ in the definition of the species pools potentially available for each of the samples, that is, in the way species were drawn from the species pool:

- Random assembly:** For all samples, all species are potentially available. They were chosen at random from the pool with equal probability.

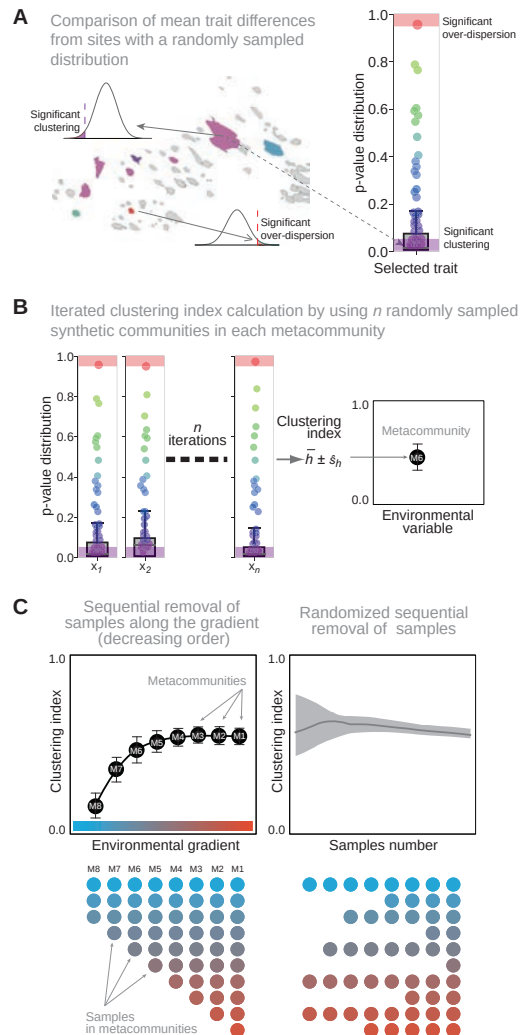


Fig. 1 Methodological framework for the Randomized Trait Community Clustering (RTCC) method. **a** For each sample from different water bodies over a given region we test the hypothesis that observed average dissimilarity (see Eq. (1)) is compatible with a given null model. These non-parametric randomizations assign a p -value for each sample. For all samples in the dataset we represent the p -value distributions obtained by the null model for a given functional trait. Significantly low p -values are related to clustering ($p < 0.05$), and significantly high p -values to over-dispersion. **b** For the whole sample set, which we define here as a metacommunity, we calculate the fraction of samples that yield significant clustering, i.e., its clustering index (see Eq. (2)). We average several realizations of the null model in order to get average clustering index and error bars. **c** Samples are sequentially removed along decreasing values of the environmental variable to define nested metacommunities; for each of them we calculate the clustering index (left). The significance of this curve is assessed by building an ensemble of curves corresponding to removing samples in random orderings, which are then used to define the shaded area (right)

- ii. *Abundance-based assembly*: For all samples, all species are potentially available, but had a probability of being chosen proportional to their abundance in the entire species pool.
- iii. *Environmental range-based assembly*: For a given sample, only those species whose ranges for the environmental factor encompass the value observed in that sample are eligible to be drawn and form the corresponding randomly generated sample. Potentially eligible species are drawn with uniform probability. Then, in this constrained way, we build a set of random synthetic samples, each characterized by a given value of the environmental factor and the same number of species as observed in each corresponding observed sample. For each species, the range of the environmental factor is defined by the minimum and maximum values of the variable across the samples in which the species is observed.
- iv. *Environmental range and abundance-based assembly*: As (iii), but, additionally, species whose ranges for the environmental factor contain the value of the empirical sample are randomly drawn proportionally to their abundances in the species pool.

For the purpose of discerning if significant clustering (or the lack of it) was influenced by the environmental gradient, we iteratively excluded the sample from the set with the highest value of the environmental variable of interest, also excluding from the pool, the corresponding species that appeared only in that sample, and recalculated the clustering index with the remaining samples (Fig. 1c, left). Additionally, we tested the significance of clustering patterns according to the four null models by applying the same sequential removal of samples, but first randomizing their salinity order (see Supplementary Materials for further details). These randomizations provided a 95% confidence interval for the clustering index (Fig. 1c, right) that allows you to decide if the observed clustering pattern as the environmental variable decreases is related to the environmental factor or not.

Statistics

Significant breakpoints of observed patterns along environmental gradients, either from ordination analyses or clustering indices, were estimated by means of the maximum F statistic derived from sequential Chow-tests (*sctest* function from *strucchange* package) [36] (see Fig. S3).

Results

We developed a detailed theoretical and methodological framework for the RTCC method (see Fig. 1, the methods

section for more details, and Fig. S1 for a summarized workflow of the methodology). To test the potential of this trait-based approach for gaining new insights about different community assembly mechanisms, we explored microbial communities along an environmental (salinity) gradient in a set of shallow saline ponds. We first evaluated 10 genomic traits that were available in public databases and that summarized genomic structural/functional variability, some of them known to respond to ecological adaptations (Fig. 2 and see Supplementary Material, section "Matching genomes"). Among them, the percentage of DNA dedicated to signal peptide synthesis exhibited the strongest clustering signal. Therefore, this trait, characterizing each species within a sample, was primarily used to further conduct the RTCC analyses. The pairwise differences for this trait were significantly smaller than those expected in synthetic communities randomly assembled from the pool according to the first null model (see "Methods" section below).

To assess the role of environmental factors in sorting species from empirical communities along the gradient, we analyzed the relationship between clustering patterns and two environmental variables with a battery of assembly null models (see the "Methods" section). Salinity and wind velocity should have different impacts on community assembly and, therefore, result in distinct patterns of clustering index values. We use wind as a control variable because it should represent a stochastic continuous perturbation, while salinity would provide a deterministic influence. We represented clustering index values by sequentially removing samples in decreasing order of the environmental variable (see Fig. 3a, b). We then tested the significance of the environmental variable by comparing the declining patterns with replicates of the same methodology based on random orderings (95% confidence intervals are shown as shaded areas). The rightmost point of the

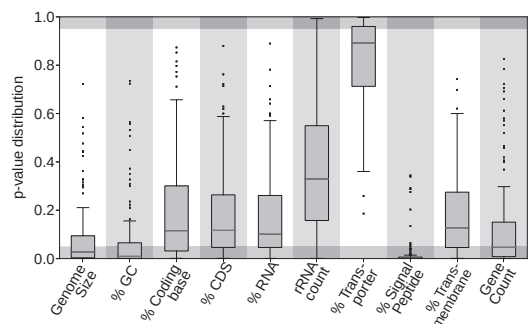


Fig. 2 Clustering distribution patterns for different functional traits. The p -value distributions obtained by the random assembly null model for 10 functional traits are shown. Colored stripes show the ranges of significant over-dispersion ($p \geq 0.95$) and significant clustering ($p \leq 0.05$). Signal peptide exhibit a strong signal of clustering

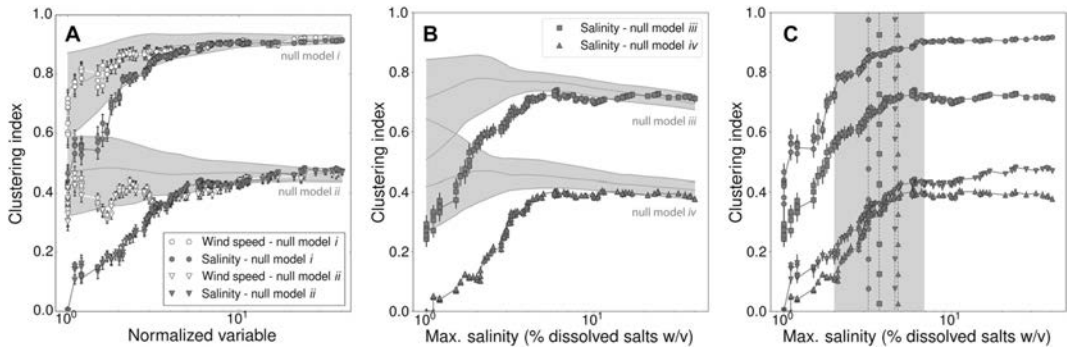


Fig. 3 Null models for community assembly. **a** Randomized signal peptide community clustering in salinity and maximum wind speed according to null models (i) and (ii), random assembly and abundance-based assembly, respectively. The clustering index for the signal peptide trait is plotted versus the maximum value of the environmental variable among all samples still present after every removal step. Samples are sequentially removed one-by-one in decreasing order of the variable. Note that wind velocities have been normalized so that their maximum values correspond to the maximum salinity observed. The shaded gray areas correspond to confidence intervals for the

different null models when the random removal of samples is repeated multiple times. **b** Same as **a** for null models (iii) and (iv), environmental range-based assembly, and environmental range and abundance-based assembly, respectively, which are meaningful only for salinity. **c** Clustering index curves are plotted again along with their thresholds, all of them lying within the rectangle (shaded in gray) yielded by the OTU-based approach, whose limits corresponds to the significant breakpoints of the observed pattern along the salinity gradient (Fig. S6)

curves corresponded to the clustering index for all samples in the data set. The random assembly null model yielded a clustering index value $h \sim 0.9$, which means that around a 90% of samples were not compatible with random assemblages, that is, showed an average trait dissimilarity significantly different from null model expectations. In other words, the vast majority of empirical samples exhibited a high degree of clustering in signal peptide compared to purely random assemblages. Such level of clustering declined when synthetic communities were built by selecting species with a probability proportional to relative abundances, and when ranges of salinity were taken into account in the assembly of simulated communities. As null model complexity increased, the fraction of samples compatible with the null hypothesis increased (see Fig. 3b). For the most constrained model, which takes into account both species abundances and salinity ranges, only at most ~30% of samples were not compatible with the null hypothesis underlying this model (see Fig. 3, lower curve). All null models showed a plateau in clustering index as the environmental variable decreased, but, at some point, a sharp decline occurred. The curves for salinity lay outside the confidence interval expected for randomized orderings, hence the null model was not able to explain the variation of clustering as salinity decreased. This result is consistent with a deterministic role of the salinity environmental gradient in community assembly. We repeated the same methodology using wind speed as a control environmental variable. Wind velocity showed a different pattern from salinity, that is, white curves laying within the shaded area in both tests (Fig. 3a). Since clustering values for ordered

sample removal lie within null model confidence intervals for randomized orderings, wind speed is not a deterministic driver of community assembly, but, as expected, plays a stochastic role (Fig. 3a). Although the OTU-based ordination approach revealed a weak positive correlation signal, this marginal correlation is well below the strong correlation observed for salinity (compare D and B panels in Fig. S6).

Chow tests were used to determine a threshold value in salinity over which environmental constraints emerge as a leading driver. We found percentages of salinity ranging from 3.2% for the random assembly null model and 4.8% for the ranges and abundances null model. These values lay within the range of salinity that presented significant breakpoints on the changing rate of communities in the ordination analysis (Fig. 3c, see also Fig. S6). We observed that traits other than signal peptide could be used to determine thresholds along the gradient (for example, the percentage of GC pairs in sequences; see Fig. S7). However, this was not necessarily true for any trait (for example, the percentage of DNA associated to transporter proteins, see also Fig. S7). Finally, signal peptide average values across genomes per sample (relative to the metacommunity average) were plotted along the salinity gradient. We observed that average trait values decreased as salinity increased (Fig. S8).

