universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Amoebae as vehicles of bacteria“

verfasst von / submitted by

Mag. rer. nat. Han-Fei Tsao

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 437
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Biologie
Betreut von / Supervisor:	Univ.-Prof. Dr. Matthias Horn

Acknowledgement

I would like to thank my supervisor Matthias Horn for giving me the opportunity to do this PhD and his mentoring not only during my time in DOME but also well beyond it. My presence in the group always felt welcomed and appreciated. I'm glad and thankful that our exchanges go way beyond scientific topics.

Special thanks to Stephan, my former minion, with whom I've shared countless weekends in the lab and many hours in the world of Sanctuary. Our shared interest in anime/nerd stuff is solely surpassed by my dear geek-friend Michael, whose knowledge about all stuff nerd/geek/NBA, and politics related is only rivaled by his FISH skills. To Christian, my man-of-honor, whose friendliness and people skills are something I'm looking up to. To Daryl for showing me the 'murican way, BBQ, whiskies, craft beers, jackets, and being a good friend. To Bela for being a valuable member of the original diploma student club, his parties, valuable advice, and a great upper-lab office-buddy. To Nadi, my second protégé; even though you've become a great admin, you'll always be the "tiny-dragon" to me. To Vincent, the time we shared was short, but always lighthearted and plenty of fun. To Patrick, the physical embodiment of friendliness and empathy; I was and am still fond of all our great conversations. To Assi, the real "boss" behind the curtains, and her calm and up-lifting encouragements. To Paul, the maltese falcon, with whom I've shared much laughter over NSFW jokes. To Geli, the 3rd generation of "Tsao Group", for her persistence and helpfulness. To Lena for our time together as FISH course tutors. To Orest for giving me the incentives to reexamine my own priorities.

My thanks to all the DOMiEs I had the honor to spend numerous hours with, whether inside or outside the lab. I'll not only miss our scientific discussions but also our cheerful banters. I was always very fortunate to be surrounded by nice people and you guys are no exception here.

But most importantly, I want to thank my beloved wife Hsin-Chieh for being on my side already long before this marathon began. You not only supported me since day one but also brought two lovely cheerleaders on-board, Viola and Aurelia, who are the apple of my eyes.

Many thanks to my parents for their support and for allowing me to experience a whole new culture.

If there is one part I dread writing, it's this part, the acknowledgment. With no data to guide the direction of my writing and being free in what I could write, the task has become more ginormous than I had imagined. Too short and it feels insincere; too long and it becomes a list of names, which again, lessens somehow the perceived appreciation. Too factual, and one might fulfill the cliche that scientists are just a bunch of robots with no soul; too emotional, and this part could become a sentimental mess about the hardships one has endured during the PhD time. But if I'm really honest with myself, I'm just afraid of the possibility, that if I take a hard look back at my time in DOME, it might stare right back at me and I might not like what I see.

As I've left the place a couple of years now, with some temporal and spatial distances in between, I can finally reevaluate and examine my time there more objectively. For people outside of DOME, the department is usually associated with its excellent scientific outputs. While there is no doubt that science has been ever-present in my everyday life at DOME, it pales in the face of my more vivid memories, which are predominantly associated with the personal connections I've made with each and everyone there. So last but not least, a big thank you to all the people on this incredible journey, who have shaped and helped me to become the person I am today.

Table of contents

Abstract

Free-living amoebae (FLA) are ubiquitous, unicellular eukaryotes living in vastly different habitats ranging from aquatic to terrestrial environments. Not only do they reside in those environments, but they also actively shape their surrounding microbial communities by predating other microbes. FLA's interaction with their primary food source, the bacteria, is manifold and not only limited to a simplistic predator-prey relationship. Many bacteria have evolved to escape the digestion process after being taken up by amoebae phagocytosis and reside within the amoeba cell, where they can thrive. Not only is the intracellular space a perfect niche full of energy sources, but it also protects these 'inhabitants' from harsh exogenous environmental conditions. Some of those bacteria form stable and long-term relationships with the amoeba, whereas others are disease-causing pathogens, occupying the host only temporarily. It is necessary to understand both interactions, as the mechanisms needed to survive in the intracellular space are often similar and the knowledge transferable. This work sought to use an array of methods, ranging from microscopy and cultivation-based approaches to state of the art electron cryotomography, genomics and bioinformatics to study the many aspects of amoeba-microbe interaction. First, I focused on the diversity of both bacteria and protists in cooling tower waters to better understand their underlying co-occurrence and putative interactions. As FLAs can harbor and protect pathogens, the generated data may contribute to shaping future public-health related decisions on the architectural design of the monitored aquatic systems. The highly dynamic microbial diversities of the investigated man-made water systems resemble natural freshwater systems. However, the basin environment selects for biofilm-forming and -associated taxa, aiding colonization of opportunistic pathogens. Co-occurrence analysis further recaptured a set of known amoeba-microbe interactions and predicted other undiscovered associations and synergies. The importance of investigating this system was further highlighted by the discovery and description of the first bacterial endosymbiont of the testate amoeba *Cochliopodium minus*. The endosymbiont was discovered inside vacuoles within the host's cytosol. Phylogenetic analysis identified this novel symbiont was only moderately related to known amoeba symbionts, and belong to a monophyletic group within the family Coxiellaceae. Besides this novel endosymbiont, we have also examined in more detail a group of *Acanthamoeba* endosymbionts , *Procabacter* species that belong to a sister-clade of the

families Neisseriaceae and Chromobacteriaceae, both including human pathogens but no known amoeba symbionts. Infection experiments unveiled a prolonged infection cycle turnover time with no real fitness loss for the host. Genome analysis revealed only moderately reduced genomes, as this group of amoeba endosymbionts retains a remarkable amount of metabolic pathways and *denovo* biosynthetic capabilities. We also discovered the presence of other genomic features of host-associated organisms such as secretion systems and effector proteins, as well as putative horizontally transferred genes with importance for survival in the intracellular niche. We speculate that this lineage resembles endosymbionts that have only recently adopted the intracellular lifestyle, as we could observe dividing bacterial cells in a complex medium in a host-free cultivation experiment. At last, we described and showed the function of a novel type six secretion system, a molecular apparatus essential for host-cell interaction, expressed by another *Acanthamoeba* endosymbiont, *Amoebophilus asiaticus*. By using cryo-focused ion beam milling, combined with electron cryotomography, and functional assays, we showed that the system facilitates phagosome escape, an important step in establishing an intracellular lifestyle. This apparatus offers an efficient mechanism for intracellular survival, and evolutionary sequence analysis suggested that not only did the T6SS evolve independently multiple times, it is also more widespread than previously suggested. Overall, this thesis contributes to our current knowledge about the diversity of protist-associated microbes in man-made water systems, the adaptation of bacteria to the intracellular lifestyle, and aspects of the underlying molecular mechanisms.

Zusammenfassung

Freilebende Amöben (FLA) sind allgegenwärtige einzellige Eukaryoten, die in den unterschiedlichsten Lebensräumen vorkommen, die von aquatischen bis zu terrestrischen Umgebungen reichen. Als ein Teil dieser Umgebungen, beeinflussen die Amöben durch Jagd auf andere Mikroben auch aktiv die umliegenden mikrobiellen Gemeinschaften. Aber die Interaktion zwischen Amöben und den Bakterien ist nicht nur auf eine simple Jäger-Beute Beziehung beschränkt sondern sehr vielfältig. Denn manche Bakterien haben Strategien entwickelt, um dem Verdauungsprozess nach der Aufnahme durch Amöben zu entgehen. Diese Bakterien können dann im Zytosol des Wirts gedeihen und sich dort vermehren. Der Raum innerhalb der Zelle bietet nicht nur Energiequellen für die intrazelluläre Bakterien, sondern dient auch als Rückzugsort um diese vor Umwelteinflüssen zu schützen. Einige Bakterienarten haben sich entwickelt um eine langfristige Beziehungen mit der Wirtsamöbe eingehen zu können, während andere Spezien, die meistens pathogene Eigenschaft besitzen, sich nur temporär im Wirt aufhalten. Es ist wichtig, beide Arten der Interaktionen mit dem Wirt zu verstehen. Denn die molekularen Grundlagen und Merkmale für die Etablierung eines intrazellulären Lebensstils, sind sich oft ähnlich und die gewonnenen Erkenntnisse für beide System anwendbar. Diese Arbeit befasst sich mit unterschiedlichsten Aspekten einer Amöbe-Symbiont Beziehung, die sowohl Umweltdiversitätsstudien als auch molekulare Mechanismen einer Infektion abdecken. Wir berichten zunächst über die Ergebnisse einer mikrobiellen Diversitätsstudie von Kühlturmwassern, die wir über einen Jahr hinweg beprobt haben. Die Vielfalt und Dynamik der Bakterien- und Protistengemeinschaft wurde durch die Amplikonsequenzierung untersucht. Wir fanden heraus, dass die mikrobielle Diversität der untersuchten Wassersysteme hoch dynamisch ist. Diese ähneln zwar der Diversität von natürlichen Süßgewässern, selektieren jedoch auf biofilmbildende und -assoziierte Spezien. Da opportunistische Krankheitserreger verstärkt in Biofilmen vorkommen, ist es empfehlenswert bei der Planung von zukünftigen Kühlturmen auf das Minimieren von Biofilmbildung zu achten. Zusätzlich berichten wir über den ersten bakteriellen Endosymbionten der Schalenamöben (testate amoeba; im Gegensatz zur Nacktamöbe) Cochliopodium minus. Der Endosymbiont lebte in Vakuolen im Zytosol des Wirts. Phylogenetische Analyse ergaben, dass dieser neuartige Endosymbiont zu einer monophyletischen Gruppe innerhalb der Familie der Coxiellaceae gehört. Die Mitglieder

dieser neuartigen Gruppe weisen nur mäßige Verwandtschaft mit bekannten Amöben-Symbionten auf. Zusätzlich zur Endosymbiont von Schalenamöben untersuchten wir auch einen Endosymbionten von einer Akanthamöbe (Nacktamöbe). Dieser Endosymbiont gehört zur Schwesterklade der Familien Neisseriaceae und Chromobacteriaceae. Beide Familien beinhalten zwar humanpathogene Bakterien, aber keinen bekannten Amöbensymbionten. Die Genomanalyse zeigte, dass der Endosymbiont nicht nur ein mäßig reduziertes Genom besitzt sondern auch das Potenzial hat verschiedene Stoffwechselwege und De-novo-Biosynthese zu nutzen. All diese Merkmale sind untypisch für Endosymbionten. Wir entdeckten bei der Genomanalyse auch genomische Merkmale, wie Sekretionssystemen und Effektorproteinen sowie mutmaßlich horizontal übertragenen Gene, die allesamt für das intrazelluläre Überleben von Bedeutung sind. Wir spekulieren, dass evolutionär gesehen diese Gruppe an Endosymbionten erst kürzlich den intrazellulären Lebensstil übernommen hat, da wir im Rahmen einer wirt-freien Kultivierungsexperiment, teilende Bakterienzellen in einem komplexen Medium beobachtet haben. Im letzten Kapitel beschrieben und demonstrierten wir anhand einer Amöben-Endosymbiont, Amoebophilus asiaticus, die Funktion eines neuartigen Typ 6-Sekretionssystems (T6SS). Durch die Verwendung von kryofokussiertem Ionenstrahlmahlen (cryo-focused ion beam milling) in Kombination mit Kryo-Elektron-Tomographie und Funktions-Assays konnte gezeigt werden, dass das System dem Symbiont hilft aus Phagosomen vom Wirt zu entkommen; eine wichtige Funktion für das Überleben im intrazellulären Raum. Zusätzlich legte die Sequenzanalyse nahe, dass sich das T6SS mehrfach unabhängig voneinander entwickelte und weiter verbreitet ist als zuvor angenommen. Insgesamt trägt diese Arbeit wesentlich zu unserem aktuellen Verständnis über die mikrobielle Vielfalt in künstlichen Wassersystemen, die Anpassung von Bakterien an den intrazellulären Raum und Teil der zugrunde liegenden molekularen Mechanismen bei.

Chapter I

Introduction

Introduction

Bacteria are omnipresent. Their traces can be found in the deep trenches at the bottom of the oceans, up to the peaking heights of the Himalayan summits. They are in the never thawing permafrost of Antarctica and on dust specks of Sahara. Their occupation of a wide range of habitats is also reflected in bacterial interactions with other species, which are manifold. In the late 19th century, the German scientist Heinrich Anton de Bary introduced a concept of an intimate, long-term association between different organisms and coined it "symbiosis" (de Bary, 1879). Symbiotic interactions involve all realms of life, but the most prevalent one is between bacteria and eukaryotes (Dale and Moran, 2006). Symbiosis impacts not only individual organisms but also whole ecosystems (McFall-Ngai et al., 2013), affecting development, reproduction, use of distinct nutrients, and defense against harmful bacteria (Moran et al., 2006). In a bacteria-eukaryote symbiotic system, we usually refer to the bacterium as the symbiont and the eukaryote as the host. A more detailed distinction can be made based on the relative location of the interaction partners to each other. Ectosymbionts reside outside on the host's surface, whereas endosymbionts occupy the inside of the host-cell. Some interactions are tightly interwoven and the partners cannot be separated for a prolonged time period without one or both partners dying. It is widely accepted within the scientific communities that it is the symbiotic interactions themselves that lead to the emergence of complex life and eventually the origin of eukaryotes (Koonin, 2010; Margulis, 2004; Martijn and Ettema, 2013).