Discussion

Competitive exclusion increasing local extinctions of sub-adapted species and strict *environmental filtering* preventing

effective colonization of non-tolerant species along an environmental gradient, lie at two extremes of a continuum. Accurately disentangling these two main drivers is challenging. We developed a method to evaluate the effect of an environmental gradient on community trait structure, and provide a conceptual framework (Fig. 4) to interpret consistent discrete environmental thresholds that helps separate an assembly regime driven by a rich combination of several biotic and abiotic factors from a regime characterized by strong environmental constraints. The observed threshold unveils the switch point above which a structuring force starts dominating along the gradient (see the highest values of the clustering index along the plateau). Conversely, the sharp decline of the clustering index below the threshold value can be regarded as a signature of the rapid decay of the predominant role of a given environmental variable in driving local community assembly (Fig. 4). Thus, the lower the environmental pressure the lower the proportion of samples exhibiting significant clustering (i.e., the lower clustering index) due to the increased dispersion of trait values within local samples.

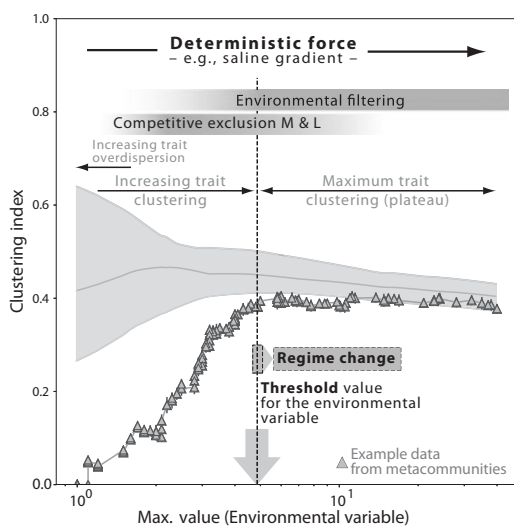


Fig. 4 Conceptual framework. A conceptual model for the interpretation of results based on the Randomized Trait Community Clustering (RTCC) method. We indicate the type of dominant deterministic forces acting along a gradient of a meaningful environmental variable. The picture shows the pattern observed along the salinity gradient by using null model (iv). The shaded gray area corresponds to confidence intervals for the same null model but randomizing sequential sample removal. A significant correlation of the trait clustering pattern with an environmental gradient exists if clustering index values for a given trait (under ordered sequential removal) lie outside the shaded area, this is, the null model fails to explain the variation of clustering along the gradient. The vertical line is the threshold salinity value separating the two community assembly regimes

Trait-based approaches have been widely used by ecologists to identify general community-level patterns of macro-organisms based on phenotypic characters [25, 37–39]. Some authors have made considerable efforts to introduce this framework in the study of microbial communities [22, 40–42], among others. Burke et al. [43] suggests that gene function, rather than the RNA taxonomy approach, holds the key to properly studying bacterial community assembly. Community composition may vary through several processes, while community functionality may remain stable due to functional redundancy of taxa [44, 45]. If stochastic forces are not too strong, selected species traits are expected to be differently favored and cluster around certain optimal values along a structuring environmental gradient. The distribution of trait values across species along a given gradient will be the result of a balance between environmental filtering and biotic interactions. Some traits can be unambiguously related to either stress-tolerance or competitive abilities. In this case, they can be used to distinguish between environmental filtering vs. competitive exclusion [17, 23] in spite of these two process producing the same trait clustering patterns [10].

Our findings indicate that the relative amount of genomic DNA coding for signal peptide shows a significant clustering pattern along the salinity gradient. Signal peptides are short peptides (5–30 aminoacids) present at the N-terminus of newly synthesized proteins that control the final destiny of these proteins, this is, whether they end up anchored to the cell membrane or are extracellularly secreted [46]. The diversity of these signaling systems decreases along salinity gradients in favor of the signal peptide-dependent *Tat* system, which dominates under high salinity conditions [47–49]. The observed decrease in average signal peptide values per sample along the salinity gradient (see Fig. S8) suggests that this trait could be involved in salinity stress tolerance, although confirming this requires further experimental research. Likewise, as salinity increases, we also observed higher similarity in GC content across locally co-occurring species (see Fig. S7). Several studies have suggested that hypersaline inhabitants are characterized by high GC content although there are some exceptions to these general patterns [50]. Moreover, it has been suggested that the stress effect of extreme environments is reflected in the nucleotide composition [40]. Other functional traits are not critically influenced by this gradient. In other words, some traits with relevant ecological roles may be related to other functional strategies not linked to the main environmental gradient, such as, for instance, the rRNA operon copy number, known to reflect bacterial growth rates [51], which are related to competitive ability [41].

Salinity has been described as one of the major environmental drivers of microbial community composition at a

global scale [52–54]. Important ecological changes have been described along salinity gradients from multipond solar salterns [55], with decreasing biodiversity as salinity increases [56], with an accompanying loss of metabolic processes [57]. Here, the specific trait pattern reveals the ecological functional adaptation of bacterial communities to the salinity gradient. Structural changes shown under a taxonomic view followed a similar pattern along the gradient (Fig. S6B). An increase in salinity beyond a threshold close to 5‰ triggers both a rapid change in community composition, and a decrease in mean relative signal peptide values (Fig. S8, upper panel). Maximum trait clustering was also reached at the same threshold. Above 5‰ salinity, community trait structure tends to stabilize, which means that there is functional redundancy in taxa while community composition still changes, but showing a slower turnover.

Between ca. 10‰ and 40‰ salinity range, community composition remained more stable. This shows that the RTCC method provided information not only about functional adaptation of communities—complementing results from their taxonomic turnover—but also about the accurate position of a salinity threshold (5‰) over which environmental constraints strongly shape community assembly. Since traits are selected based on the degree of clustering signal rather than on previous trait knowledge, our method opens new possibilities to objectively determine both whether or not and where a functional adaptation occurs along an environmental gradient.

It is reasonable to think that, in principle, if an environmental variable changes very smoothly, community composition, and, consequently, community trait structure, should also change smoothly. However, the opposite seems to be true in microbial communities along a salinity gradient. The RTCC accurately finds the position of a salinity threshold.

Ecological theory predicts that the end points of community assembly can be multiple [58]. The possibility that several stable species assemblages can exist under the same environmental conditions opens the door to observing abrupt transitions in community composition as a single environmental variable slowly changes. The study of the origins and underlying causes of these break points in community composition along environmental gradients deserves further empirical and theoretical work.

To conclude, we provide a new conceptual approach with the RTCC method that could be useful in guiding hypothesis-driven studies coping with the high complexity of biological systems. Although we used a data set from microbial communities, we highlight the potential of this method to be applied more generally to quantitative phenotypic or genotypic trait data measured for macro-organismal communities as well. A combination of theoretical modeling through trait-based analyses allows us to

go beyond description and deepen our mechanistic understanding of community assembly. Quantifying how community assembly is shaped by the environment is critical to predicting the ecological impact of environmental changes, and our approach provides a powerful tool to advance this analysis.

Acknowledgements This work was funded by the Spanish "Ministerio de Economía y Competitividad" under the projects BRIDGES (CGL2015-69043-P, to EOC and DA), and the Ramón y Cajal Fellowship program (DA). Both VJO and MM-S have been supported by Ph.D. contracts funded by the Spanish "Ministerio de Economía y Competitividad" under the projects BRIDGES and SITES (CGL2015-69043-P, CGL2012-39964 to EOC and DA). We thank Joan Cáliz and Gerard Funosas for help in bioinformatic analyses and fruitful discussions and the facilities and warm and creative atmosphere provided by the "white room" in the Computational Biology Lab (CBL) of the Center for Advanced Studies of Blanes.

Author contributions EOC, JAC, and DA conceived the study. EOC and XT-M carried out the sampling and field work. JAC, VJO and DA designed the null models. RO-A, MM-S, and XT-M conducted bioinformatic analyses. MM-S, VJO, JAC, and XT-M contributed with scripts and analyses. JAC and XT-M led the analyses of the data. XT-M, JAC, and MM-S assembled and organized the supplementary materials. XT-M, JAC, and DA led the writing of the manuscript. All authors contributed to developing the concepts along several "Bridges" workshops, discussed and contributed critically to the drafts and gave final approval for publication.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. van der Plas F, Janzen T, Ordóñez A, Fokema W, Reinders J, Etienne RS, et al. A new modeling approach estimates the relative importance of different community assembly processes. *Ecology*. 2015;96:1502–15.
2. Vellend M. Conceptual synthesis in community ecology. *Q Rev Biol*. 2010;85:183–206.
3. Vellend M. The theory of ecological communities (MPB-57). Princeton University Press; 2016.
4. Gravel D, Canham CD, Beaudet M, Messier C. Reconciling niche and neutrality: the continuum hypothesis. *Ecol Lett*. 2006; 9:399–409.
5. MacArthur RH. Species packing and what interspecific competition minimizes. *Proc Natl Acad Sci USA*. 1969;64:1369–71.
6. Hubbell SP. The unified theory of biodiversity and biogeography. Princeton: Princeton University Press; 2001.
7. Tilman D. Resource competition and community structure. Princeton, NJ: Princeton University Press; 1982.
8. Law R, Morton RD. Permanence and the assembly of ecological communities. *Ecology*. 1996;77:762–75.
9. Law R, Weatherby AJ, Warren PH. On the invasibility of persistent protist communities. *Oikos*. 2000;88:319–26.