Free-living amoebae

Free-living amoebae (FLA) are ubiquitous unicellular eukaryotes, occupying a broad range of habitats. They play an important ecological role, because of their ability to influence and shape microbial communities by not only predating other microorganisms, but also harboring bacteria as potential hosts. An example of potential hosts are the free-living amoebae (FLA). It is important to note that the term FLA does not refer to their taxonomy but constitutes a summary term for phylogenetically diverse and heterogeneous unicellular organisms, e.g., *Acanthamoeba*, *Vermamoeba, Naegleria,* and *Balamuthia* (Walochnik & Aspöck, 2007). FLA are important bacterial predators as well as regulators of the microbial communities and populations, including fungi and other protozoa (Weekers et al., 1993). FLA can change between distinct forms, the stationary cyst, the amoeboid trophozoite form, and the

flagellated form for some FLA like the *Naegleria* (Dingle, 1966). The motile and metabolically active trophozoites graze via phagocytosis, a process where extracellular particles are first engulfed and then internalized. In contrast, the cyst stage exhibits elevated resistance towards exogenous harm, not unlike spores. At this stage, FLA can defy unfavorable conditions such as anaerobic conditions (Tomov et al., 1999), desiccation, chemical toxicities (Kilvington and Price, 1990), radiation (Aksozek et al. 2002), heat, and starvation. Encystation allows the amoeba to cope with highly dynamic conditions and thus inhabit various environments, not only in freshwater and soil environments but also in air samples and man-made environments (Fouque et al., 2012; Mergeryan, 1991; Miltner and Bermudez, 2000).

FLA as a reservoir for bacteria

In the 1970s, the first study about long-term symbiotic relationships between *Acanthamoeba* sp. and bacteria was published (Proca-Ciobanu et al., 1975). Today, we know that FLA like *Acanthamoeba* spp., *Vermamoeba vermiformis,* Thecamoebidae, and *Naegleria* spp. are all able to harbor bacteria (Horn et al., 2000, Michel et al., 1999; Michel et al., 2009; Walochnik et al., 2005). These bacteria can survive the uptake by phagocytosis, evade the downstream digestion in phagolysosomes, and exploit the host as a niche for replication (Hoffmann et al., 1997; Koehsler et al., 2007; Verani et al., 2016). These microbes then reside within the host cell, either in the cytosolic space, vacuoles, cyst walls or, in some rare cases, the nucleus (Steinert et al., 1998; Michel et al., 2009; Walochnik et al., 2005; Schulz et al., 2014). Living intracellularly grants the endosymbiont several benefits. The first advantage is nutrient availability, and the second is the protective nature of the intracellular niche. We differentiate between two categories of amoeba-associated bacteria. First, there are the facultative amoeba endosymbionts including bacterial pathogens such as *Legionella* and *Mycobacterium* spp., *Pseudomonas aeruginosa*, *Francisella tularensis*, *Coxiella burnetii*, and *Vibrio cholerae* (Abd et al., 2003; La Scola and Raoult, 2001; Sandström et al., 2010; Thomas and McDonnell, 2007; Thom et al., 1992; Van der Henst et al., 2016). These pathogens only reside temporarily within FLAs as they are able to survive outside of the host. In addition, their interactions are usually unsustainable over time as the facultative symbionts pursue conflicting fitness interests with their host. The second category are the obligate endosymbionts of FLAs. Those symbionts are reliant on their host to provide shelter and metabolites. Already back in the 1990s, a report has described that up to 25% of all *Acanthamoeba* isolates have a stable symbiotic association

with intracellular bacteria (Fritsche et al., 1993). Today known FLA-associated endosymbionts belong to diverse bacterial lineages: Alphaproteobacteria (Fritsche et al., 1999; Horn et al., 1999; Schrallhammer et al., 2013; Xuan et al., 2007), Betaproteobacteria (Heinz et al., 2007; Horn et al., 2002), Gammaproteobacteria (Horn et al., 1999; Schulz et al., 2015; Tsao et al., 2017), Bacteriodetes (Horn et al., 2001), Depedentiae (formerly known as TM6; Delafont et al., 2015; Pagnier et al., 2015) and the Chlamydiae (Collingro et al., 2005; Corsaro et al., 2010; Horn et al., 2004; Horn, 2008).

Implications for public-health

Encystation of FLA boosts their resistance toward hazardous environments, simultaneously sheltering all intracellular organisms residing inside. This has become a particularly notorious problem for effective disinfection of drinking water systems, as bacteria inside amoebae can survive disinfection procedures and quickly recolonize water systems (Lau and Ashbolt 2009; Marciano-Cabral et al., 2010). Another threat is the high concentration of potential pathogens within a single host cell, easily crossing the minimal infective dose for establishing a successful infection. Worse, *Legionella pneumophila,* the causative agent of Legionnaire's disease, either grown within amoebae or in contact with such, exhibit elevated pathogenicity towards macrophages and mice cells compared to agar-plate grown counterparts (Brieland et al., 1997; Cirillo et at., 1994; Cirillo et al., 1999). This constitutes a significant health issue, especially in public buildings, such as hospitals with risk groups such as immunocompromised or elderly people. Therefore, FLAs are considered vectors or "trojan horses" for bacterial pathogens for spreading diseases to humans (Barker and Brown, 1994). Further studies have implicated that the relationship goes deeper than mere sheltering. FLA's phagocytic trait, a function shared with the metazoan immune response, might have also driven the adaptation of bacteria towards intracellular survival in eukaryotic cells, effectively broadening the range of potential hosts (Albert-Weissenberger et al., 2007; Barker and Brown, 1994; Cosson & Soldati, 2008; Darby et al., 2007; Harb et al., 2000; Horn et al., 2004; Molmeret et al., 2005). All things considered, there is substantial evidence for a vital role of amoebae as vehicles for known and potentially emerging pathogens, especially in drinking water and cooling water systems. Yet, microbial communities and their interactions in those systems are likely more complex than realized and surveys on the overall microbial community composition and dynamics are rare (Llewellyn et al., 2017; Pereira et al., 2017; Wang et al., 2013).

References

Abd, H., Johansson, T., Golovliov, I., Sandström, G., and Forsman, M. (2003) Survival and growth of Francisella tularensis in Acanthamoeba castellanii. *Appl Environ Microbiol* **69**: 600–606.

Aksozek, A., McClellan, K., Howard, K., Niederkorn, J.Y., and Alizadeh, H. (2002) Resistance of Acanthamoeba castellanii Cysts to Physical, Chemical, and Radiological Conditions. *J Parasitol* **88**: 621–623.

Albert-Weissenberger, C., Cazalet, C., and Buchrieser, C. (2007) Legionella pneumophila - A human pathogen that co-evolved with fresh water protozoa. *Cell Mol Life Sci* **64**: 432–448.

Barker, J. and Brown, M.R.W. (1994) Trojan Horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. *Microbiology* **140**: 1253–1259.

Brieland, J.K., Fantone, J.C., Remick, D.G., LeGendre, M., McClain, M., and Engleberg, N.C. (1997) The role of Legionella pneumophila-infected Hartmannella vermiformis as an infectious particle in a murine model of Legionnaires' disease. *Infect Immun* **65**: 5330–5333.

Cirillo, J.D., Falkow, S., and Tompkins, L.S. (1994) Growth of Legionella pneumophila in Acanthamoeba castellanii enhances invasion. *Infect Immun* **62**: 3254–3261.

Cirillo, J.D., Cirillo, S.L.G., Yan, L., Bermudez, L.E., Falkow, S., and Tompkins, L.S. (1999) Intracellular growth in Acanthamoeba castellanii affects monocyte entry mechanisms and enhances virulence of Legionella pneumophila. *Infect Immun* **67**: 4427–4434.

Collingro, A., Toenshoff, E.R., Taylor, M.W., Fritsche, T.R., Wagner, M., and Horn, M. (2005) "Candidatus Protochlamydia amoebophila", an endosymbiont of Acanthamoeba spp. *Int J Syst Evol Microbiol* **55**: 1863–1866.

Corsaro, D., Michel, R., Walochnik, J., Müller, K.D., and Greub, G. (2010) Saccamoeba lacustris, sp. nov. (Amoebozoa: Lobosea: Hartmannellidae), a new lobose amoeba, parasitized by the novel chlamydia "Candidatus Metachlamydia lacustris" (Chlamydiae: Parachlamydiaceae). *Eur J Protistol* **46**: 86–95.

Cosson, P. and Soldati, T. (2008) Eat, kill or die: when amoeba meets bacteria. *Curr Opin Microbiol* **11**: 271–276.

Dale, C. and Moran, N.A. (2006) Molecular Interactions between Bacterial Symbionts and Their Hosts. *Cell* **126**: 453–465.

Darby, A.C., Cho, N.H., Fuxelius, H.H., Westberg, J., and Andersson, S.G.E. (2007) Intracellular pathogens go extreme: genome evolution in the Rickettsiales. *Trends Genet* **23**: 511–520.

de Bary, A. (1879) DIE ERSCHEINUNG DER SYMBIOSE, Verlag von Karl J. Trübner.

Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L., and Héchard, Y. (2015) Shedding light on microbial dark matter: A TM6 bacterium as natural endosymbiont of a free-living amoeba. *Environ Microbiol Rep* **7**: 970–978.

Dingle, A.D. and Fulton, C. (1966) DEVELOPMENT OF THE FLAGELLAR APPARATUS OF NAEGLERIA. *J Cell Biol* **31**: 43–54.

Fouque, E., Trouilhé, M.C., Thomas, V., Hartemann, P., Rodier, M.H., and Hécharda, Y. (2012) Cellular, biochemical, and molecular changes during encystment of free-living amoebae. *Eukaryot Cell* **11**: 382–387.

Fritsche, T.R., Gautom, R.K., Seyedirashti, S., Bergeron, D.L., and Lindquist, T.D. (1993) Occurrence of bacterial endosymbionts in Acanthamoeba spp. isolated from corneal and environmental specimens and contact lenses. *J Clin Microbiol* **31**: 1122–1126.

Fritsche, T.R., Horn, M., Seyedirashti, S., Gautom, R.K., Schleifer, K.H., and Wagner, M. (1999) In situ detection of novel bacterial endosymbionts of Acanthamoeba spp. phylogenetically related to members of the order Rickettsiales. *Appl Environ Microbiol* **65**: 206–12.

Harb, O.S., Gao, L.Y., and Kwaik, Y.A. (2000) From protozoa to mammalian cells: A new paradigm in the life cycle of intracellular bacterial pathogens. *Environ Microbiol* **2**: 251–265.

Heinz, E., Kolarov, I., Kästner, C., Toenshoff, E.R., Wagner, M., and Horn, M. (2007) An Acanthamoeba sp. containing two phylogenetically different bacterial endosymbionts. *Environ Microbiol* **9**: 1604–1609.

Hoffmann, R. (1997) Natural infection with microsporidian organisms (KW19) in Vannella spp. (Gymnamoebia) isolated from a domestic tap-water supply. *Parasitol Res* **84**: 164–166.

Horn, M., Harzenetter, M.D., Linner, T., Schmid, E.N., Müller, K.D., Michel, R., and Wagner, M. (2001) Members of the Cytophaga-Flavobacterium-Bacteroides phylum as intracellular bacteria of acanthamoebae: proposal of "Candidatus Amoebophilus asiaticus". *Environ Microbiol* **3**: 440–9.

Horn, M. (2008) Chlamydiae as Symbionts in Eukaryotes. *Annu Rev Microbiol* **62**: 113–131.

Horn, M., Fritsche, T.R., Gautom, R.K., Schleifer, K.H., and Wagner, M. (1999) Novel bacterial endosymbionts of Acanthamoeba spp. related to the Paramecium caudatum symbiont Caedibacter caryophilus. *Environ Microbiol* **1**: 357–367.

Horn, M. and Wagner, M. (2004) Bacterial endosymbionts of free-living amoebae. In *Journal of Eukaryotic Microbiology*. pp. 509–514.

Horn, M., Wagner, M., Müller, K.D., Schmid, E.N., Fritsche, T.R., Schleifer, K.H., and Michel, R. (2000) Neochlamydia hartmannellae gen. nov., sp. nov. (Parachlamydiaceae), an endoparasite of the amoeba Hartmannella vermiformis. *Microbiology* **146**: 1231–1239.

Kilvington, S. and Price, J. (1990) Survival of Legionella pneumophila within cysts of Acanthamoeba polyphaga following chlorine exposure. *J Appl Bacteriol* **68**: 519–525.

Koehsler, M., Walochnik, J., Michel, R., Lugauer, J., and Wylezich, C. (2007) Molecular identification and classification of Cochlonema euryblastum, a zoopagalean parasite of Thecamoeba quadrilineata. *Mycologia* **99**: 215–221.