10. Mayfield MM, Levine JM. Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecol Lett*. 2010;13:1085–93.
11. Chesson PL. Mechanisms of maintenance of species diversity. *Annu Rev Ecol Syst*. 2000;31:343–66.
12. Liebholt AM, Chase J. Metacommunity ecology. Princeton: Princeton University Press; 2018.
13. Bertness M, Callaway RM. Positive interactions in communities. *Trends Ecol Evol*. 1994;9:191–3.
14. Maestre FT, Callaway RM, Valladares F, Lortie CJ. Refining the stress-gradient hypothesis for competition and facilitation in plant communities. *J Ecol*. 2009;97:199–205.
15. Bastolla U, Fortuna MA, Pascual-García A, Ferrera A, Luque B, Bascompte J. The architecture of 290 mutualistic networks minimizes competition and increases biodiversity. *Nature*. 2009;458:1018–20.
16. Helmus MR, Savage K, Diebel MW, Maxted JT, Ives AR. Separating the determinants of phylogenetic community structure. *Ecol Lett*. 2007;10:917–25.
17. Kraft NJB, Adler PB, Godoy O, James EC, Fuller S, Levine JM. Community assembly, coexistence and the environmental filtering metaphor. *Funct Ecol*. 2014;29:592–9.
18. Peres-Neto PR, Dray S, Braak CJ. Linking trait variation to the environment: critical issues with community-weighted mean correlation resolved by the fourth-corner approach. *Ecography*. 2017;40:806–16. arXiv:arXiv:0902.0132v1
19. Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, et al. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol Rev*. 2013;88:15–30.
20. Townsend Peterson A, Soberón J, Pearson RG, Anderson RP, Martínez-Meyer E, Nakamura M, et al. Ecological niches and geographic distributions. Oxford: Princeton University Press; 2011.
21. Barberán A, Casamayor EO, Fierer N. The microbial contribution to macroecology. *Front Microbiol*. 2014;5:203.
22. Dini-Andreote F, Stegen JS, van Elsas JD, Salles JF. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci USA*. 2015;112:E1326–E1332.
23. Goberna M, Navarro-Cano JA, Valiente-Banuet A, García C, Verdú M. Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecol Lett*. 2014;17:1191–201.
24. Barberán A, Ramirez KS, Leff JW, Bradford MA, Wall DH, Fierer N. Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecol Lett*. 2014;17:794–802.
25. Ter Braak CJ, Cormont A, Dray S. In the fourth-corner problem reports. *Ecology*. 2012;93:1525–6.
26. Edwards KF, Litchman E, Klausmeier CA. Functional traits explain phytoplankton responses to environmental gradients across lakes of the United States. *Ecology*. 2013;94:1626–35.
27. Etienne RS, Alonso D. A dispersal-limited sampling theory for species and alleles. *Ecol Lett*. 2005;8:1147–56. arXiv.
28. Webb CO, Ackerly DD, McPeck MA, Donoghue MJ. Phylogenies and community ecology. *Annu Rev Ecol Systemat*. 2002;33:475–505.
29. Gotelli NJ, Graves GR. Null models in ecology. Washington: Smithsonian Institution Press; 1996.
30. Casamayor EO, Triadó-Margarit X, Castañeda C. Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. *FEMS Microbiol Ecol*. 2013;85:503–18.
31. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods*. 2013;10:996–8.
32. Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, et al. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res*. 2014;42:581–91.
33. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, et al. IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res*. 2014;42:D560–7.
34. Langille MGI, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnol*. 2013;31:814–21.
35. Nemergut DR, Knelman JE, Ferrenberg S, Bilinski T, Melbourne B, Jiang L, et al. Decreases in average bacterial community rRNA operon copy number during succession. *ISME J*. 2015;10:1147–56.
36. Zeileis A, Kleiber C, Krämer W, Hornik K. Testing and dating of structural changes in practice. *Comput Stat Data Anal*. 2003;44:109–23.
37. McGill BJ, Enquist BJ, Weiher E, Westoby M. Rebuilding community ecology from functional traits. *Trends Ecol Evol*. 2006;21:178–85.
38. Adler PB, Fajardo A, Kleinhesselink AR, Kraft NJB. Trait-based tests of coexistence mechanisms. *Ecol Lett*. 2013;16:1294–306.
39. Thakur MP, Wright AJ. Environmental filtering, niche construction, and trait variability: the missing discussion. *Trends Ecol Evol*. 2017;32:884–6.
40. Barberán A, Fernández-Guerra A, Bohannan BJM, Casamayor EO. Exploration of community traits as ecological markers in microbial metagenomes. *Mol Ecol*. 2012;21:1909–17.
41. Krause S, Le Roux X, Niklaus PA, Van Bodegom PM, Lennon JT, Bertilsson S, et al. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front Microbiol*. 2014;5:251.
42. Ortiz-Álvarez R, Fierer N, de los Ríos A, Casamayor EO, Barberán A. Consistent changes in the taxonomic structure and functional attributes of bacterial communities during primary succession. *ISME J*. 2018;12:1658–67.
43. Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T. Bacterial community assembly based on functional genes rather than species. *Proc Natl Acad Sci*. 2011;108:14288–93.
44. Grossmann L, Beisser D, Bock C, Chatzinotas A, Jensen M, Preisfeld A, et al. Trade-off between taxon diversity and functional diversity in European lake ecosystems. *Mol Ecol*. 2016;25:5876–88.
45. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and functional redundancy in microbial systems. *Nat Ecol Evol*. 2018;2:936–43.
46. von Heijne G. Signal sequences: the limits of variation. *J Mol Biol*. 1985;184:99–105.
47. Wesley RR, Brüser T, Kissinger JC, Pohlschröder M. Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol*. 2002;45:943–50.
48. van der Ploeg R, Mäder U, Homuth G, Schaffer M, Denham EL, Monteferrante CG, et al. Environmental salinity determines the specificity and need for tat-dependent secretion of the YwbN protein in *Bacillus subtilis*. *PLoS One*. 2011;6:1–10.
49. van der Ploeg R, Monteferrante CG, Piersma S, Barnett JP, Kouwen TRHM, Robinson C, et al. High-salinity growth conditions promote tat-independent secretion of tat substrates in *Bacillus subtilis*. *Appl Environ Microbiol*. 2012;78:7733–44.
50. Paul S, Bag SK, Das S, Harvill ET, Dutta C. Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol*. 2008;9:R70.

51. Klappenbach JA, Dunbar JM, Schmidt TM. rRNA operon copy number reflects ecological strategies of bacteria. *Appl Environ Microbiol.* 2000;66:1328–33.
52. Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci USA.* 2007;104:11436–40.
53. Auguet J-C, Barberán A, Casamayor EO. Global ecological patterns in uncultured Archaea. *ISME J.* 2010;4:182–90.
54. Barberán A, Casamayor EO. Global phylogenetic community structure and beta-diversity patterns in surface bacterioplankton metacommunities. *Aquat Microb Ecol.* 2010;59:1–10.
55. Gasol JM, Casamayor EO, Joint I, Garde K, Gustavson K, Benlloch S, et al. Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat Microb Ecol.* 2004;34:193–206.
56. Casamayor EO, Triadó-Margarit X. Microbial diversity and novelty along salinity gradients. New York: Springer; 2013. p. 1–8.
57. Oren A. Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev.* 1999;63:334–48.
58. Bunin G. Ecological communities with Lotka–Volterra dynamics. *Phys Rev E.* 2017;95:1–8.

Related works



Contents lists available at ScienceDirect

Science of the Total Environment

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

Microbial composition, potential functional roles and genetic novelty in gypsum-rich and hypersaline soils of Monegros and Gallocanta (Spain)

Mateu Menéndez-Serra ^a, Xavier Triadó-Margarit ^a, Carmen Castañeda ^b,
Juan Herrero ^b, Emilio O. Casamayor ^{a,*}

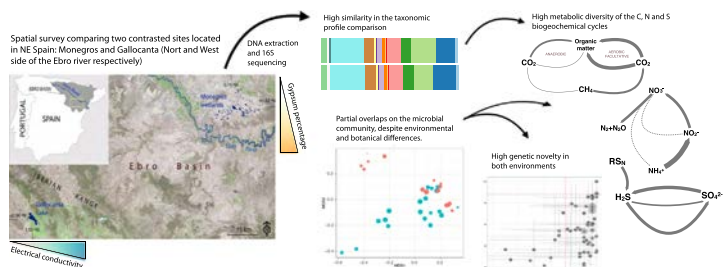
^a Integrative Freshwater Ecology Group, Centro de Estudios Avanzados de Blanes (CEAB-CSIC), Acces Cala Sant Francesc 14, Blanes 17300, Spain

^b Estación Experimental de Aula Dei (EEAD-CSIC), Av. Montañana 1005, 50059 Zaragoza, Spain

HIGHLIGHTS

- Gypsum content ranging from 1% up to 92% do not significantly affect the microbial composition.
- Similar microbial composition and functionality exist despite of the different geochemical and botanical environments.
- High genetic novelty was found highlighting the value of the microbial heritage in saline and gypsum rich soils.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 23 July 2018

Received in revised form 3 September 2018

Accepted 4 September 2018

Available online 05 September 2018

Editor: Charlotte Poschenrieder

Keywords:

Microbial community

Gypsum

Hypersaline soils

Genetic novelty

ABSTRACT

Soil microbial communities (both Bacteria and Archaea) were studied after 16S rRNA genes massive sequencing in two hypersaline and gypsum-rich contrasted sites located in NE Spain. Soil microbial communities were also locally analysed according to environmental variables, including geological, physico-chemical, biogeochemically, and climatic data. Typical soil characteristics, climate data, and plant composition clearly split the two sites and major differences among the microbial communities for the areas were initially expected. Overall, high values of microbial species richness (up to 2300 taxa) and ecological diversity was detected in both sites. High genetic novelty levels were found mostly to environmental sequences, highlighting the high potential for microbiological studies. In contrast to the initial expectations, a substantial overlapping between Monegros and Gallocanta microbes was observed, indicating a high similarity despite of the geographical, botanical and environmental distances between sites, in agreement with both high dispersal and local selection inherent to the microbial world. The potential biogeochemical cycling showed small differences between sites, with presence of photosynthetic green and purple sulfur bacteria, cyanobacteria and aerobic and anaerobic chemolithotrophs. Potential for aerobic methane oxidation and anaerobic methanogenesis was observed in both sites, with predominance of potential nitrification mostly by ammonia-oxidizing archaea, nitrite oxidation and denitrification, and minor contribution for nitrate reduction and nitrate ammonification. The predicted functions based on the taxonomic composition showed high overlapping between the two studied regions, despite their difference in gypsum richness.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Soils are one of the most complex and taxon-rich microbial habitats on Earth (Torsvik et al., 1990; Fierer and Jackson, 2006; Barberán and

* Corresponding author.

E-mail address: casamayor@ceab.csic.es (E.O. Casamayor).

Casamayor, 2014) and it is well known that soil microorganisms play fundamental roles in the biological function, physico-chemical structure and fertility and sustainability of soil ecosystems. Local soil biodiversity is strongly driven by spatial heterogeneity (Bardgett and van der Putten, 2014), although the presence of big scale patterns as latitudinal gradients in soil diversity is still unclear (Andam et al., 2016). Soils of hot arid regions cover 31.5% of the world and compare in diversity with those present in humid regions (Dregne, 1976). However, the potential for biodiversity conservation in arid and semiarid environments is still poorly documented despite the fact that the ecosystem vulnerability to desertification and land degradation is worldwide recognized for such habitats, and concerns for environmental protection have raised in the last decades (Williams, 2002). In these regions, soils support high-value habitats threatened by landscape transformations, and pose a constraint to intensive agricultural production (Davidson, 2014; Kingsford et al., 2016; Greed et al., 2017).

Soils associated to saline continental wetlands usually occur in endorheic drainage basins, which approximately cover 1/10 of the Earth's surface area (Waiser and Roberts, 2009), and have been largely

ignored until recently (Williams, 2002). The behaviour and ecology of hypersaline soils associated to saline athalassohaline lakes (dissolved salts composition substantially different from marine waters), are dominated by high concentrations of highly soluble salts with major effects on the abundance and growth of microbes (Oren, 1999; Gasol et al., 2004). Under these conditions, organisms cope with osmotic stress and ion-specific toxicity, in addition to other frequent unfavourable conditions, like alternation of long lasting dry and wet periods, high UV radiation doses, or extreme temperatures including periods with liquid water below 0 °C due to the freezing-point depression in the saline solutions. Also in arid regions, gypsum-rich soils pose other different limitations for life. Gypsum soils are usually related to marginal lands and a lack of enough attention for biodiversity conservation concerns and for studies of life under extreme conditions (Connon et al., 2007; Azua-Bustos et al., 2012; Palacio et al., 2014). While gypsum ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) does not cause appreciable toxicity or ionic stress for most organisms, the low retention capacity of water and nutrients of gypsum soils is a major drawback for life. The global surface of gypseous lands can be estimated from one to two million square kilometers

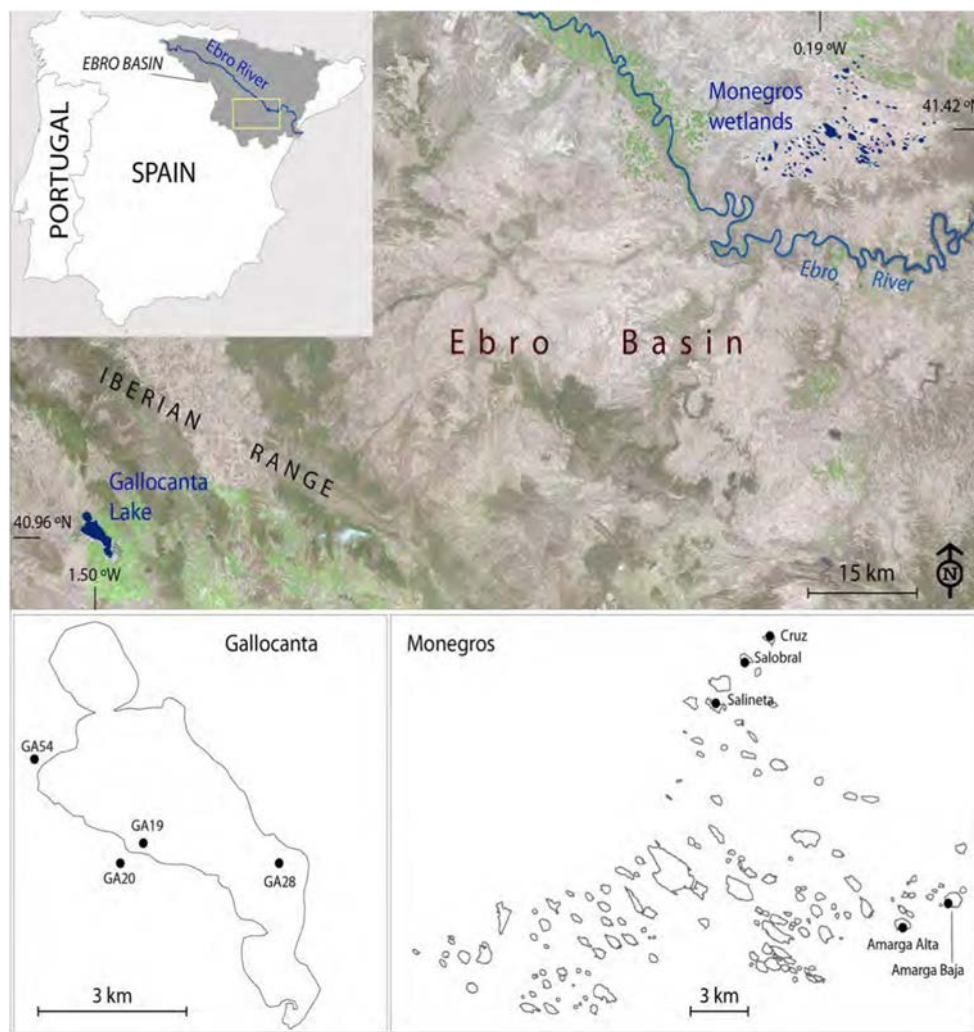


Fig. 1. Location of Monegros and Gallocanta Ramsar sites in the Ebro basin (inset) and soil sites sampled. SPOT 5 image from 2009 (RGB 123) as background.

(Casby-Horton et al., 2015). These marginal soils have received limited attention from researchers and gypsum-rich soils in arid lands remain often unknown on the biodiversity they host, including endemisms.

In this work, we explored the soil microbial composition, potential functional roles, and genetic novelty in two hypersaline and gypsum-rich contrasted sites located in NE Spain, that represent good examples of large arid saline environments in Europe. Europe is second after Africa in the number of inland intermittent saline wetlands included in the Ramsar list, and almost 38% of these sites are in Spain (Ramsar Convention Secretariat, 2010). Major differences on the microbial communities were expected between sites due to its high environmental idiosyncrasies, and also along both the electrical conductivity and gypsum percentage gradients. Microbial communities (both Bacteria and Archaea) were studied after 16S rRNA genes massive sequencing and were taxonomically identified. The results were compared with the previously reported idiosyncratic plant species composition. Soil microbial communities were also analysed according to local environmental variables, including geological, physico-chemical, biogeochemistry, and climatic data.

2. Materials and methods

2.1. Description of the Monegros and Gallocanta study zones

Gallocanta and Monegros, NE Spain (Fig. 1) are hypersaline Ramsar wetlands protected under the EU Birds and Habitats Directives, and are subject to agricultural intensification. They provide habitats for rare and threatened plants and animals, and many studies have unveiled a large number of endemisms mainly of invertebrates, vascular plants, lichens, and bryophytes (Braun-Blanquet and Bolòs, 1958; Casamayor et al., 2013 and references therein). Gallocanta Lake, with 14.4 km², is the largest well-preserved saline lake in Western Europe. It is located on the west side of the Ebro River, at 1000 m a.s.l. in the Iberian Range (Fig. 1). Climate is dry semiarid with a mean rainfall of 488 mm yr⁻¹, a mean annual temperature of 11.2 °C, and a mean hydric deficit of 605 mm yr⁻¹ (García-Vera and Martínez-Cob, 2004). The Monegros Desert wetlands are located within the Central Ebro Basin, on the east side of the Ebro River at a mean elevation of 328 m a.s.l. The territory is an endorheic platform (c. 400 km²) with 149 scattered saline depressions ranging in size from <2 to 239 ha. Some of these depressions host hypersaline wetlands. This is one of the most arid regions in Europe, with a mean annual rainfall of 350 mm, a mean annual temperature of 14.9 °C, and 1255 mm yr⁻¹ of mean reference evapotranspiration (Faci and Martínez-Cob, 1991). Soils developed in these lake basins, under arid and semiarid climates, frequently become seasonally or intermittently dry due to the limited precipitation and high evapotranspiration rates (Castañeda et al., 2015, 2017).

2.2. Sampling and environmental data

We sampled soil from five sites in Monegros (i.e., Amarga Alta, Amarga Baja, Cruz, Salobral and Salineta) and four sites in Gallocanta (see details in Fig. 1). Sample sites were selected in order to cover the

largest environmental gradients on gypsum percentage and electrical conductivity. Soil samples were collected from auger holes and pits opened (Table 1). The plant species compositions were obtained from Conesa et al. (2011) and Gómez et al. (1983) for Monegros and Gallocanta, respectively (Supplementary Table S1). Geological data were obtained from the Spanish Geological Survey maps MAGNA series, scale 1:50,000. Climatic data were provided by the Spanish Meteorological Agency (AEMET), and elevation data were obtained from the LiDAR derived digital elevation model supplied by the Spanish Geographical Institute, with a vertical precision of 0.2 m. Additional ancillary data on soils composition and agricultural practices are also shown in supplementary Table S2.

2.3. Samples processing

Overall, 34 soil samples were taken in Monegros (19 samples) and Gallocanta (15 samples), to cover the widest range of saline and gypsum contents (Fig. 2). Samples were collected from different depths, from 5 to 150 cm. Soil samples were air-dried at room temperature and sieved through a 2-mm mesh. Soil salinity was measured as the electrical conductivity of the 1:5 soil:water extract (EC1:5) and the soil saturation extract (ECe), and expressed in dS m⁻¹ at 25 °C. Soil pH was measured with a pH electrode (Orion 9157BNMD) in the 1:2.5 soil:water extract and in the saturated paste. Calcium carbonate equivalent (CCE) was quantified by gasometry and gypsum content was quantified by thermogravimetry (Artieda et al., 2006).

2.4. DNA extraction and sequencing

For DNA analyses, 0.25 g of soil was used following the PowerSoil® DNA Isolation Kit (Qiagen) manufacturer's protocol. Purified DNA samples were amplified and sequenced through Illumina MiSeq 2 × 250 following protocols of the RTSF-MSU genomic core facilities (Michigan State University, USA. <https://rtsf.natsci.msu.edu/>). The V4 regions of the 16S rRNA genes (ca. 250 nucleotides, both Bacteria and Archaea) were amplified with the primers F515 (5'-GTGCCAGCAGCGCGCGTAA-3') and R806 (5'-GGACTACHVGGGTWTCTAAT-3') (Caporaso et al., 2010). The whole gene sequence dataset is available in the NCBI Sequence Read Archive under BioProject ID: PRJNA420947 (<http://www.ncbi.nlm.nih.gov/bioproject/420947>), and under accession numbers from SAMN08126969 to SAMN08127002 (see Supplementary Table S3 for details).

2.5. 16S rRNA gene sequences analyses

Raw sequences were processed using the UPARSE pipeline (Edgar, 2013). After excluding samples with <5000 reads, the dataset reached 1,683,696 sequences, 74.3% of which had ≥Q30. After merging of read pairs the dataset was filtered by setting a maximum number of expected errors of 0.25, and discarding reads with higher error probabilities (Edgar and Flyvbjerg, 2015). As a result, 819,379 sequences (93.9%) with a length of 250 bp were kept for the downstream analysis. Chimeric reads were discarded by UCHIME following two strategies: de novo and reference-

Table 1
Dates, location and sampling information for soils analysed in Gallocanta and Monegros.