Koonin, E. V. (2010) The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biol* **11:** 209

Lau, H.Y. and Ashbolt, N.J. (2009) The role of biofilms and protozoa in legionella pathogenesis: Implications for drinking water. *J Appl Microbiol* **2**: 368-378

Llewellyn, A.C., Lucas, C.E., Roberts, S.E., Brown, E.W., Nayak, B.S., Raphael, B.H., and Winchell, J.M. (2017) Distribution of Legionella and bacterial community composition among regionally diverse US cooling towers. *PLoS One* **12**: e0189937.

Marciano-Cabral, F., Jamerson, M., and Kaneshiro, E.S. (2010) Free-living amoebae, Legionella and Mycobacterium in tap water supplied by a municipal drinking water utility in the USA. *J Water Health* **8**: 71–82.

Margulis, L. (2004) Serial endosymbiotic theory (SET) and composite individuality. *Microbiol Today* **31**: 172–175.

Marrie, T.J., Raoult, D., La Scola, B., Birtles, R.J., De Carolis, E., Duperval, R., et al. (2001) Legionella-like and other amoebal pathogens as agents of community-acquired pneumonia. *Emerg Infect Dis* **7**: 1026–1029.

Martijn, J. and Ettema, T.J.G. (2013) From archaeon to eukaryote: The evolutionary dark ages of the eukaryotic cell. In *Biochemical Society Transactions*.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H. V., Domazet-Lošo, T., Douglas, A.E., et al. (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci U S A* **110**: 3229–3236.

Mergeryan, H. (1991) The prevalence of acanthamoeba in the human environment. *Rev Infect Dis* **13**: S390–S391.

Michel, R., Hauröded r, B., and Zöller, L. (2009) Isolation of the amoeba thecamoeba quadrilineata harbouring intranuclear spore forming endoparasites considered as fungus-like organisms. *Acta Protozool* **48**: 41–49.

Michel, R., Hauröder, B., Müller, K.D., and Zöller, L. (1999) An environmental Naegleria-strain, unable to form cysts - Turned out to harbour two different species of endocytobionts. *Endocytobiosis Cell Res* **13**: 115–118.

Miltner, E.C. and Bermudez, L.E. (2000) Mycobacterium avium grown in Acanthamoeba castellanii is protected from the effects of antimicrobials. *Antimicrob Agents Chemother* **44**: 1990–1994.

Molmeret, M., Horn, M., Wagner, M., Santic, M., Abu Kwaik, Y., Kwaik, Y.A., and Abu Kwaik, Y. (2005) Amoebae as Training Grounds for Intracellular Bacterial Pathogens. *Appl Environ Microbiol* **71**: 20–28.

Pagnier, I., Yutin, N., Croce, O., Makarova, K.S., Wolf, Y.I., Benamar, S., et al. (2015) Babela massiliensis, a representative of a widespread bacterial phylum with unusual adaptations to parasitism in amoebae. *Biol Direct* **10**: 13.

Pereira, R.P.A., Peplies, J., Höfle, M.G., and Brettar, I. (2017) Bacterial community dynamics in a cooling tower with emphasis on pathogenic bacteria and Legionella species using universal and genus-specific deep sequencing. *Water Res* **122**: 363–376.

Proca-Ciobanu, M., Lupascu, G., Petrovici, A., and Ionescu, M.D. (1975) Electron microscopic study of a pathogenic Acanthamoeba castellani strain: The presence of bacterial endosymbionts. *Int J Parasitol* **5**: 49–56.

Sandström, G., Saeed, A., and Abd, H. (2010) Acanthamoeba polyphaga is a possible host for Vibrio cholerae in aquatic environments. *Exp Parasitol* **126**: 65–8.

Schrallhammer, M., Ferrantini, F., Vannini, C., Galati, S., Schweikert, M., Görtz, H.D., et al. (2013) "Candidatus Megaira polyxenophila" gen. nov., sp. nov.: Considerations on Evolutionary History, Host Range and Shift of Early Divergent Rickettsiae. *PLoS One* **8**:.

Schulz, F., Lagkouvardos, I., Wascher, F., Aistleitner, K., Kostanjšek, R., and Horn, M. (2014) Life in an unusual intracellular niche: A bacterial symbiont infecting the nucleus of amoebae. *ISME J* **8**: 1634–1644.

Schulz, F., Tyml, T., Pizzetti, I., Dyková, I., Fazi, S., Kostka, M., and Horn, M. (2015) Marine amoebae with cytoplasmic and perinuclear symbionts deeply branching in the Gammaproteobacteria. *Sci Rep* **5**: 1–10.

Thom, S., Warhurst, D., and Drasar, B.S. (1992) Association of Vibrio cholerae with fresh water amoebae. *J Med Microbiol* **36**: 303–306.

Thomas, V. and McDonnell, G. (2007) Relationship between mycobacteria and amoebae: Ecological and epidemiological concerns. *Lett Appl Microbiol* **45**: 349–357.

Tomov, A.T., Tsvetkova, E.D., Tomova, I.A., Michailova, L.I., and Kassovski, V.K. (1999) Persistence and multiplication of obligate anaerobe bacteria in amebae under aerobic conditions. *Anaerobe* **5**: 19–23.

Tsao, H.F., Scheikl, U., Volland, J.M., Köhsler, M., Bright, M., Walochnik, J., and Horn, M. (2017) "Candidatus Cochliophilus cryoturris" (Coxiellaceae), a symbiont of the testate amoeba Cochliopodium minus. *Sci Rep* **7**: 3394.

Van Der Henst, C., Scrignari, T., Maclachlan, C., and Blokesch, M. (2016) An intracellular replication niche for Vibrio cholerae in the amoeba Acanthamoeba castellanii. *ISME J* **10**: 897–910.

Verani, J.R., Lorick, S.A., Yoder, J.S., Beach, M.J., Braden, C.R., Roberts, J.M., et al. (2009) National outbreak of Acanthamoeba keratitis associated with use of a contact lens solution, United States. *Emerg Infect Dis* **15**: 1236–1242.

Walochnik, J. and Aspöck, H. (2005) Die diagnostik von Infektionen mit freilebenden amöben (FLA). *LaboratoriumsMedizin* **29**: 446–456.

Walochnik, J. and Aspöck, H. (2007) Amöben : Paradebeispiele für Probleme der Phylogenetik , Klassifikation und Nomenklatur. In *Denisia* **20**: 323-350

Wang, J., Liu, M., Xiao, H., Wu, W., Xie, M., Sun, M., et al. (2013) Bacterial community structure in cooling water and biofilm in an industrial recirculating cooling water system. *Water Sci Technol* **68**: 940–947.

Xuan, Y.H., Yu, H.S., Jeong, H.J., Seol, S.Y., Chung, D. Il, and Kong, H.H. (2007) Molecular characterization of bacterial endosymbionts of Acanthamoeba isolates from infected corneas of Korean patients. *Korean J Parasitol* **45**: 1–9.

Chapter II

Overview of publications / Manuscripts

Research goal

This thesis aims to elucidate the association of microbes with amoeba at multiple scales, ranging from microbial communities to relationships between to partners and the molecular mechanisms underlying these associations.

Chapter III

Open cooling towers of factories, hospitals, or other large public buildings are known sources of legionellosis-outbreaks by spreading aerosols over vast distances. Despite the release of the guideline ÖNORM B 5020:2013 to improve water quality and reduce bacterial growth in cooling towers, a routinely performed standardized disinfection procedure and screening for microbial contaminants are not obligatory. The overall microbial community and protist predators largely determine the survival and dispersal of bacterial pathogens in this aquatic niche, but such information in these systems are mostly elusive. What is the microbial community composition, and how stable is it over time, and how similar is it between different cooling towers? How do these man-made systems compare to natural aquatic systems? Can known or novel protist-bacteria interactions be discovered using state-of-the-art cultivation-independent approaches? To address these questions, three different cooling towers located in Vienna were investigated over the time course of one year. I used amplicon sequencing techniques on the collected water samples and applied various bioinformatic analysis approaches to elucidate bacterial and protist diversity. Comparison between the different sampling locations as well as time points was performed to gain insight into the bigger picture. In addition, co-occurrence analysis was used to predict interactions between bacteria and protists.

Manuscript title:

The cooling tower water microbiota: Seasonal dynamics and co-occurrence of bacterial and protist phylotypes

Author names:

Han-Fei Tsao, Ute Scheikl, Craig Herbold, Alexander Indra, Julia Walochnik, Matthias Horn

Reference:

Water Res. (2019), doi:10.1016/j.watres.2019.04.028.

Author contributions:

H.F.T., U.S., A.I., J.W., and M.H. conceived the study. H.F.T. and U.S. performed the sampling and DNA isolation; H.F.T. performed PCR and prepared sequencing reactions. C.H. performed the demultiplexing, filtering and clustering of the reads. H.F.T. performed the analysis of the data and prepared the figures; H.F.T. and M.H. wrote the manuscript. All authors reviewed and edited the manuscript.

Chapter IV

Bacteria-amoeba relationships have been studied almost exclusively in isolates of free-living naked amoebae such as Acanthamoeba, Vanella, and Vermamoeba. From one of the cooling water samples described in Chapter III, I managed to isolate the first bacterial endosymbiont of a testate amoeba. Primarily found in the freshwater and marine environment, the amoeba host *Cochliopodium minus* is covered in a scale-like carbohydrate coat. The endosymbiont was named '*Candidatus* Cochliophilus cryoturris', due to its affiliation with a *Cochliopodium* host and the sample origin. In this chapter, my goal was to reveal the phylogeny of this novel endosymbiont and the exact intracellular location they are residing in. How do they compare to other amoeba-associated-bacteria?

To answer these questions, I used 16S rRNA gene sequencing and phylogenetic analysis to infer the taxonomic placement of C. cryoturris. An endosymbiont-specific probe targeting the 16S ribosomal RNA was designed and applied for fluorescence in situ hybridization assays to visualize the amoeba-symbiont interaction. More data were gathered by using transmission electron microscopy to reveal the intracellular structures of a *Cochliopodium* host infected with *C. cryoturris*.

Manuscript title:

'*Candidatus* Cochliophilus cryoturris' (Coxiellaceae), a symbiont of the testate amoeba *Cochliopodium minus*

Author names:

Han-Fei Tsao, Ute Scheikl, Jean-Marie Volland, Martina Köhsler, Monika Bright, Julia Walochnik, Matthias Horn

Reference:

Sci. Rep. **7**, 3394 (2017). doi: 10.1038/s41598-017-03642-8

Author contributions:

H.F.T., U.S., J.W., and M.H. conceived the study. H.F.T. and U.S. performed the sampling; U.S. and M.K. isolated and characterized the amoeba. H.F.T. identified and characterized the bacterial symbiont. J.M.V. and M.B. performed electron microscopy analysis. H.F.T. prepared the figures; H.F.T. and M.H. wrote the manuscript. All authors reviewed and edited the manuscript.

Chapter V

Procabacter spp. are endosymbionts of *Acanthamoeba,* first observed in 1975 (Proca-Ciobanu et al. 1975). The Gram-negative, rod-shaped members were found in various clinical and environmental amoeba isolates from all over the world. FISH assays have shown their presence in the host's cytosol, and phylogenetic analysis determined their placement within the Betaproteobacteria (Horn et al. 2002). Chapter V describes in more detail the infection process of *Procabacter acanthamoebae*, the representative *Procabacter* strain. How long is the infection process, and what impact on fitness does an infection have? What is the host spectrum? Can *Procabacter* survive without a host cell? The chapter also dives into the genomic detail and makeup of the *Procabacter* group. What is the proper *Procabacter* phylogeny within the Betaproteobacteria? How does the genome compare to genomes of other amoeba-associated bacteria? Do the genomes exhibit typical hallmarks of organisms with an intracellular lifestyle? I aimed to address these questions by conducting infection experiments under various conditions and with different host cell types. With FISH in combination with a *Procabacter* - specific probe, we visualized the infection cycle. For the comparative genome analysis, I first used next-generation sequencing to obtain draft genome sequences. The genome sequences were then annotated with an in house annotation pipeline and then compared to each other and to genomes of other amoeba endosymbionts.

Manuscript title:

Better in than out: *Procabacter acanthamoebae*, an amoebae endosymbiont in the making

Author names:

Han-Fei Tsao, Stephan Köstlbacher, Matthias Horn

Reference:

Draft manuscript

Author contributions:

H.-F.T., S.K., and M.H. conceived the study. H.-F.T. and S.K. performed the genomic DNA isolation. S.K. performed the infection cycle, biofilm-related experiments, and the genome assembly. H.-F.T. and S.K. did the phylogenetic analysis. H.-F.T. did the genome annotation and ancestor genome reconstruction. H.-F.T. prepared the figures and tables. H.-F.T. wrote the manuscript.

Chapter VI

Secretion systems are essential facilitators of interactions between bacteria and their eukaryotic hosts, especially for the transfer of effector protein (Green and Mecsas 2016). The *Acanthamoeba* endosymbiont *Amoebophilus asiaticus* belongs to the diverse phylum Bacteroidetes. The *Amoebophilus* genome is unique, as it contains a high percentage of mobile genetic elements and encodes for a large number of proteins predicted to be secreted and to function in host-cell interaction. Yet, no known secretion systems, usually responsible for the transfer of effector-proteins into the host cell, were found (Schmitz-Esser et al. 2010). However, a genomic region with similarity to prophages was discovered in the *Amoebophilus* genome, which might act as a putative secretion system (Penz et al. 2010). In this chapter, we aimed to characterize the structure and function of this putative secretion system. By using cryo-focused ion beam milling combined with electron cryotomography, the apparatus' structure was revealed. In combination with infection assays, transcriptomics data, and functional assays, the interaction of the apparatus with membranes was shown. Further, sequence analysis unveiled the evolutionary history of this novel secretion system.

Manuscript title:

In situ architecture, function, and evolution of a novel contractile injection system

Author names:

Désirée Böck, João M. Medeiros, Han-Fei Tsao, Thomas Penz, Gregor L. Weiss, Karin Aistleitner, Matthias Horn, Martin Pilhofer

Reference:

Science (80-.). **717**, 713–717 (2017). doi: 10.1126/science.aan7904

Author contributions:

D.B., J.M.M., H.-F.T., T.P., G.L.W., K.A., and M.P. collected data. All authors analyzed and discussed data. D.B. and M.P. wrote the paper with input from the co-authors. D.B. and J.M.M. contributed equally to the study.

References

Green, E.R. and Mecsas, J. (2016) Bacterial Secretion Systems: An Overview. *Microbiol Spectr* **4**:.

Horn, M., Fritsche, T.R., Linner, T., Gautom, R.K., Harzenetter, M.D., and Wagner, M. (2002) Obligate bacterial endosymbionts of Acanthamoeba spp. related to the β-Proteobacteria: Proposal of "Candidatus Procabacter acanthamoebae" gen. nov., sp. nov. *Int J Syst Evol Microbiol* **52**: 599–605.

Penz, T., Horn, M., and Schmitz-Esser, S. (2010) The genome of the amoeba symbiont "candidatus amoebophilus asiaticus" encodes an afp-like prophage possibly used for protein secretion. *Virulence* **1**: 541–545.

Proca-Ciobanu, M., Lupascu, G., Petrovici, A., and Ionescu, M.D. (1975) Electron microscopic study of a pathogenic Acanthamoeba castellani strain: The presence of bacterial endosymbionts. *Int J Parasitol* **5**: 49–56.

Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., and Horn, M. (2010) The genome of the amoeba symbiont "Candidatus Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J Bacteriol* **192**: 1045–1057

Chapter III

The cooling tower water microbiota: Seasonal dynamics and co-occurrence of bacterial and protist phylotypes

Water Research 159 (2019) 464–479

Contents lists available at ScienceDirect

Water Research

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

The cooling tower water microbiota: Seasonal dynamics and co-occurrence of bacterial and protist phylotypes

Han-Fei Tsao [a], Ute Scheikl [b], Craig Herbold [a], Alexander Indra [c], Julia Walochnik [b], Matthias Horn [a, *]

[a] Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria
[b] Institute of Specific Prophylaxis and Tropical Medicine, Medical University of Vienna, Vienna, Austria
[c] Department of Mycobacteriology and Clinical Molecular Biology, AGES, Vienna, Austria

ARTICLE INFO

Article history:
Received 18 September 2018
Received in revised form
4 April 2019
Accepted 13 April 2019
Available online 22 April 2019

Keywords:
Microbial diversity
Legionella
Mycobacteria
Chlamydia
Protists
Free-living amoeba
Built environment
HVAC cooling tower

ABSTRACT

Cooling towers for heating, ventilation and air conditioning are ubiquitous in the built environment. Often located on rooftops, their semi-open water basins provide a suitable environment for microbial growth. They are recognized as a potential source of bacterial pathogens and have been associated with disease outbreaks such as Legionnaires' disease. While measures to minimize public health risks are in place, the general microbial and protist community structure and dynamics in these systems remain largely elusive. In this study, we analysed the microbiome of the bulk water from the basins of three cooling towers by 16S and 18S rRNA gene amplicon sequencing over the course of one year. Bacterial diversity in all three towers was broadly comparable to other freshwater systems, yet less diverse than natural environments; the most abundant taxa are also frequently found in freshwater or drinking water. While each cooling tower had a pronounced site-specific microbial community, taxa shared among all locations mainly included groups generally associated with biofilm formation. We also detected several groups related to known opportunistic pathogens, such as *Legionella*, *Mycobacterium*, and *Pseudomonas* species, albeit at generally low abundance. Although cooling towers represent a rather stable environment, microbial community composition was highly dynamic and subject to seasonal change. Protists are important members of the cooling tower water microbiome and known reservoirs for bacterial pathogens. Co-occurrence analysis of bacteria and protist taxa successfully captured known interactions between amoeba-associated bacteria and their hosts, and predicted a large number of additional relationships involving ciliates and other protists. Together, this study provides an unbiased and comprehensive overview of microbial diversity of cooling tower water basins, establishing a framework for investigating and assessing public health risks associated with these man-made freshwater environments.

1. Introduction

Most modern commercial, industrial, and residential buildings rely on cooling towers as cost-efficient measures to remove excess heat. Containing large semi-open water volumes at a rather constant temperature, cooling towers are suitable environments for microbial growth throughout the year and have been implicated in bacterial outbreaks (Kurtz et al., 1982; Pagnier et al., 2009; Torvinen et al., 2013; Yamamoto et al., 1992). One of the major public health risks associated with cooling towers is Legionnaires' disease. This respiratory tract infection is caused by *Legionella pneumophila*, a gammaproteobacterial pathogen acquired through inhalation of aerosols and is potentially fatal for immunocompromised patients (Hamilton et al., 2018; Walser et al., 2014). Other opportunistic bacterial pathogens detected in cooling towers include *Mycobacterium* spp., *Pseudomonas* spp., *Burkholderia* spp., and *Pantoea* spp. (Ceyhan and Ozdemir, 2008).

Protists, such as free-living amoebae, are common members of microbial communities in cooling towers (Barbaree et al., 1986; Berk et al., 2006; Delafont et al., 2016; Pagnier et al., 2009). They are predators of other microbes in the system and may also be reservoirs for bacterial pathogens such as *Legionella* and *Mycobacterium*

* Corresponding author.
E-mail address: horn@microbial-ecology.net (M. Horn).

https://doi.org/10.1016/j.watres.2019.04.028

spp., *Pseudomonas aeruginosa*, *Francisella tularensis*, *Coxiella burnetii*, and *Vibrio cholerae* (Abd et al., 2003; La Scola and Raoult, 2001; Sandström et al., 2010; Thomas and McDonnell, 2007; Thom et al., 1992; Van der Henst et al., 2016). These microbes are able to escape the regular phagolysosomal pathway and transiently replicate within amoeba trophozoites. Moreover, many free-living amoebae can form cysts, a resistant life stage providing protection from unfavorable environmental conditions including biocides (Fouque et al., 2012; Miltner and Bermudez, 2000). Efficient disinfection procedures targeting not only bacteria but also protists have thus been recognized as important in reducing the public health risk associated with cooling towers (Critchley and Bentham, 2009; Scheikl et al., 2016).

Bacteria-protist relationships are manifold (Cirillo et al., 1997, 1994; Fritsche et al., 1998; Greub and Raoult, 2004; Hess, 2016; Hess, 2017; Molmeret et al., 2005). They also include obligate bacterial symbionts, such as *Caedibacter* and *Holospora* species in ciliates (Dziallas, 2012; Fokin, 2004; Goertz, 2001), or *Protochlamydia* and *Parachlamydia* species, also known as environmental chlamydiae (Horn, 2008; Taylor-Brown et al., 2015). Yet, microbial communities and their interactions in cooling tower water are likely far more complex than recognized. The presence of specific bacterial pathogens has been investigated (Inoue et al., 2015; Li et al., 2015; Liu et al., 2011, 2009; States et al., 1987; Torvinen et al., 2013; Ulleryd et al., 2012), but surveys on the overall microbial community composition are rare (Llewellyn et al., 2017; Wang et al., 2013). Studies specifically investigating community assembly and dynamics in these systems are largely lacking, with currently only one recent report (Pereira et al., 2017). Except for a few notable studies (Llewellyn et al., 2017; Pereira et al., 2017; Wang et al., 2013), work on the microbial community of water towers has been based on cultivation dependent methods or PCR assays targeting specific microbial groups, which are inherently limited in both taxonomic resolution and scope. High-throughput DNA sequencing of 16S/18S ribosomal RNA gene PCR products, termed amplicon sequencing, is able to circumvent these limitations and provides detailed data on both bacterial and protist communities (Thompson et al., 2017). After all, it is the wide range of interactions - from antagonistic to synergistic - that shape microbial assemblages and ultimately determine the chances for pathogen proliferation.

In this study, we used amplicon sequencing to investigate seasonal and geographic variation in the composition of bacterial and protist communities found in three cooling towers over a one-year period. We focused our analysis on the bulk water in the water basins, because it is the source of aerosols formed during cooling tower operation (Nhu Nguyen et al., 2005). Bacteria-protist relationships were inferred from co-occurrence network analysis. These approaches identified key microbial players and provided detailed insight into the microbial community dynamics in these systems including putative bacteria-protist interactions.

2. Materials and methods

2.1. Site description, sample collection, and processing

Three different heating, ventilation and air conditioning (HVAC) cooling towers located on the rooftop of different buildings in the city of Vienna, Austria, were studied. Two buildings, a hospital and a business complex, were located near the city center (1.7 km apart from each other) whereas a second hospital was located at Vienna's periphery (5 km from the other two locations). The two hospitals are hereafter referred to as ("cooling tower 1") CT-1 and CT-2, the business complex as CT-3. For all three sites, the water temperature was recorded throughout the sampling campaign. The cooling towers were subjected to disinfection with chlorine/bromine-based disinfectant (1–3 g/m^3) over an automated dosimeter. CT-1 was disinfected three times a week, CT-2 every 18 h and CT-3 every four days.

During the sampling period between September 2013 and September 2014, water samples were taken on a biweekly basis. CT-3 was not operating during the winter season, resulting in fewer samples. From the tank basin, 3 L of bulk water was sampled and stored at 4 °C for up to a day before further processing. The water was stirred first and then 2 L per sample were vacuum-filtered onto a cellulose nitrate filter (diameter 0.2 μm; 12.5 cm^2; Sartorius stedim) in 500 ml steps. In case of rapid filter clogging, filters with a slightly larger pore size (diameter 0.45 μm) were used in addition for further filtration of the same sample. Up to four filters were used per sample. All filters were subsequently used for downstream processing.

2.2. DNA extraction, and 16S/18S rRNA gene amplicon sequencing

The PowerWater® DNA Isolation Kit (Qiagen, Hilden, Germany) was used for DNA extraction from the filters. All steps were performed according to the standard centrifuge-based protocol recommended by the manufacturer. Briefly, the filters were inserted in a lysis buffer-containing bead beating tube where the cells were mechanically and chemically lysed. The lysates were transferred to a DNA-retaining spin column; lysates from the same sample were loaded onto the same column to pool the DNA. After washing, DNA was eluted in the DNA elution buffer and used for PCR.

A two-step barcoded amplicon sequencing approach was used as described previously (Herbold et al., 2015). Briefly, fragments of the bacterial and eukaryotic small subunit rRNA gene were amplified, resulting in 26 libraries each for CT-1 and CT-2, and 14 libraries for CT-3. V3 and V4 regions of the bacterial 16S rRNA gene were amplified with the primers Bakt_341F (5′-CCTACGGGNGGCWGCAG-3′) and Bakt_805R (5′-GACTACHVGGGTATCTAATCC-3'; Herlemann et al., 2011). Archaea were not targeted by this primer set and were not analysed in this study. The eukaryotic 18S rRNA gene primer-pair EUK_1391F (5′-GTACACACCGCCCGTC-3′) and EUK_1510R (5′-CCTTCYGCAGGTTCACCTAC-3′) based on the Earth Microbiome Project (Version 5, 2012; Amaral-Zettler et al., 2009; Gilbert et al., 2014) were used to amplify the 18S–V9 region. For each primer set and sequencing run, negative controls without the addition of DNA were performed and sequenced.

Each PCR reaction included 1x DreamTaq Green Buffer (Fermentas, Thermo Fisher Scientific, Vienna, Austria), 2 mM $MgCl_2$, 0.2 mM dNTP mix (Fermentas), 0.1 mg mL^{-1} bovine serum albumin, 1 μM of each of the forward and reverse primers, 0.025 U DreamTaq polymerase (Fermentas), and 1 μL of DNA template. This first PCR amplification (94 °C for 3 min; 25 cycles of 45 s at 94 °C, 30 s at 52 °C (16S rRNA gene) or 57 °C (18S rRNA gene), 90 s at 72 °C; and 72 °C for 10 min) was performed in triplicates; PCR products were pooled and served as the template for the second barcoding PCR (95 °C for 3 min; 10 cycles of 30 s at 95 °C, 30 s at 55 °C, 60 s at 72 °C; and 72 °C for 7 min). Purification and quantification of PCR products were done as described (Herbold et al., 2015). The TruSeq Nano DNA Library Prep Kit (Illumina) was used for adaptor ligation and PCR without the fragmentation step. Sequencing was performed by Microsynth AG (Balgach, Switzerland) on a MiSeq system (Illumina) using the MiSeq Reagent kit V3. Resulting sequence datasets were deposited in the NCBI Sequence Read Archive under study accession number PRJEB21563.