Sampling date (dy/mo/yr)	Coord_X UTM ED50	Coord_Y UTM ED50	Sites	Sampling point	Sampling type	Description
23/10/10	625910,41	4535971,64	GALLOCANTA	GA19	Pit	Halophytes fringe
23/10/10	625698,73	4535611,18	GALLOCANTA	GA20	Pit	Halophytes fronge
01/09/2011	628205,18	4535807,17	GALLOCANTA	GA28	Pit	Saline bare soil
17/07/2014	623531,16	4537468,42	GALLOCANTA	GA54	Auger hole	Vegetation fringe
7/10/10	746154,15	4587148,85	AMARGA ALTA	AMA2	Auger hole	Saline bare soil
7/10/10	748606,60	4588197,24	AMARGA BAJA	AMB2	Auger hole	Saline bare soil
27/3/13	739808,39	4599264,49	CRUZ	CRZ	Auger hole	Saline bare soil
22/5/14	738654,14	4598426,39	SALOBRAL	SLB	Auger hole	Saline bare soil
22/5/14	737212,68	4596429,08	SALINETA	SLN	Auger hole	Saline bare soil

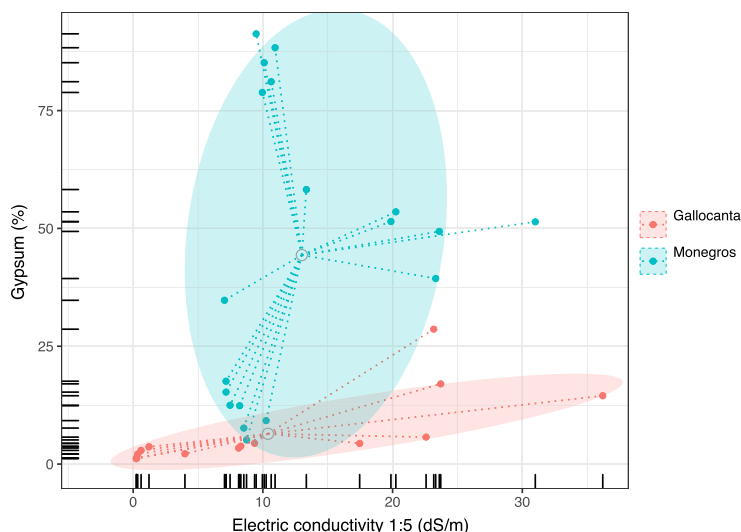


Fig. 2. Environmental distribution of the 34 saline soil samples analysed according to electrical conductivity and gypsum content. Ellipses represent confidence interval at 95% for Monegros (blue dots) and Gallocanta (red dots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

based (against 'GOLD' reference database available for ChimeraSlayer, available on https://drive5.com/uchime/uchime_download.html). Filtered sequences were clustered into Operational Taxonomic Units (OTUs) by using a cut-off of 97% identity threshold. A total of 12,292 OTUs were obtained. The taxonomic assignment was carried out with SINA aligner v.1.2.11 (Pruesse et al., 2012), using SILVA_128 reference database (Gurevich et al., 2013) with the de-replicated version at 99% sequence identity (SSURef_NR99_128_SILVA). OTUs were filtered by their alignment quality score and sequences classified as mitochondria and chloroplasts were excluded. A final filtering by library size was carried out, keeping only samples with >10,000 reads. The resulting OTU table consisted of 10,805 OTUs for 34 samples. Data were normalized by rarefying the reads of all samples to the minimum threshold previously established, 10,000 reads/sample. Rarefaction at the same depth was repeated 100 times in order to avoid the loss of less abundant OTUs, and the resulting rarefactions were unified on an average rarefied OTU table (Caliz et al., 2015).

2.6. Statistical analysis

Statistical analyses were run in R environment (<http://www.r-project.org>). Community ecology related parameters as species richness or Shannon-Weaver diversity index were calculated using the vegan package (Oksanen et al., 2016). Bray-Curtis index after Hellinger standardization was used to create a bacterial community distance matrix which was further used for the Non-metric multidimensional scaling analysis run with MetaMDS functions in vegan package. The distance matrix was also used to calculate the Analyses of Similarities (ANOSIM). Permutational Multivariate Analysis of Variance Using Distance Matrix (adonis test) based on 1000 permutation was used to test for differences between groups belonging to different factor categories and variations along environmental gradients. The Gower General Coefficient of Similarity (Gower, 1971), which allows the calculation of a distance matrix using both logical and quantitative data, was used to calculate distance matrices based on the vegetal composition and environmental variables, including geological, physico-chemical, biogeochemical, and climatic data. The same weight was assigned to all variables when distances between

samples were calculated. The resulting distance matrices were represented using a Classical multidimensional scaling (MDS) (Gower, 1966) with the function cmdscale in the stats package. Spearman's Rank order correlation coefficient was carried out to test monotonic relationships between numeric variables.

2.7. Microbial genetic novelty

The 16S rRNA gene novelty was tested by BLAST search (GenBank survey carried out on June 2017) to both the closest environmental match (CEM) and the closest cultured match (CCM) available in global databases. OTUs matching <97% identity with databases ("species-level" identity threshold) were considered highly novel (Casamayor et al., 2013). A phylogenetic tree including the most abundant and novel OTUs (identity between 90 and 95% with CCM) was carried out with ARB 6.0. (Ludwig et al., 2004) in SILVA_128 reference database. OTUs sequences were inserted in the optimized tree provided by default using the maximum parsimony method.

2.8. Functional groups

A preliminary analysis on the functional groups potentially involved in the carbon (C), nitrogen (N), and Sulfur (S) biogeochemical cycling under both anaerobic and aerobic conditions, was carried out according the relative abundance of previously identified target taxa (Belser, 1979; Borrel et al., 2014; Caliz et al., 2015; Carlson and Ingraham, 1983; Iguchi et al., 2015; Llorens-Marès et al., 2015; Lückner et al., 2013; Tardy-Jacquenod et al., 1998; Torrentó et al., 2010; Verbaendert et al., 2011; Vigneron et al., 2017; Yao and Peng, 2017). For the C cycle, specific targeted groups were photo- and chemoautotrophs, methanogens and aerobic methane oxidizers. For the N cycling, we targeted potential denitrification, nitrate reduction, nitrate ammonification, nitrification (ammoxia-oxidizers archaea – AOA- and bacteria – AOB-, Fernández-Guerra and Casamayor, 2012), and nitrite oxidation (NOB). Finally, for the S cycling, we focused on groups involved in sulfur oxidation, dissimilatory sulfate reduction, and polysulfide reduction.

3. Results

3.1. Community composition and environmental variables

The soil samples dataset from the two sites covered a wide range of environmental variables mainly for gypsum content (from 1 to 91%), and electrical conductivity 1:5 soil:water extract (from 0.24 to 36.20 dS m⁻¹), whereas pH of the saturated paste showed a narrow interval and ranged between 7.71 and 8.57. A more heterogeneous electrical conductivity distribution was observed for samples from Gallocanta, whereas a wider variation for the gypsum content was observed in Monegros (Fig. 2). Overall, high values of microbial species richness (from ~850 to ~1800 OTUs) and Shannon-Weaver diversity index (from 5.4 to 6.4) were found in both sites (see details in Supplementary Table S3). The multivariate analysis showed that microbial composition was only directly affected by the local sampling location (adonis, $R^2 = 0.42$, p -value < 0.001). In contrast to the initial expectations, other factors such as the regional comparison –Monegros and Gallocanta– (adonis, $R^2 = 0.07$, p -value < 0.01), gypsum content (adonis, $R^2 = 0.09$, p -value < 0.01), sample depth (adonis, $R^2 = 0.04$, p -value > 0.05), and electrical conductivity (adonis, $R^2 = 0.08$, p -value < 0.01), showed negligible contributions as drivers of microbial composition. A series of correlation tests did not show significant correlations between any tested factor and ecological parameters (Spearman, p -value > 0.05, data not shown).

As expected, a strong segregation between Gallocanta and Monegros was observed for the environmental variables tested (Supplementary Table S2), including geological, physico-chemical, biogeochemical, and climatic data (Gower General Coefficient of Similarity, Fig. 3A). Some sampling points showed specific properties, such as 'Gallocanta 28' because of its lower electrical conductivity values and presence of gypsum, or 'Salineta' separated from the other sampling points in Monegros because of the substantially higher electrical conductivity (Supplementary Table S2). The same strong segregation detected for environmental variables was observed for the plant species composition (Fig. 3B, data available in Supplementary Table S1). Unexpectedly, the ordination analysis for bacterial taxa composition (OTU level) showed a substantial overlapping between Monegros and Gallocanta sites (Fig. 3C), indicating a high similarity despite of the geographically and environmentally distant microbial communities, ruling out our initial hypothesis of major differences expected for the microbial communities between sites.

3.2. Microbial taxonomic and functional composition

Samples were grouped at different taxonomic resolutions for Monegros and Gallocanta to compare the whole community composition (both Bacteria and Archaea). High rank taxonomic levels showed a high homogeneity with common predominance of a few Phyla (Fig. 4A), i.e., Acidobacteria (range 14–17% between sites), Actinobacteria (18–13%), Bacteroidetes (7–8%), Chloroflexi (11–10%), Planctomycetes (7–8%), and Proteobacteria (24–25%). The same overlapping was observed both at the Class and the Order levels (Fig. 4C and D, respectively). In agreement, the most abundant OTUs at both sites were the same, i.e., unclassified Acidobacteria Subgroup 6 and uncultured Anaerolineaceae (Chloroflexi). Overall, 24 out of the 30 most abundant OTUs were shared between Monegros and Gallocanta communities (rank-abundance curves shown in Supplemental Fig. S1).

The potential biogeochemical cycling for Monegros and Gallocanta soils according to targeted microorganisms showed small differences between sites (Fig. 5). Autotrophic synthesis of organic matter was potentially carried out both aerobically and anaerobically, with presence of photosynthetic green and purple sulfur bacteria, cyanobacteria and aerobic and anaerobic chemolithotrophs. Potential for aerobic methane oxidation and anaerobic methanogenesis was observed in both sites. For the sulfur cycling, a balance between all targeted pathways was

observed in Monegros whereas Gallocanta showed a predominance of sulfur oxidation potential (i.e., 67%, in terms of abundance of the taxa potentially involved in each S pathway). For nitrogen cycling, predominance of potential nitrification mostly by ammonia-oxidizing archaea (AOA), nitrite oxidation and nitrogen assimilation, and minor contribution for nitrate reduction and nitrate ammonification was observed.

3.3. Genetic novelty

Overall, high genetic novelty (i.e., identity <97% against both CEM –closest environmental match– and CCM –closest cultured match– databases) was found for c. 10% of the retrieved 16S rRNA gene sequences (i.e., 13% in Monegros, Fig. 6A, and 9.4% in Gallocanta, Fig. 6B) that reached c. 3% of the total abundance (i.e., 2.3% Monegros, and 3.6% Gallocanta). Most (i.e., >70%) of the OTUs in Monegros and Gallocanta placed into the cultured gap plot area (i.e., those sequences showing >97% sequence identity for CEM and <97% for CCM), and they were the most important in terms of total abundance (c. 60% in both sites). The mean percentage of identity in Monegros and Gallocanta against CCM was close to 92%, highlighting the great potential to bring into culture substantially new microbes still. This lack of culture representatives was apparent even for some of the most studied bacterial Phyla such as Proteobacteria, Firmicutes, and Cyanobacteria (Supplementary Fig. S2). OTUs already detected in previous studies (>97% identity for CEM) were more likely to be shared by Monegros and Gallocanta than those considered more novel (<97% identity for CEM) (Supplementary Fig. S3), suggesting again high similarity between sites but with the presence of some site-specific populations.