2.3. Data compilation, filtering, clustering, and classification

Paired-end reads were demultiplexed and trimmed as described

(Herbold et al., 2015) and were assembled using fastq-join within Qiime (Caporaso et al., 2010). 18S rRNA reads were forced to be trimmed by at least 88 nucleotides prior to trimming and assembly because the expected amplicon length of 178 bp was shorter than the average read length. Clustering into operational taxonomic units (OTUs) and chimera-checking were performed using Uparse (Edgar, 2013), using a distance of 0.03 and disallowing singletons to serve as a centroid. Taxonomic classification of OTU centroids was carried out using the Bayesian classifier implemented in mothur (Schloss et al., 2009) and the Silva SSU database (Version 132; Quast et al., 2013) for 16S rRNA data sets, or the protist ribosomal reference database PR^2 (Version 4.11.1; Guillou et al., 2013) for 18S rRNA data sets. OTUs with a classification confidence score lower than 80% for 16S rRNA data sets and 60% for 18S rRNA data sets were changed to 'unclassified'. The class Betaproteobacteria has recently been proposed to be reclassified as the new order Betaproteobacteriales within the class Gammaproteobacteria (Parks et al., 2018). For clarity and consistency with previous studies, the Betaproteobacteria were still treated as a separate class, but both names are provided. 16S rRNA OTUs classified as mitochondria and chloroplasts were omitted from subsequent analysis; this did not change the overall estimates of alpha and beta diversity measures. 16S rRNA OTUs present only in the negative controls or strongly overrepresented (50 fold) in the negative controls compared to all samples were considered to be cross-contamination and thus removed from the dataset (Table S1).

All 18S rRNA OTUs only present in the negative control were removed from the dataset. OTUs present in both control and samples were only retained if their sample-abundance exceed the control-abundance by 1000 fold (Table S2). Taking into account the known relationships between protists and pathogenic bacteria in the water-borne environment, we restricted our analysis of 18S rRNA sequences to the groups Amoebozoa (Lobosa, Conosa), Excavata (Discoba, Metamonada), Rhizaria (Cercozoa, Foraminifera) and Alveolata (Apicomplexa, Ciliophora, Dinoflagellata). Both data sets were rarefied individually to standardize sequence numbers using the rarify_even_depth function in the R-package phyloseq (McMurdie and Holmes, 2013) with replacement option turned off. For the 16S rRNA data set a minimum of 1713 sequences was used, representing the 15% quantile of the data; for the 18S rRNA data set, which had a greater library size variation, a cutoff of 59 sequences equivalent to 45% quantile of the data was used. Samples not fulfilling these criteria were removed from the subsequent analysis.

2.4. Alpha and beta diversity measures and co-occurrence analysis

Alpha diversity was assessed for both data sets with the R-package *phyloseq* (McMurdie and Holmes, 2013) using the estimators Chao1 and the Shannon diversity index. Bacterial community composition was compared by non-metric multidimensional scaling (NMDS) analysis based on Bray-Curtis dissimilarity.

For the OTU co-occurrence analysis, only those samples for which both 16S and 18S rRNA sequence data were available after rarefying were included. In addition, only OTUs present in at least 3 samples (for each cooling tower) and with a total number of sequences greater than 10 were kept to remove the number of infrequent OTUs. For each cooling tower, co-occurrence analysis of 16S rRNA and 18S rRNA OTUs was performed using CoNet (Faust and Raes, 2016), a plug-in of the software Cytoscape (Shannon et al., 2003). To calculate consensus networks, a combination of different correlation and distance measures including Pearson and Spearman correlation, mutual information, Kullback-Leibler divergence, and Bray-Curtis dissimilarity was used. The maximum lag was set to 1 to include slightly shifted associations. Distribution of all pair-wise scores was generated, and the 7.5% top-scoring edges supported by at least four out of five (positive correlations) or three out of five (negative correlations) measures were kept. This process was performed with 1000 renormalized permutations, resampling by row-shuffling, and p-value merging with Brown's method. After applying Benjamini-Hochberg's false discovery rate correction, edges with a corrected p-value below 0.05 were kept.

The R-package ampvis (Albertsen et al., 2015) was used to generate heatmaps; MetacodeR (Foster et al., 2017) was used to generate heat trees.

3. Results and discussion

In this study, we monitored the microbial community in the water basins of three heating, ventilation and air conditioning (HVAC) cooling towers - in the following referred to as CT-1, CT-2, and CT-3 - over the course of one year. The temperature in the basins ranged between 20.1 °C and 32.3 °C for CT-1 (99 percentile of all temperature data points), and between 19.9 °C and 29 °C for CT-2 throughout the sampling period. The average temperature during the summer months was slightly higher than during winter; 27.8 °C vs. 26.4 °C for CT-1, and 26.6 °C vs. 23.1 °C for CT-2. Temperature data for CT-3 were only available after mid-April 2014; its temperature range was 10.6–27.4 °C, larger than that of the other two towers. Elevated temperature in the water basin is a characteristic feature of cooling towers and together with the semi-open design of these systems provide good conditions for microbial growth. Cooling tower water tanks, like other freshwater or drinking water systems, develop biofilms (Liu et al., 2009). Yet, for the purpose of this study we focused on the analysis of bulk water for two reasons: (i) bulk water represents the seed community for biofilm formation in the basin, and (ii) it is the source of aerosols formed during cooling tower operation (Nhu Nguyen et al., 2005).

3.1. Complex but reduced bacterial diversity compared to open freshwater environments

In total, we collected 26 samples each for CT-1 and CT-2, and 14 samples for CT-3. Amplicon sequencing of 16S and 18S rRNA genes revealed taxonomically diverse bacterial and protist communities for all three cooling towers. To study seasonal dynamics of microbial diversity in the three cooling towers, we first estimated species-level richness and evenness using a threshold of 97% sequence similarity for operational taxonomic unit (OTU) definition. Good's coverage estimator was 0.984 ± 0.007 (mean $\pm$ SEM) for 16S OTUs and 0.958 ± 0.018 (mean $\pm$ SEM) for 18S OTUs, indicating that sequencing depth was sufficient for reliable community composition analysis. After rarefying, our data set included 23 CT-1 samples, 21 CT-2 samples, and 12 CT-3 samples. In total 590 bacterial OTUs and 321 protist OTUs were considered for our analysis (Tables S3 and S4).

The three cooling towers showed clear differences in bacterial and protist species richness and evenness (Fig. 1). As a general pattern, protist diversity was much lower than bacterial diversity for all samples and sampling sites. Microbial diversity fluctuated strongly throughout the sampling period, and there was no clear common trend observed among the different cooling towers. It is interesting that while CT-1 had the highest bacterial diversity values (as inferred from Chao1 and Shannon estimators; Welch's *t*-test p-value < 0.0001), it also showed the lowest protist diversity and evenness compared to the other two locations, which were also more similar with each other (Fig. 1). This was also reflected in the NMDS-based beta-diversity analysis (Welch-corrected *t*-test of the average Bray-Curtis distance within one sampling site compared to other sites; p-value < 0.0001), revealing clear sample clusters according to sampling site and time point (Fig. 2). At all sampling

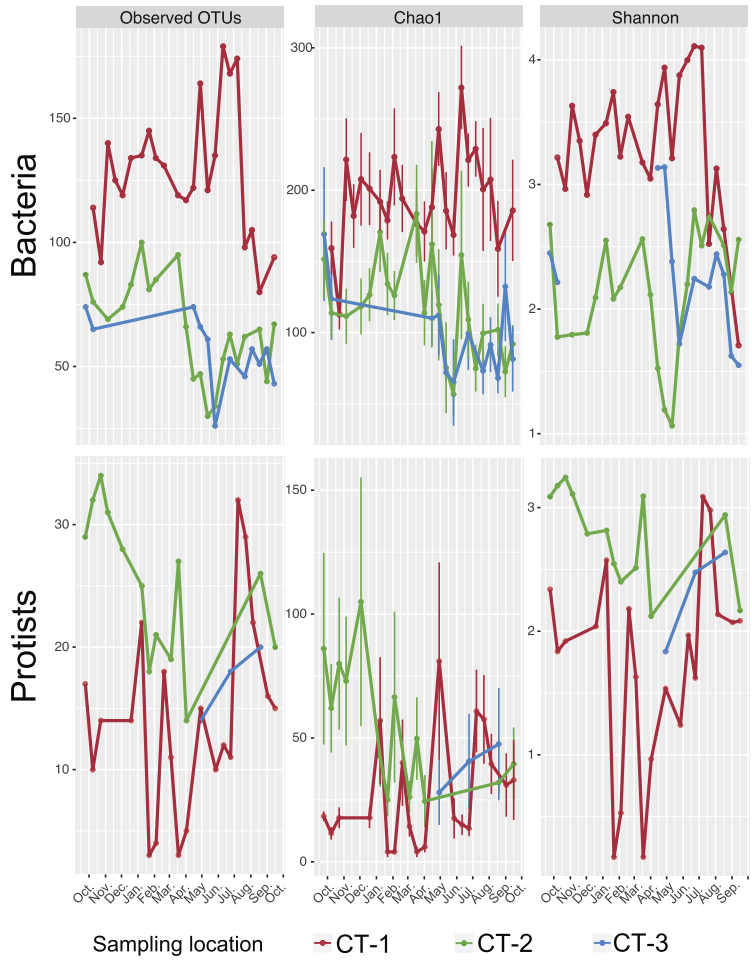


Fig. 1. Bacterial and protist diversity in three cooling tower water basins. The number of observed OTUs and diversity indices for estimation of species richness and evenness of bacterial and protist communities are indicated for cooling towers CT-1, CT-2, and CT-3 over a sampling period of one year (September 2013 to September 2014). Pronounced differences between sampling sites and time points in both bacterial and protist diversity can be observed.

sites, we observed gradual community shifts over time reflecting seasonal changes in bacterial and protist diversity. In CT-1, the high similarity between autumn, winter, and spring samples indicated a gradual rather than sudden change. In contrast, the community shift in CT-2 was more pronounced (Fig. 2).

To better understand community structure we next classified all OTUs, grouped them at the genus level, and focused on the 20 most abundant genera (Fig. 3). CT-2 and CT-3 are both dominated by a few abundant genera, whereas in CT-1, abundances of the top genera are more evenly distributed. This analysis showed that the general bacterial community composition is different between the three cooling towers, yet *Flavobacterium* (Bacteroidetes), *Hyphomicrobium,* members of the Rhizobiales and Sphingomonadales (Alphaproteobacteria), *Pseudomonas* (Gammaproteobacteria), and *Methyloversatilis* (Betaproteobacteria/Betaproteobacteriales; Parks et al. 2018) are among the top abundant taxa in all three towers (Fig. 3). The most striking difference between the cooling towers is the extremely high *Pseudomonas* abundance in CT-2 (Fig. 3), which makes up almost 31.8% of all sequence reads (compared to up to 3% at the other two sampling sites) while other Gammaproteobacteria were generally not well represented in our data set. Our observations are consistent with findings from the only cooling tower for which a similar analysis is available so far (Pereira et al., 2017). In this study, members of the Rhizobiales, Burkholderiales, Methylophilales, and Cytophagales were also among the most abundant taxa. In contrast, we didn't detect Xanthomonadales among the top taxa, but found other groups to be well represented, including Corynebacteriales, Rhodocyclales, and Pseudomonadales (Fig. 3). It is noteworthy, that Betaproteobacteria (Betaproteobacteriales), commonly found in freshwater systems (Newton et al., 2011), are with two exceptions absent in the CT-1 top genera list, whereas they are abundant in CT-2 and CT-3.

Two members of the recently identified Candidate Phyla Radiation (CPR) are among the most abundant taxa in CT-1: Parcubacteria (formerly OD1) and Saccharimonadia (formerly known as TM7), both now classified within the phylum Patescibacteria. Members of both groups show reduced genomes and only limited metabolic capabilities, likely endorsing a parasitic or symbiotic

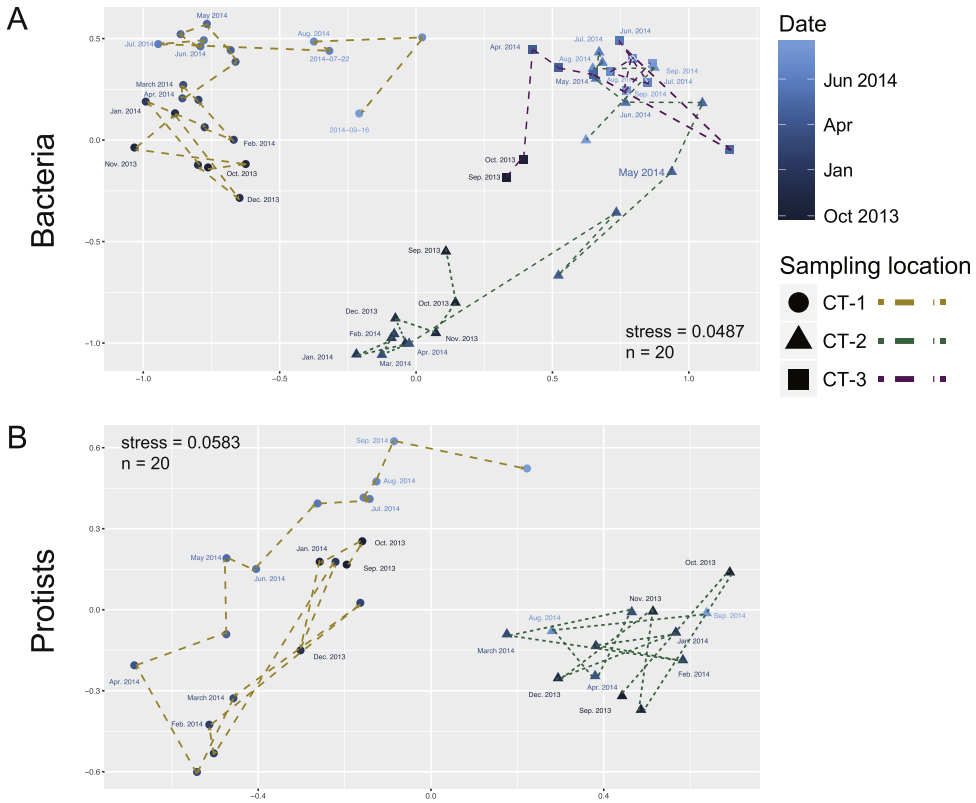


Fig. 2. Similarity of bacterial and protist community structure across time and space. Nonmetric multidimensional scaling (NMDS) analysis illustrating beta diversity among all samples based on the Bacterial or protist community composition (protist diversity was not included in CT-3 samples due to low sample number). For each cooling tower, the samples were sequentially connected with arrows. The first sample of each month are annotated with the date in the plot. Microbial communities of cooling tower CT-1 are clearly separated in this analysis, indicating a unique community composition. Shifts in bacterial community structure apparently follow a seasonal trend.

lifestyle (He et al., 2014; Nelson and Stegen, 2015). They are known to have small cell sizes and are able to pass through a 0.22 μm filter (Luef et al., 2015). CT-1 also included an unclassified member of the Rickettsiales (Alphaproteobacteria) among the top abundant taxa, again a group of microbes generally associated with an intracellular lifestyle.

Overall, the most abundant bacterial orders identified in the cooling tower samples are well in agreement with other reports on man-made water distribution systems, where orders belonging to Alpha-, Beta-, and Gammaproteobacteria are usually the prevalent taxa (Berry et al., 2006; Brown et al., 2015; Farenhorst et al., 2017; Kalmbach et al., 1997; Llewellyn et al., 2017; Martiny et al., 2005; McCoy and VanBriesen, 2012; Pereira et al., 2017; Revetta et al., 2010; Tokajian et al., 2005; Vaz-Moreira et al., 2012; Wang et al., 2013; Zhang et al., 2012). Bacterial communities in more natural settings such as lakes, rivers, or streams are often dominated by Betaproteobacteria (Betaproteobacteriales), but also include Actinobacteria, Bacteroidetes, Cyanobacteria, Verrucomicrobia, and Saccharimonadia (Boucher et al., 2006; Crump and Hobbie, 2005; Liu et al., 2012; Lozupone and Knight, 2005; Mueller-Spitz et al., 2009; Newton et al., 2011; Tang et al., 2015; Zwart et al., 2002). The dominance of these groups in the cooling tower microbial community structure is consistent with what has been reported for other freshwater systems. This suggests that community structure in cooling towers may be primarily influenced by the seed community in the incoming drinking water and then shaped by the specific conditions in the cooling tower water basin.

3.2. Seasonal microbial community dynamics

To further analyse temporal changes in community structure, we examined the relative abundance of phyla over time. Consistent with our previous analysis, we observed pronounced changes in bacterial community composition over time (Fig. 4). The most dramatic seasonal change occurred in cooling tower CT-2, where an abrupt decline of Gammaproteobacteria (*Pseudomonas*) was observed during June 2014 (Fig. 4). Pronounced variations were also observed in the more diverse community of CT-1. In CT-3, the overall bacterial community composition was rather stable compared to the other cooling towers (Fig. 4). Protistan community members of the Rhizaria, the Excavata, the Alveolata, and the Amoebozoa were found at all three sampling sites, with the latter two phyla being most abundant and showing pronounced temporal variations. Notably, changes in bacterial and protist diversity do not necessarily follow the same trend, indicating complex interactions and differences in response to environmental conditions (Figs. 3 and 4).

Next, we aimed to understand differences and similarities between bacterial and protist community composition and dynamics in the three cooling towers in more detail. To this end, we focused

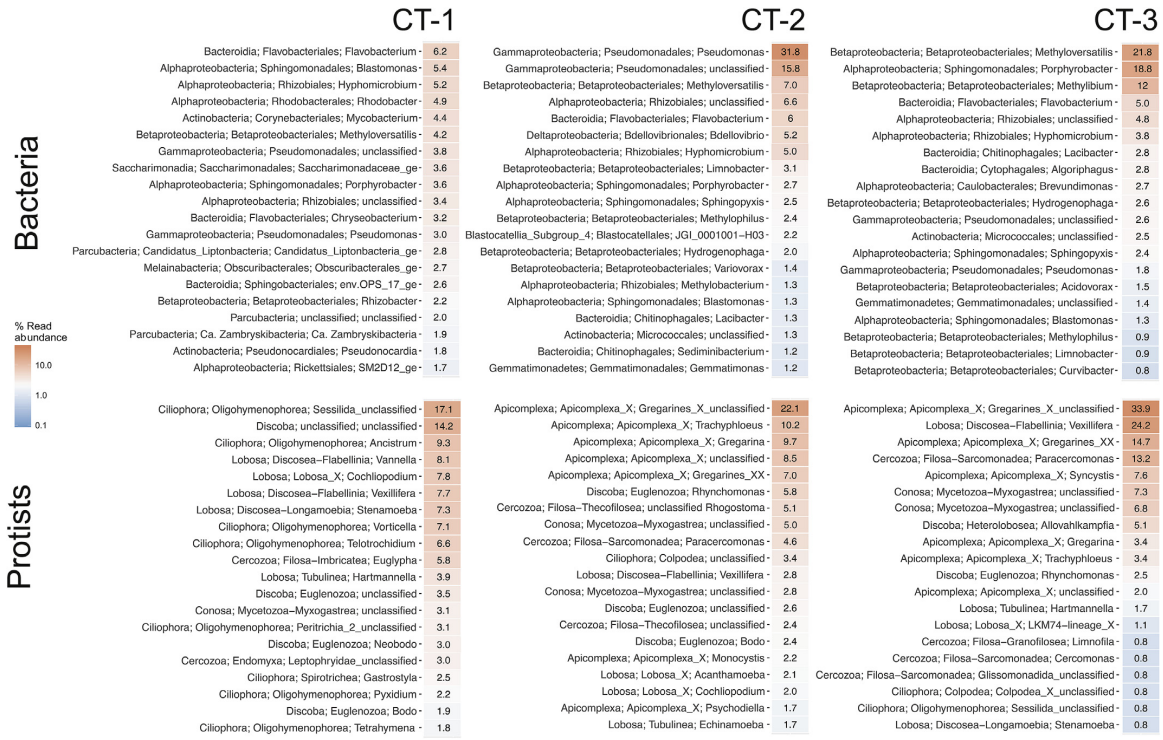


Fig. 3. The most abundant bacterial and protist genus at three cooling towers. For each sampling site, the average relative abundance of the 20 most frequently observed bacteria and protist taxa are listed at the genus level. In total 79% of all bacterial OTUs and 59% of all protist OTUs could be classified at the genus level (Tables S3 and S4). For clarity, the phylum as well as the order level are also indicated (the Proteobacteria were split into its representative classes). The Sphingomonadales, Pseudomonadales, Rhizobiales, and Burkholderiales including many biofilm associated taxa are among the top bacterial groups.

on various sets of OTUs: We defined a set of persistent OTUs as taxa that were detected at a given location at nearly all time points (taxa were allowed to be missing from only two samples, corresponding to at least 90% of all samples for CT-1 and CT-2, and 80% for CT-3). The second set was referred to as site-specific OTUs and included those taxa that were exclusively found at a single location. The third set comprised taxa detected at all cooling towers and was referred to as shared OTUs.

Consistent with our observed seasonal changes, there were few persistent OTUs, 13 in CT-1, 10 in CT-2, and 18 in CT-3 (Fig. 5). They included Actinobacteria (*Mycobacterium* sp., or members of the Microbacteriaceae), Proteobacteria (*Sphingopyxis, Methyloversatilis, Blastomonas, Defluviimonas*, *Porphyrobacter,* Xanthobacteraceae), Patescibacteria (Saccharimonadia, formerly TM7; Albertsen et al., 2013), Bacteroidetes, or Acidobacteria. Interestingly, most of the top five abundant taxa in each of the cooling towers also belonged to the set of persistent OTUs (Figs. 3 and 5).

The differences between the cooling towers were further analysed by focusing on the set of site-specific taxa. Consistent with earlier observations (Fig. 1), CT-1 had by far the largest number of site-specific bacterial OTUs (245, 53.1.%), compared to CT-2 (66, 14.3%) and CT-3 (9, 1.9%; Fig. 6A). It is noteworthy, that there is no overlap between the sets of site-specific taxa and the persistent OTUs for any of the cooling towers. In direct comparison, the site-specific taxa comprised a wider range of diverse phyla, including members of the Patescibacteria, Chlamydiae, Chloroflexi, Planctomycetes, Verrucomicrobia, and Depedentiae (TM6; Delafont et al., 2015; Table S3). The transient character of the site-specific OTUs was further illustrated by most of them being detected at a maximum of nine time points. This indicated that there was a substantial fraction of taxa specific to individual cooling towers that failed to establish long-term colonization.

3.3. The core bacterial microbiome of cooling towers consists of biofilm-forming taxa

To better understand the common features of cooling tower microbial communities, we next focused on the set of shared OTUs, i.e. those detected at least once at more than one site. This set represents 30.6% of all bacterial OTUs in this study, and the fraction of shared OTUs varied between the individual cooling towers (28.6%, 44.9%, and 72.6%, respectively). The majority of shared OTUs belongs to the Alphaproteobacteria (Sphingomonadaceae and Hyphomicrobiaceae), Betaproteobacteria/Betaproteobacteriales (Burkholderiaceae), Actinobacteria, Bacteroidetes (Bacteroidia), and Gammaproteobacteria (Pseudomonadaceae) (Fig. 6B and C). From this large set of shared OTUs only six bacterial taxa were nearly always present, i. e. also belong to the set of persistent OTUs (Fig. 5). This included members of the genera *Pseudomonas, Methyloversatilis*, *Blastomonas, Sphingopyxis,* and *Porphyrobacter*, and the family Xanthobacteraceae (formerly classified as Bradyrhizobiaceae).