4. Discussion

Microbial communities are fundamental components of the soil and its functioning, and soil microbial ecology studies are abundant and diverse in the scientific literature (Bardgett and van der Putten, 2014). Despite the fact that aridity may reduce soil microbial diversity and abundance at the global scale (Maestre et al., 2015), biodiversity studies are still of interest because of the physiological adaptations and ecological peculiarities to thrive in the arid and saline realms. In the case of plant communities, it is well established the presence of idiosyncratic species adapted to aridity, and the importance of drylands for landscape and biodiversity conservation policies (SBSTTA, 1999). In our study, the analysis of the soil microbial composition in gypsum-rich and hypersaline soils of Monegros and Gallocanta has evidenced its ecological importance with high values of diversity and species richness similar to those usually found in highly relevant microbial ecosystems such as microbial mats (Bolhuis et al., 2014). Monegros and Gallocanta showed also a similar microbial composition, despite of the geographical and environmental (i.e., geology, physico-chemical parameters, biogeochemistry, and climatic conditions) distances between them. This fact is in agreement with both high dispersal and local selection inherent to the microbial world (Barberán et al., 2014). However, still a substantial number (i.e., 10%) of the microorganisms found showed a high genetic novelty, potentially suggesting idiosyncratic local combinations for the two study areas that deserve additional microbiological studies. These environments are of especial interest for microbial cultivation because >70% of the OTUs had not closely related cultured counterparts. Interestingly, many of these not previously cultured (and in some cases not previously detected) microbes are found at high relative abundances, and belonged to some of the most studied bacterial taxa such as Proteobacteria, Chloroflexi and Bacteroidetes, although this result should be substantiated with larger length sequences for more accurate phylogenetic analysis. Other less abundant potentially highly novel OTUs matched important and well-studied taxa in soils such as Cyanobacteria and Firmicutes. The survival of these islands of biodiversity and biological speciation is however threatened by many kinds of urban or agricultural developments combined with the scarcity of

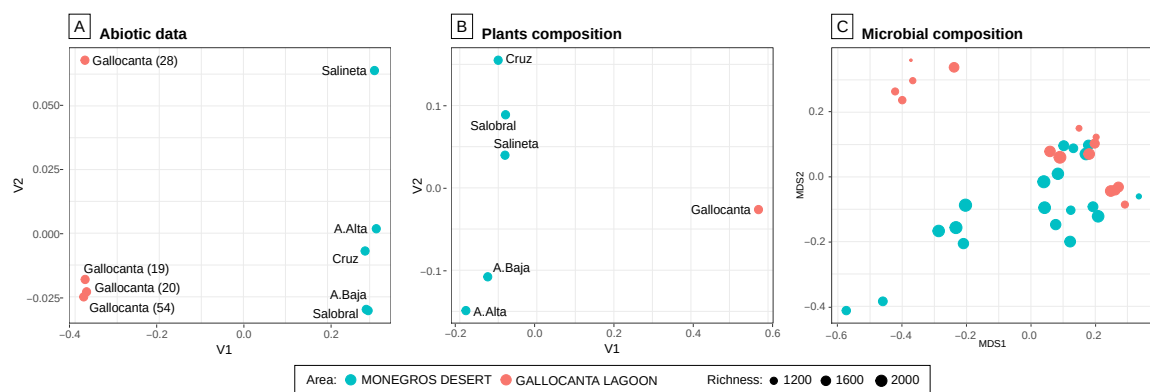


Fig. 3. Multidimensional scaling ordination analyses showing the dissimilarity between Monegros (green dots) and Gallocanta (red dots) samples according to environmental (A), plant species (B) and microbial composition (C) datasets. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

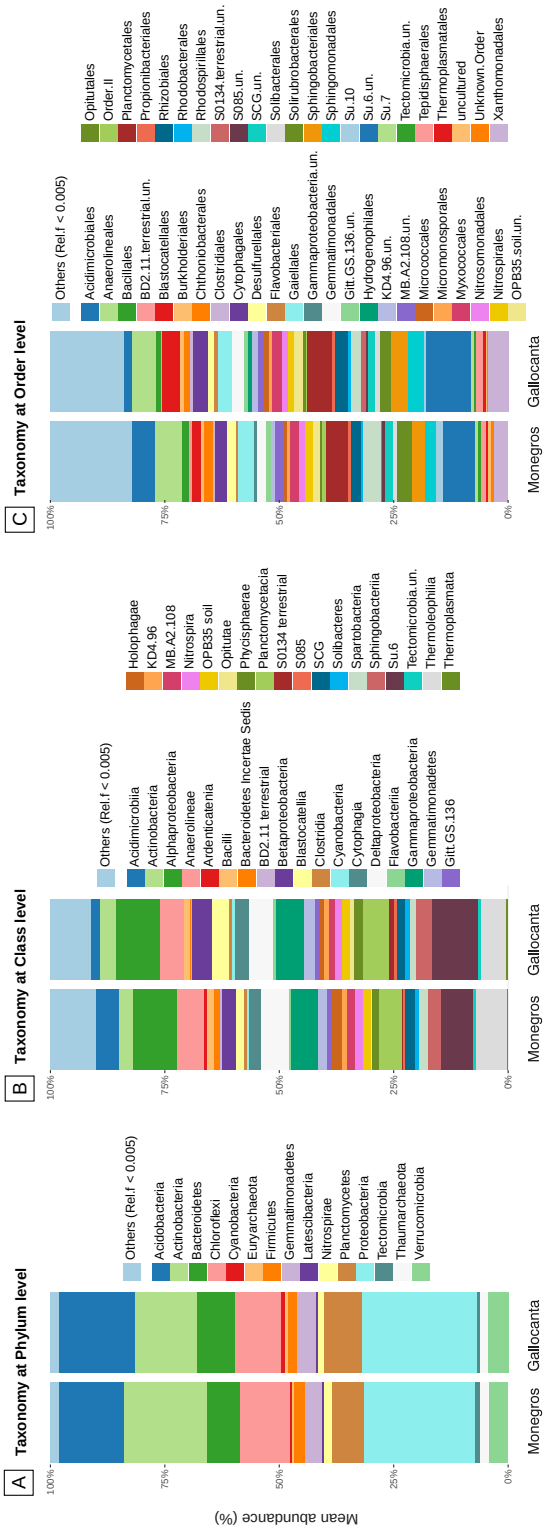


Fig. 4. Taxonomic composition of the prokaryotic microbial communities of Monegros and Gallocañta shown at Phylum (A), Class (B) and Order (C) levels.

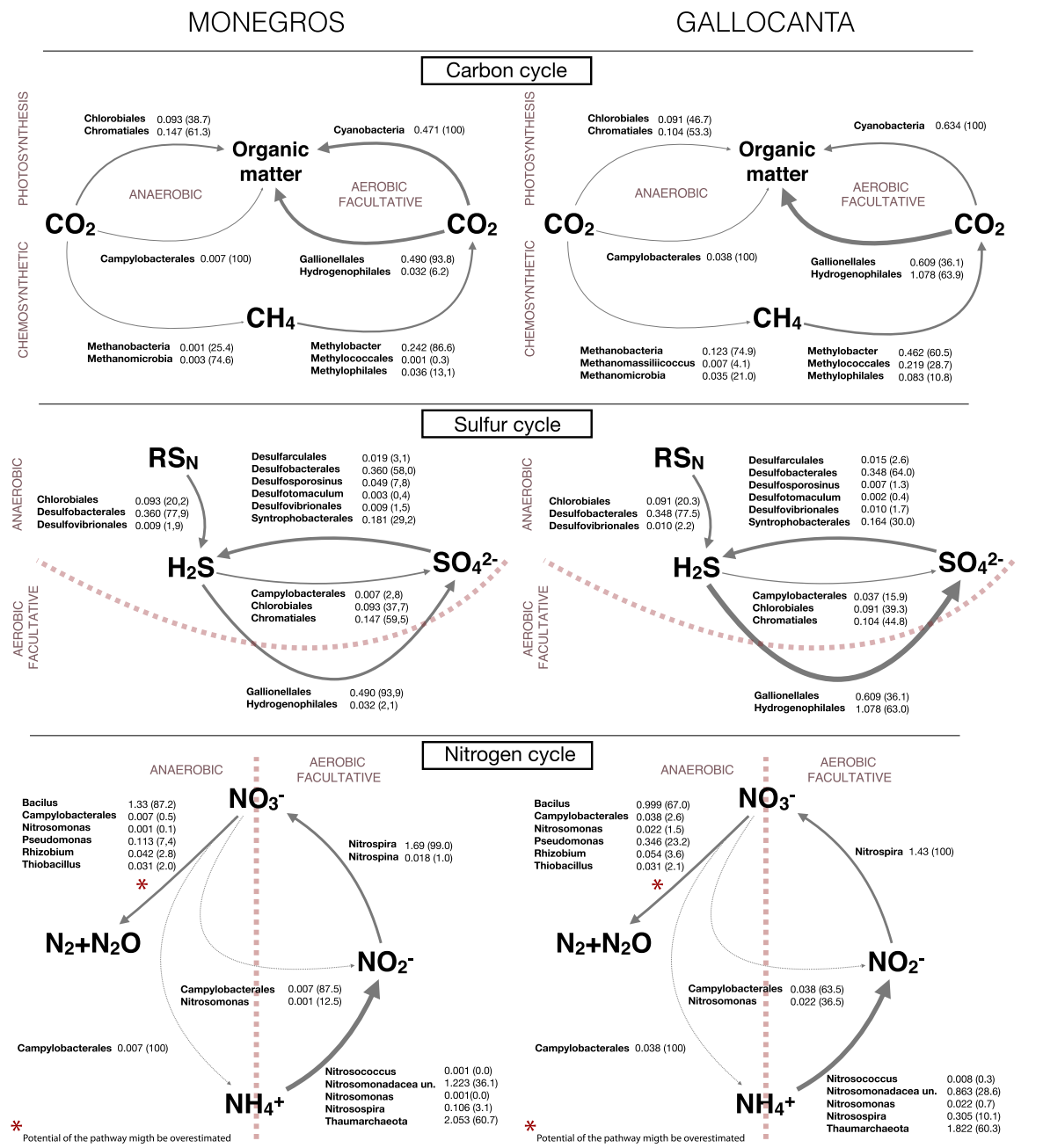


Fig. 5. Diagram representation for Carbon, Nitrogen and Sulfur biogeochemical cycling. Taxonomic groups potentially involved on the different pathways are quantified in terms of mean abundance vs. total sequences (%). A second value is given, inside parentheses, as the mean abundance inside each pathway (%). Arrows size shows the potential of each biogeochemical step according to relative abundances.

inventories and maps. The microbial novelty captured in the present study highlights important gaps both in culturing and in sequencing efforts for dry and saline soils, and more detailed investigations are needed. Hypersaline and gypsum rich drylands require the development of active conservation strategies to preserve microbial biodiversity (Casamayor, 2017), to unveil biotechnological applications (Oren,

2002), and to potentiate exobiology studies (Azua-Bustos et al., 2012) in areas often considered of reduced environmental interest. Under a geochemical perspective, the investigated wetland soils and their associated aquifers have a high level of mineralization mainly due to the influence of the substrata, especially in Monegros where lutites and gypsum materials are abundant. The high solubility of gypsum

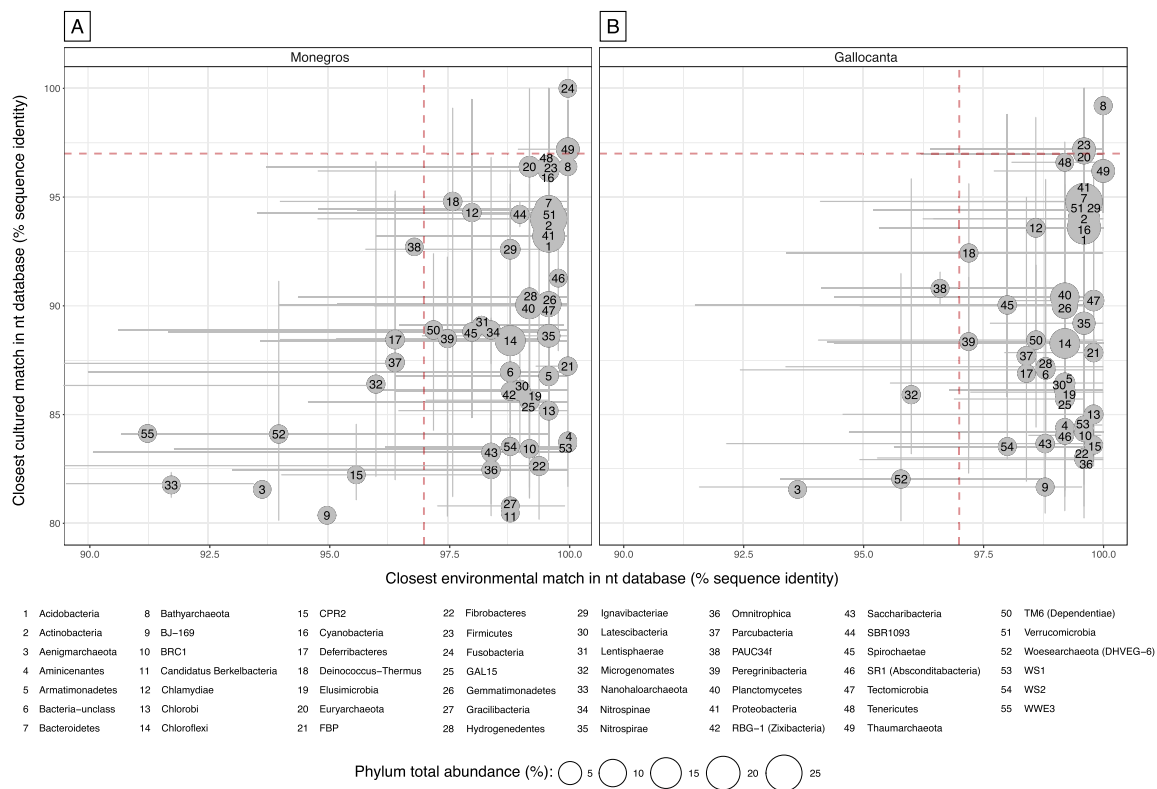


Fig. 6. Novelty level according to the Closest Environmental Match (CEM) and Closest Cultured Match (CCM) in GenBank, for Monegros (A) and Gallocanta (B). Error bars correspond to the 95th percentile. Red-dotted line marks the “species-level” 97% identity cut-off. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(2 g L⁻¹) produces the saturation in both calcium and sulfate ions in the soil solution, even for modest content of gypsum in soil. These ions are not harmful for microorganisms, as shown by the fact that gypsum crystals can host microbes (Vítek et al., 2016) in extremely harsh environments (Crits-Christoph et al., 2016). This observation is in agreement with the fact that we did not detect all major influence of gypsum content in the soil microbial composition or in the ecological diversity indexes. Most soils surveyed in the present investigation contained gypsum plus high concentrations of more soluble salts, with sodium, magnesium and chloride as the most toxic ions that also produce high osmotic stress to microorganisms, and probably the differences in concentration of these ions in the soil determined the prosperity of the different microbial communities (Pajares et al., 2016). The effects of highly soluble salts are supposed to supersede all other life-stressing factors, but the antagonistic effect of Ca²⁺ and Mg²⁺ against monovalent ions can be surmised. Moreover, wetlands soils are vulnerable to diffuse pollution from the surrounding agricultural fields, especially in the study area, with risk of failure to comply with the environmental objectives of the Water Framework Directive (Carceller-Layel et al., 2007). Diffuse pollution and eutrophication may also threaten the idiosyncratic and most novel microbial populations of Monegros and Gallocanta. Nutrients coming from massive agricultural fertilization and intensive pig farming move downslope and accumulate in water and soils. Because of the chronic scarcity of water in arid lands, dilution potential is low and such inputs have an ecological impact much greater than in wet regions. This transfer of nutrients towards the soil wetlands can be favoured by the predominantly sandy and loamy texture (Castañeda et al., 2015), especially in the case of the gypsum-rich soils (Castañeda et al., 2015; Moret-Fernández et al., 2013).

The functional roles obtained from the potential biogeochemical metabolic pathways, with larger abundance and diversity of microorganisms involved in N than in S cycling, and small differences in the main metabolic pathways for Monegros and Gallocanta agreed with the intermittent flooding conditions prevailing in these hypersaline and gypsum rich soils. In both regions, predominant N pathways were related to nitrification, nitrite oxidation, and nitrogen assimilation. This is probably related to high presence of reactive nitrogen compounds in soils due to their use as fertilizers that could be exacerbated in the future. The potential for denitrification of the soil microbial assemblages should be however carefully interpreted, because the relationship with taxonomy is not so clear like with other groups. It remains to be tested whether or not most members of the phyla associated with denitrification are likely to be denitrifiers. We observed the potential for an effective biogeochemical switching from aerobic to anaerobic conditions and vice versa for the carbon and sulfur cycles also probably related with the intermittent flooding conditions prevailing in these areas. Potential for methanogenesis (anaerobic archaeal methanogens) and methane oxidation (methylobacteria), and for carbon fixation by oxygenic (Cyanobacteria), anoxygenic (purple and green sulfur bacteria) photosynthesis and chemolithotrophs (autotrophic colorless sulfur bacteria and iron oxidizers) was widely observed. In most cases, the potential for carbon cycling was coupled to both anaerobically nitrate respiration (autotrophic and heterotrophic denitrifiers), sulfur oxidation (colorless sulfur bacteria) and anaerobic sulfate respiration (sulfate reducers) highlighting the high versatility of these microbial populations.

Overall, saline and arid soils contained a metabolically rich microbial community, and substantial microbial novelty with new groups of unknown microorganisms that should be addressed by microbiological culturing methods for detailed ecophysiological characterization. Frequent fluctuations between flooding and desiccation periods probably provided the heterogeneous environmental conditions that supported a large repertoire of aerobic and anaerobic metabolisms that can fuel active biogeochemical cycling despite of the high fluctuating scenarios. The importance of microbes in the biological structure and functioning of unusual arid zones such as those studied here, should be more widely

considered in conservation and management plans as a fundamental component of the natural heritage threatened by extensive changes in land use.

Declarations of interest

None.

Acknowledgments

This research was granted by Spanish projects BRIDGES CGL2015-69043-P to EOC, and BASIL (PCIN-2014-106) funded by the Ministry of Economy, Industry and Competitiveness (MINECO). MMS holds a FPI predoctoral fellowship from BRIDGES-MINECO. We acknowledge J.A. Conesa for the revision of plant species composition in Monegros and Gallocanta.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scitotenv.2018.09.050>.

References

- Andam, C.P., Doroghazi, J.R., Campbell, A.N., Kelly, P.J., Choudoir, M.J., Buckley, D.H., 2016. A latitudinal diversity gradient in terrestrial bacteria of the genus *Streptomyces*. *MBio* 7, e02200–e02215.
- Artieda, O., Herrero, J., Drohan, P.J., 2006. Refinement of the differential water loss method for gypsum determination in soils. *Soil Sci. Soc. Am. J.* 70, 1932–1935. <https://doi.org/10.2136/sssaj2006.0043N>.
- Azua-Bustos, A., Urrejola, C., Vicuña, R., 2012. Life at the dry edge: microorganisms of the Atacama Desert. *FEBS Lett.* 586 (18), 2939–2945. <https://doi.org/10.1016/j.febslet.2012.07.025>.
- Barberán, A., Casamayor, E.O., 2014. A phylogenetic perspective on species diversity, β -diversity and biogeography of the microbial world. *Mol. Ecol.* 23, 5868–5876.
- Barberán, A., Casamayor, E.O., Fierer, N., 2014. The microbial contribution to macroecology. *Front. Microbiol.* 5, 203.
- Bardgett, R.D., van der Putten, W.H., 2014. Belowground biodiversity and ecosystem functioning. *Nature* 515, 505–511.
- Belser, L.W., 1979. Population ecology of nitrifying bacteria. *Annu. Rev. Microbiol.* 33 (1), 309–333. <https://doi.org/10.1146/annurev.mi.33.100179.001521>.
- Bolhuis, H., Cretoiu, M.S., Stal, L.J., 2014. Molecular ecology of microbial mats. *FEMS Microb. Ecol.* 90, 335–350.
- Borrel, G., Parisot, N., Harris, H.M., Peyretailade, E., Gaci, N., Tottey, W., Brugère, J.F., 2014. Comparative genomics highlights the unique biology of Methanomassiliococcales, a Thermoplasmatales-related seventh order of methanogenic archaea that encodes pyrrolysine. *BMC Genomics* 15 (1), 679. <https://doi.org/10.1186/1471-2164-15-679>.
- Braun-Blanquet, J., Bolòs, O., 1958. Les groupements végétaux du bassin moyen de l'Ebre et leur dynamisme. *Annales Est. Exp. Aula Dei* 5 (1–4), 1–266.
- Caliz, J., Montes-Borrego, M., Triadó-Margarit, X., Metsis, M., Landa, B.B., Casamayor, E.O., 2015. Influence of edaphic, climatic, and agronomic factors on the composition and abundance of nitrifying microorganisms in the rhizosphere of commercial olive crops. *PLoS One* 10 (5). <https://doi.org/10.1371/journal.pone.0125787>.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336.
- Carceller-Layel, T., Costa-Alandí, C., Coloma-López, P., García-Vera, M.A., San Román-Saldaña, J., 2007. Groundwater in the central sector of the Ebro Basin. *Int. J. Water Res. Dev.* 23 (1), 165–187. <https://doi.org/10.1080/07900620601185649>.
- Carlson, C.A., Ingraham, J.L., 1983. Comparison of denitrification by *Pseudomonas stutzeri*, *Pseudomonas aeruginosa*, and *Paracoccus denitrificans*. *Appl. Environ. Microbiol.* 45, 1247–1253.
- Casamayor, E.O., 2017. Towards a microbial conservation perspective in high-mountain lakes. In: Catalan, J., Ninot, J.M., Aniz, M.M. (Eds.), *Challenges for High Mountain Conservation in a Changing World*. Book series “Advances in Global Change Research” Vol. 62. Springer, pp. 157–180.
- Casamayor, E.O., Triadó-Margarit, X., Castañeda, C., 2013. Microbial biodiversity in saline shallow lakes of the Monegros Desert, Spain. *FEMS Microbiol. Ecol.* 85 (3), 503–518. <https://doi.org/10.1111/1574-6941.12139>.
- Casby-Horton, S., Herrero, J., Rolong, N.A., 2015. Gypsum soils – their morphology, classification, function, and landscapes. *Adv. Agron.* 130, 231–290. <https://doi.org/10.1016/b.s.agron.2014.10.002>.
- Castañeda, C., Gracia, F.J., Luna, E., Rodríguez-Ochoa, R., 2015. Edaphic and geomorphic evidences of water level fluctuations in Gallocanta Lake, NE Spain. *Geoderma* 239–240, 265–279. <https://doi.org/10.1016/j.geoderma.2014.11.005>.