A common feature of these microbes is their ability to form or to be associated with microbial biofilms (Rickard et al., 2004; Soto-Giron et al., 2016; Zhang et al., 2012). *Pseudomonas* species are ubiquitous environmental bacteria well-known for their capability

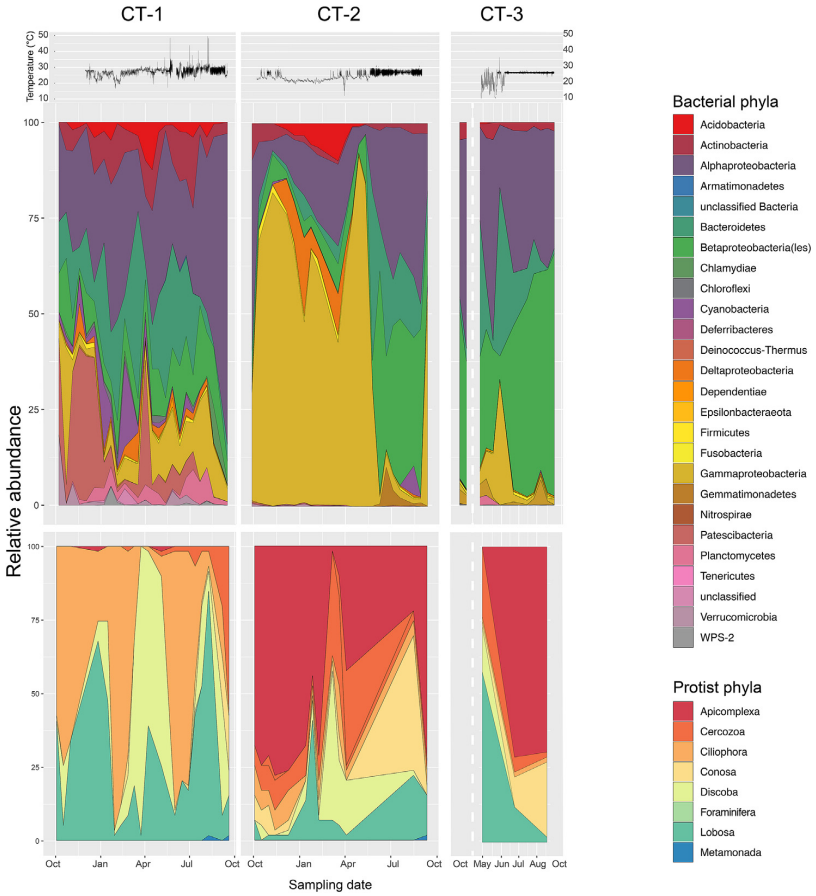


Fig. 4. Temporal dynamics of bacterial and protists diversity over a one year sampling period. Changes in relative abundance of bacteria and protists at the phylum level for cooling towers CT-1, CT-2, and CT-3 reveals temporal dynamics of community structure despite relatively constant environmental conditions. The average daily temperature data in the water basin is shown in the uppermost panel; lack of data for CT-3 is indicated by a white dashed line.

to form biofilms (Costerton, 1999; Jensen et al., 2010; Sauer et al., 2002). They are considered pioneers, specifically facilitating the initiation of biofilm formation (Douterelo et al., 2014). Members of the genus *Methyloversatilis* are found in various environments including drinking water and biofilms (Kalyuzhnaya et al., 2006; Liu et al., 2012; Shpigel et al., 2015; van der Kooij et al., 2017; Williams et al., 2004). *Blastomonas* species have the ability to coaggregate with various bacteria, an important feature for biofilm formation (Katharios-Lanwermeyer et al., 2014; Rickard et al., 2002). Additionally, other members of the order Sphingomonadales such as *Sphingopyxis*, *Porphyrobacter*, and the family Xanthobacteraceae (formerly Bradyrhizobiaceae) were repeatedly found to be associated with bacterial biofilms and are commonly present in drinking water environments (Bereschenko et al., 2010; Buse et al., 2014; Chao et al., 2015; Douterelo et al., 2014; Hong et al., 2010; Muchesa et al., 2017; Revetta et al., 2010; Rożej et al., 2015; Singh et al., 2003; Vaz-Moreira et al., 2011).

The occurrence and persistence of these microbes at all cooling towers suggest that biofilm formation is an important process in cooling tower microbial communities, likely contributing to maintenance and resistance against disinfection measures. Cooling tower basins offer a large surface area and are characterized by a slow water flow speed, a rather constant and elevated temperature, as well as increased concentrations of organic matter compared to closed drinking water systems. These conditions are well-suited for biofilm formation (Türetgen and Cotuk, 2007; Zacheus et al., 2000).

Biofilms are preferred niches of protists because of the high density of microbes serving as a potential food source, and grazing is a major factor modulating the structure and function of bacterial biofilms (Flemming et al., 2016; Matz et al., 2005; Parry, 2004). Biofilms are also a known reservoir and potential source for a number of bacterial pathogens (Delafont et al., 2014; Garcia et al., 2013; Greub and Raoult, 2004; Wingender and Flemming, 2011).

3.4. Opportunistic bacterial pathogens

Over the course of our one-year sampling campaign we detected few OTUs classified as genera comprising known (opportunistic) bacterial pathogens, as defined in a previous study (Pereira et al., 2017; Table S5). The abundance of these OTUs ranges from 0.01% to 28% depending on the sampling location, but the majority are below 1%. We observed a total of 29 OTUs belonging to three different phyla, Actinobacteria, Firmicutes, and Proteobacteria. This includes *Bosea*, *Helicobacter*, and *Stenotrophobacter* and several

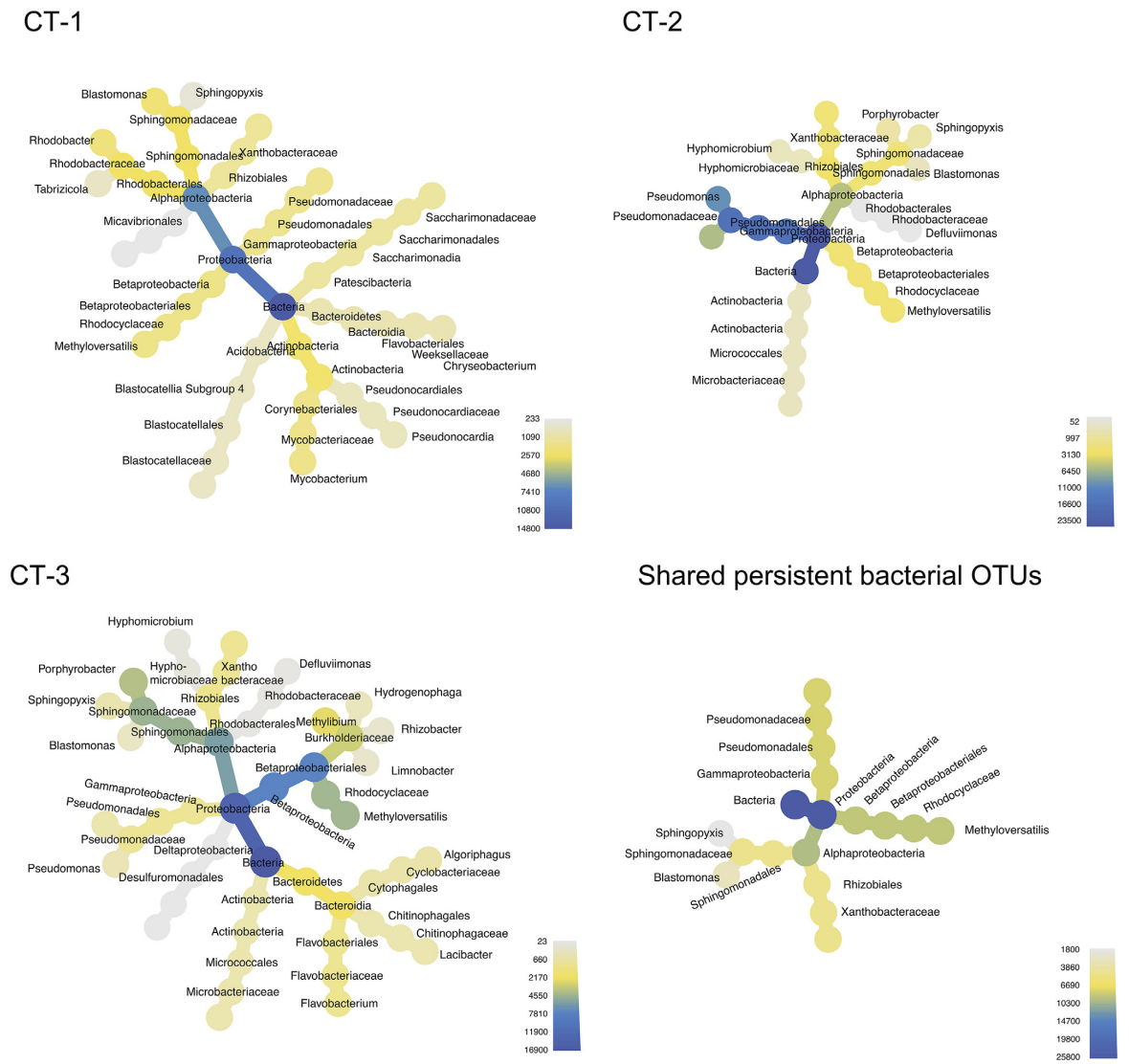


Fig. 5. Persistent bacterial OTUs in three cooling tower water basins. For each sampling site, a heat tree depicts the persistent bacterial taxa at the genus level (unclassified taxa are unlabeled). Persistent taxa are defined as being present in nearly all samples of one sampling location. The color corresponds to the total number of sequences detected for each sampling site at the respective taxonomic level. Shared bacterial OTUs are persistent taxa found in all three sampling sites, all belonging to the Proteobacteria. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Arcobacter, *Bacillus*, *Sphingomonas*, *Legionella*, *Mycobacterium*, and *Pseudomonas* species (Table S5). In comparison to the cooling tower studied earlier, the cooling towers investigated here display a less diverse set of genera including potential pathogens (Pereira et al., 2017).

Of note, none of the *Legionella* sequences showed a particularly high similarity to the human pathogen *Legionella pneumophila*. This is consistent with no *L. pneumophila* outbreak having been reported in the vicinity of our sampling locations during this study. The *Legionella* species detected belong to a group commonly referred to as *Legionella*-like amoebal pathogens, a large number of species primarily found as parasites in free-living amoebae (Greub and Raoult, 2004; Marrie et al., 2001; Newsome et al., 1988). All *Legionella* OTUs belong to the set of site-specific taxa, and their relative abundances were below 0.1%, rendering them members of the rare biosphere in these systems.

We found four OTUs assigned to the genus *Mycobacterium*, two of which occurred in two cooling towers. One of those, *M. wolinsky*, showed a relatively high abundance in CT-1 (4%) and belonged to the persistent taxa in this cooling tower (Table S3, Fig. 5). *M. wolinsky* can cause bacteremia and, like the other mycobacteria detected in this study, is a member of the nontuberculous mycobacteria (NTM) group. This group includes opportunistic human pathogens found in a variety of habitats, ranging from soil, water,

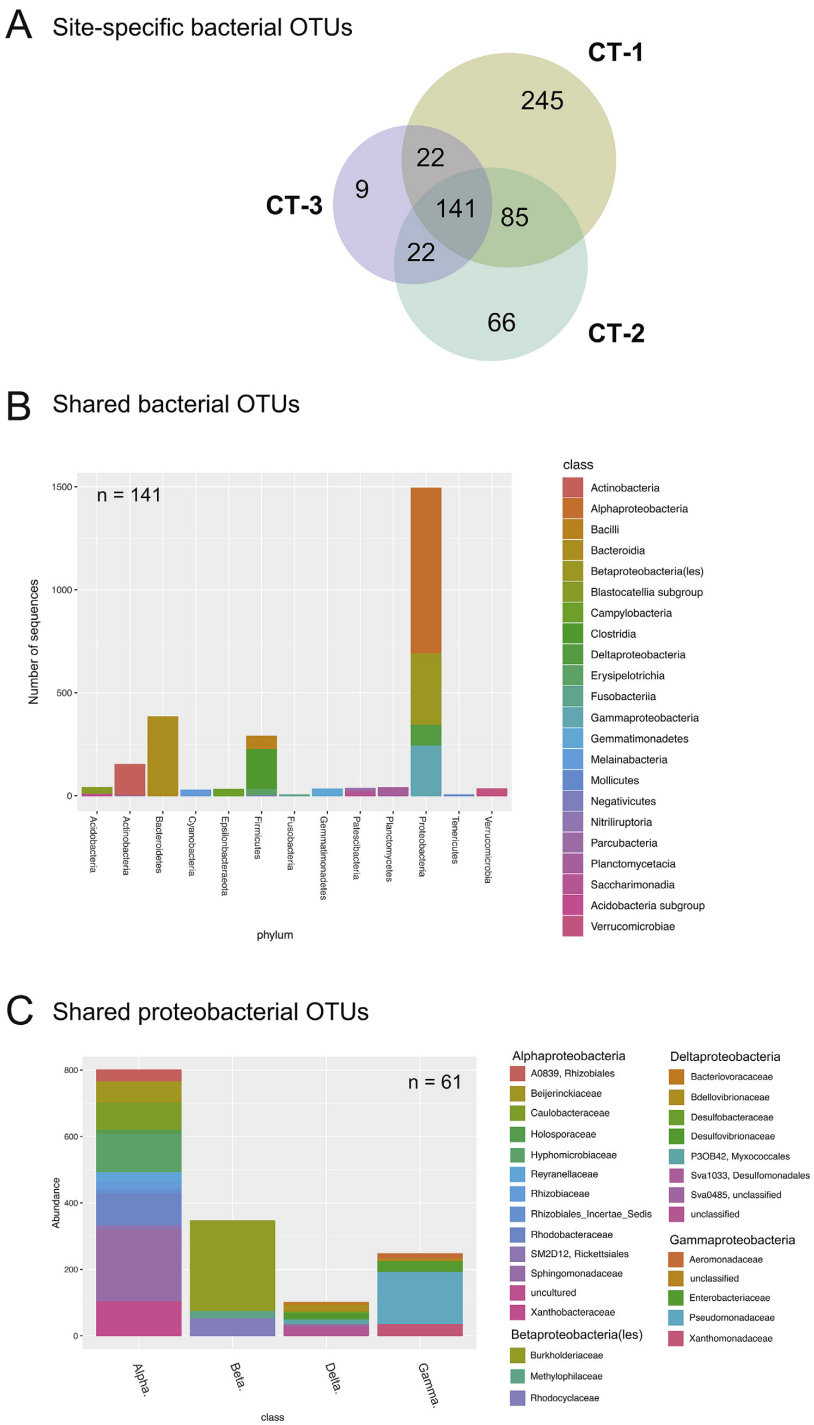


Fig. 6. Site-specific and shared bacterial OTUs in three cooling towers. (A) The Venn diagram represents the proportion of shared and site-specific bacterial OTUs at the three sampling sites. Reflecting its higher species richness, CT-1 showed the greatest number of site-specific bacterial OTUs. The number of OTUs shared by all three sampling sites clearly exceeds the number of OTUs shared individually by two cooling towers. **(B)** The bar chart shows the taxonomic classification of the shared OTUs (occurring at least once at each sampling site). The majority of the OTUs belongs to the Proteobacteria, Bacteroidetes, Actinobacteria, and Firmicutes. **(C)** A more detailed view of the shared Proteobacteria on the family level. The most abundant groups are the Bradyrhizobiaceae, Caulobacteraceae, Hyphomicrobiaceae, Sphingomonadaceae, Comamonadaceae, and Pseudomonadaceae.

and man-made systems to aerosols (Hilborn et al., 2006; Parker et al., 1983; Vaerewijck et al., 2005). They are well-known colonizers of water distribution systems, and due to their unique cell wall composition may be involved in surface colonization and biofilm formation.