- Castañeda, C., Luna, E., Rabenhorst, M., 2017. Reducing conditions of soil of Gallocanta Lake, NE Spain. *Eur. J. Soil Sci.* 68, 249–258. <https://doi.org/10.1111/ejss.12407>.
- Conesa, J.A., Castañeda, C., Pedrol, J., 2011. Las saladas de Monegros y su entorno. *Hábitats y paisaje vegetal* (in Spanish). Consejo de Protección de la Naturaleza de Aragón. Zaragoza. (540 pp.). 978-84-89862-76-0.
- Connon, S.A., Lester, E.D., Shafaat, H.S., Obenhuber, D.C., Ponce, A., 2007. Bacterial diversity in hyperarid Atacama desert soils. *J. Geophys. Res. Biogeosci.* 112 (4), 1–9. <https://doi.org/10.1029/2006JG003311>.
- Cris-Christoph, A., Robinson, C.K., Bing, M., Jacques, R., Wierchos, J., Ascaso, C., Octavio, O., Souza-Egipsy, V., Casero, M.C., Diruggiero, J., 2016. Phylogenetic and functional substrate specificity for endolithic microbial communities in hyper-arid environments. *Front. Microbiol.* 7, 301. <https://www.frontiersin.org/article/10.3389/fmicb.2016.00301>.
- Davidson, N.C., 2014. How much wetland has the world lost? Long-term and recent trends in global wetland area. *Mar. Freshw. Res.* 65, 934–941. <https://doi.org/10.1071/MF14173>.
- Dregne, H.E., 1976. *Soils of the Arid Regions*. Elsevier, Amsterdam.
- Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10 (10), 996–998 Retrieved from. <https://doi.org/10.1038/nmeth.2604>.
- Edgar, R.C., Flyvbjerg, H., 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* 31 (21), 3476–3482 Retrieved from. <https://doi.org/10.1093/bioinformatics/btv401>.
- Faci, J.M., Martínez-Cob, A., 1991. The Reference Evapotranspiration Estimation in Aragón (in Spanish). Diputación General de Aragón, Zaragoza, Spain Available at: <http://digital.csic.es/handle/10261/73507> (115 pp.).
- Fernández-Guerra, A., Casamayor, E.O., 2012. Habitat-associated phylogenetic community patterns of microbial ammonia oxidizers. *PLoS One* 7, e47330.
- Fierer, N., Jackson, R.B., 2006. The diversity and biogeography of soil bacterial communities. *Proc. Natl. Acad. Sci. U. S. A.* 103, 626.
- García-Vera, M.A., Martínez-Cob, A., 2004. Evolución del contenido de humedad y de la tasa de evaporación en humedales: aplicación a la Laguna de Gallocanta. 2004-PH-14-I, Zaragoza, Spain (In Spanish). Available at: <http://digital.csic.es/bitstream/10261/2976/1/InformeGallocanta2004-DEF.pdf>.
- Gasol, J.M., Casamayor, E.O., Joint, I., Schauer, M., Massana, R., Pedrós-Alió, C., 2004. Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat. Microb. Ecol.* 34, 193–206.
- Gómez, D., Monserrat, G., Ferrer, J., 1983. Aportación al estudio de la flora y vegetación en la cuenca endorreica de Gallocanta. In: Aranzadi, E., Guiral, J.J. (Eds.), *Estudio de la Biocenosis de la Laguna de Gallocanta y su cuenca*. MOPU, Madrid, España (982 pp.).
- Gower, J.C., 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* 53 (3–4), 325–338 Retrieved from. <https://doi.org/10.1093/biomet/53.3.4.325>.
- Gower, J.C., 1971. A general coefficient of similarity and some of its properties. *Biometrics* 27 (4), 857–871. <https://doi.org/10.2307/2528823>.
- Greedy, I.F., Lane, C.R., Serran, J.N., Alexander, L.C., McLaughlin, D.L., Raanan-Kiperwas, H., Rains, M.C., Rains, K.C., Smith, L., 2017. Enhancing protection for vulnerable waters. *Nat. Geosci.* 10 (11), 809–815. <https://doi.org/10.1038/ngeo3041>.
- Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29 (8), 1072–1075 Retrieved from. <https://doi.org/10.1093/bioinformatics/btt086>.
- Iguchi, H., Yurimoto, H., Sakai, Y., 2015. Interactions of methylotrophs with plants and other heterotrophic bacteria. *Microorganisms* <https://doi.org/10.3390/microorganisms3020137>.
- Kingsford, R.T., Basset, A., Jackson, L., 2016. Wetlands: conservation's poor cousins. *Aquat. Conserv. Mar. Freshwat. Ecosyst.* 26, 892–916.
- Llorens-Marès, T., Yooseph, S., Goll, J., Hoffman, J., Vila-Costa, M., Borrego, C.M., Dupont, C.L., Casamayor, E.O., 2015. Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. *ISME J.* 9, 1648–1661.
- Lücker, S., Nowka, B., Rattei, T., Spieck, E., Daims, H., 2013. The genome of *Nitrospina gracilis* illuminates the metabolism and evolution of the major marine nitrite oxidizer. *Front. Microbiol.* 4 (FEB). <https://doi.org/10.3389/fmicb.2013.00027>.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Schleifer, K.-H., 2004. ARB: a software environment for sequence data. *Nucleic Acids Res.* 32 (4), 1363–1371. <https://doi.org/10.1093/nar/gkh293>.
- Maestre, F.T., Delgado-Baquerizo, M., Jeffries, T.C., Eldridge, D.J., Ochoa, V., Gozalo, B., Quero, J.L., García-Gómez, M., Gallardo, A., Ulrich, W., Bowker, M.A., Arredondo, T., Barraza-Zepeda, C., Bran, D., Florentino, A., Gaitán, J., Gutiérrez, J.R., Huber-Sannwald, E., Jankju, M., Mau, R.L., Miriti, M., Naseri, K., Ospina, A., Stavi, I., Wang, D., Woods, N.N., Yuan, X., Zaady, E., Singh, B.K., 2015. Increasing aridity reduces soil microbial diversity and abundance in global drylands. *Proc. Natl. Acad. Sci. U. S. A.* 112, 15684–15689.
- Moret-Fernández, D., Castañeda, C., Paracuellos, E., Jiménez, S., Herrero, J., 2013. Hydrophysical characterization of contrasting soils in a semiarid zone of the Ebro river valley (NE Spain). *J. Hydrol.* 486, 403–411. <https://doi.org/10.1016/j.jhydrol.2013.02.009>.
- Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Henry, M., Stevens, H., 2016. *vegan Community Ecology Package* version 2.4-1 September 2016.
- Oren, A., 1999. Bioenergetic aspects of halophilism. *Microbiol. Mol. Biol. Rev.* 63, 334–348.
- Oren, A., 2002. Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. *J. Ind. Microbiol. Biotechnol.* 28, 56–63.
- Pajares, S., Escalante, A.E., Noguez, A.M., García-Oliva, F., Martínez-Piedragil, C., Cram, S.S., Eguarte, L.E., Souza, V., 2016. Spatial heterogeneity of physicochemical properties explains differences in microbial composition in arid soils from Cuatro Ciénegas, México. *PeerJ* 4, e2459. <https://doi.org/10.7717/peerj.2459>.
- Palacio, S., Azorín, J., Montserrat-Martí, G., Ferrio, J.P., 2014. The crystallization water of gypsum rocks is a relevant water source for plants. *Nat. Commun.* <https://doi.org/10.1038/ncomms5660>.
- Pruesse, E., Peplies, J., Glöckner, F.O., 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28 (14), 1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>.
- Ramsar Convention Secretariat (Ed.), 2010. *Ramsar Handbooks for the Wise Use of Wetlands*, fourth ed. Vol. 1. Ramsar Convention Secretariat, Gland (Switzerland) (56 pp.).
- SBSTTA, 1999. *Biological Diversity of Drylands, Arid, Semi-arid, Savannah, Grassland and Mediterranean Ecosystems*. Draft Recommendations to COP5. Montreal, Canada (16 pp.).
- Tardy-Jacquenod, C., Magot, M., Patel, B.K., Matheron, R., Caumette, P., 1998. *Desulfotomaculum halophilum* sp. nov., a halophilic sulfate-reducing bacterium isolated from oil production facilities. *Int. J. Syst. Bacteriol.* 48 (2), 333–338. <https://doi.org/10.1099/00207713-48-2-333>.
- Torrentó, C., Cama, J., Urmeneta, J., Otero, N., Soler, A., 2010. Denitrification of groundwater with pyrite and *Thiobacillus denitrificans*. *Chem. Geol.* 278 (1–2), 80–91. <https://doi.org/10.1016/j.chemgeo.2010.09.003>.
- Torsvik, V., Goksoyr, J., Daas, F.L., 1990. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* 56, 782–787.
- Verbaendert, I., Boon, N., De Vos, P., Heylen, K., 2011. Denitrification is a common feature among members of the genus *Bacillus*. *Syst. Appl. Microbiol.* 34 (5), 385–391. <https://doi.org/10.1016/j.syapm.2011.02.003>.
- Vigneron, A., Bishop, A., Alsop, E.B., Hull, K., Rhodes, I., Hendricks, R., Head, I.M., Tsismetis, N., 2017. Microbial and isotopic evidence for methane cycling in hydrocarbon-containing groundwater from the Pennsylvania region. *Front. Microbiol.* 8. <https://doi.org/10.3389/fmicb.2017.00593>.
- Vitek, P., Ascaso, C., Artieda, O., Wierchos, J., 2016. Raman imaging in geomicrobiology: endolithic phototrophic microorganisms in gypsum from the extreme sun irradiation area in the Atacama Desert. *Anal. Bioanal. Chem.* 408 (15), 4083–4092. <https://doi.org/10.1007/s00216-016-9497-9>.
- Waiser, M.J., Robarts, R.D., 2009. Saline Inland Waters. In: Likens, Gene E. (Ed.), *Encyclopedia of Inland Waters*. Vol. 2. Elsevier, Oxford, pp. 634–644.
- Williams, W.D., 2002. Environmental threats to salt lakes and the likely status of inland saline ecosystems in 2025. *Environ. Conserv.* 29, 154–167. <https://doi.org/10.1017/S0376892902000103>.
- Yao, Q., Peng, D.-C., 2017. Nitrite oxidizing bacteria (NOB) dominating in nitrifying community in full-scale biological nutrient removal wastewater treatment plants. *AMB Express* 7 (1), 25. <https://doi.org/10.1186/s13568-017-0328-y>.