Six *Pseudomonas* OTUs were detected. *P. stutzeri* was the most abundant OTU in CT-2 (28.4%) and was also found in the other two cooling towers. *P. toyotomiensis* (15.7%) represented the second most abundant OTU in CT-2, and was among the persistent shared OTUs (Table S3, Fig. 5). Additional *Pseudomonas* species were seen, but at much lower frequencies. Not considered to be strict pathogens, members of this genus are ubiquitous and metabolically versatile; in addition to their prominent role in biofilm formation some are known as opportunistic human pathogens and cause various infections in e.g. the urinary tract, soft tissues, or the respiratory tract (Goetz et al., 1983; Keys et al., 1983; Potvliege et al., 1987).

Taken together, we couldn't find any evidence for the presence of *sensu stricto* human pathogens in the cooling tower microbial communities analysed in this study, not even among the rare biosphere.

3.5. A highly dynamic protist community

We detected the presence of four groups of protists, the Rhizaria, the Excavata, the Alveolata, and the Amoebozoa at all three sampling sites. Although the top taxa in all cooling towers belonged mainly to the Alveolata (Ciliophora and Apicomplexa) and Amoebozoa (Lobosa), we observed pronounced differences in relative abundances between the cooling towers (Fig. 3, Table S4). The most abundant protist groups in cooling tower CT-1 were classified into diverse taxa in the Ciliophora, like *Ancistrum* (9.3%) and *Vorticella* (7.1%), and Lobosa taxa including *Vanella* (8.1%), *Cochliopodium* (7.8%), and *Vexillifera* species (7.7%). In contrast, cooling towers CT-2 and CT-3 were dominated by diverse gregarines. These apicomplexan parasites are often associated with small invertebrates. Although the source of the gregarines in our cooling tower water system remains unknown, they were likely introduced together with their potential animal hosts. Still, there were also several well-known free-living amoebae among the most abundant taxa, such as *Vexillifera*, *Stenamoeba*, *Vannella*, *Hartmannella*, *Echinamoebida*, *Cochliopodium*, *Acanthamoeba*, and *Allovahlkampfia* (Fig. 3, Table S4). Taken together, the protist community detected in the three cooling towers was dominated by taxa frequently found in diverse freshwater habitats and able to persist in an environment characterized by steady water movement.

The protist community composition was more variable than the bacterial microbiome. No persistent protist taxa were observed, and the most frequently recurring OTUs were detected in less than 50% of the samples (Table S4). These included *Stenamoeba, Ancistrum,* and *Vexillifera* (CT-1) and gregarines (CT-2). The repeated occurrence of gregarines (often associated with small invertebrates) suggests that in addition to bacteria-protists interactions, small invertebrates (though not analysed in this study) may also affect the structure of cooling tower microbiomes.

By far, the largest proportion of taxa was site-specific (Fig. 7A, Table S4). This included representatives of the genera *Stenamoeba*, *Vannella*, *Echinamoeba*, and *Cochliopodium* (Amoebozoa), *Tetrahymena*, *Chilodonella,* and gregarines (Alveolata), *Neobodo* and Vahlkampfiidae (Excavata), and *Paracercomonas* (Rhizaria). More than half of these site-specific OTUs do not occur at more than one timepoint, further indicating that the protist community in the water basin is more variable than the bacterial community. Possible reasons for this may include (i) varying protist community composition of the source water, (ii) competition between protists, or (iii) random shearing of protists associated with bacterial biofilms.

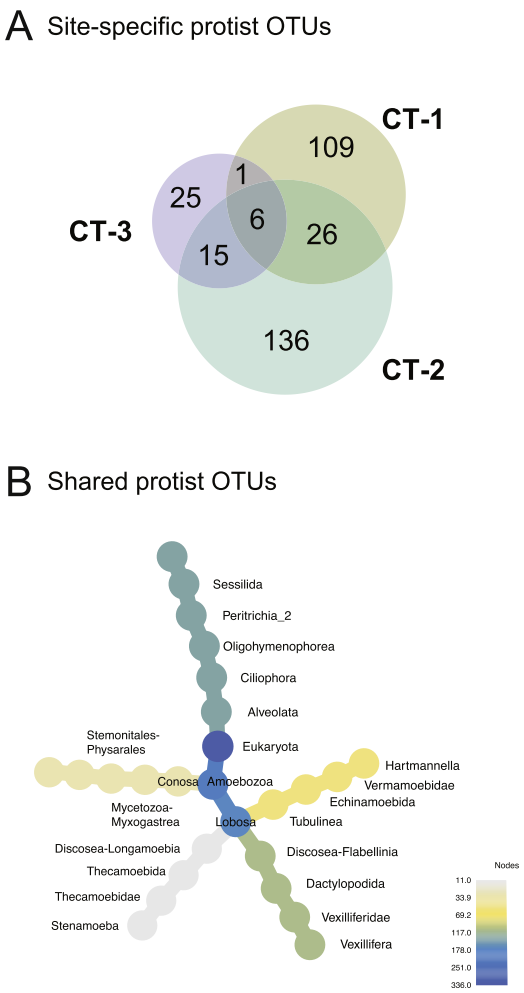


Fig. 7. Site-specific and shared protist OTUs in three cooling towers. The Venn diagram in **(A)** represents the proportion of shared and site-specific protist OTUs at the three sampling sites. In contrast to the less variable bacterial community composition, the majority of the protist OTUs are site-specific, and the number of OTUs shared by all three sampling sites constitutes only a very small fraction of all OTUs. **(B)** The heat tree indicates the taxonomic classification of the six shared OTUs, most of which are free-living amoebae (*Hartmannella*, *Vexillifera*, *Stenamoeba*). Unclassified taxa are unlabeled.

There were few OTUs found at all three cooling towers (6 corresponding to 1.7% of all protist OTUs). These included a member of the Sessilida, and free-living amoebae such as *Hartmannella*, *Stenamoeba*, *Vexillifera*, and a conosa member. (Fig. 7B). Taken together, the protist microbiome of cooling tower water is surprisingly diverse, with a species richness lower than the bacterial community but in the same order of magnitude. At the same time, it is highly dynamic and only showed a small core set of persistent and shared taxa.

3.6. Bacteria-protist interactions uncovered by co-occurrence networks

To gain insights into putative interactions between bacteria and

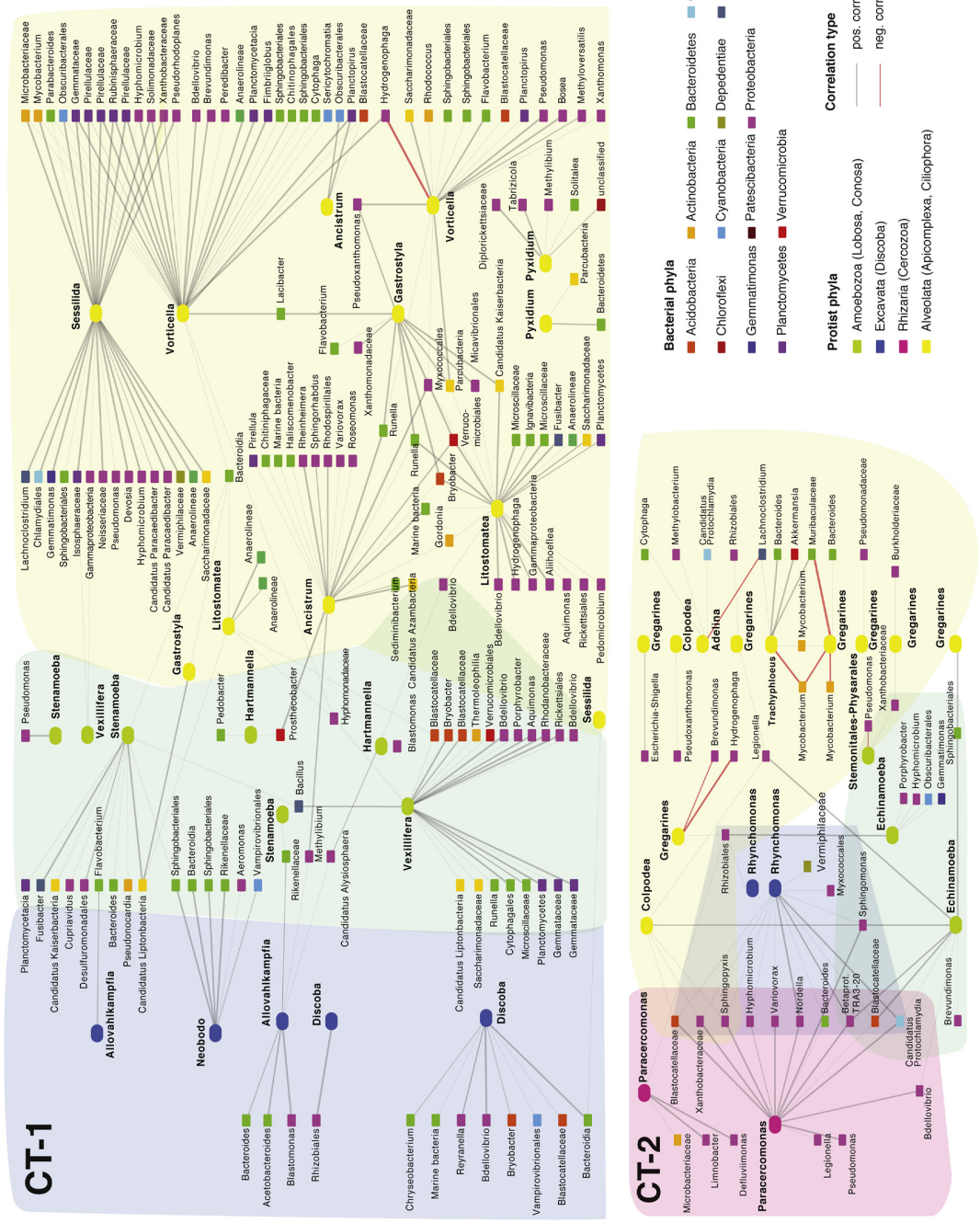


Fig. 8. Co-occurrence network indicating bacteria-protist relationships in cooling tower water microbiomes. Co-occurrence networks including only interactions between bacterial and protist OTUs are shown for each sampling location; positive correlations are indicated by grey connections (edges), negative correlations are indicated by red edges. The number of negative correlations might be underestimated due to the conservative thresholds applied in this study. The width of the edges negatively correlates to the p-value of each predicted interaction. In CT-1 the majority of the protist nodes are members of the Alveolata, whereas all four protist group are represented in the CT-2 network. This analysis revealed interactions involving known protist-associated bacteria, constituting 15.6% (CT-1) and 24% (CT-2) of all predicted relationships. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

protists in the cooling tower water microbiomes, we constructed and analysed co-occurrence networks for each cooling tower using all bacterial and protist OTUs, and a comprehensive approach for sequence abundance-based network inference (Faust and Raes, 2016). In our analysis, we only considered relationships between bacteria and protists (Fig. 8). The resulting co-occurrence networks were characterized by a high number of nodes (189 OTUs for CT-1, 61 for CT-2) and correlations (207 edges for CT-1, 78 for CT-2). Due to the low sample number, CT-3 was not included in this analysis. The networks are composed of mainly positive correlations, in which the abundances of bacterial and protist OTUs followed the same trend.

Protists in the CT-1 network are mainly ciliates (Alveolata; Oligohymenophorea), followed by naked amoebae (Amoebozoa; Discosea). Most protists were correlated with a number of bacterial taxa, with *Vorticella* and an unclassified Sessilida showing highest centrality (i.e. number of edges) identifying them as key players (Fig. 8). Each protist node had its own individual set of bacterial interaction partners. Some bacterial nodes were shared between different protists but the majority (73%) of bacterial nodes were linked to only one protist, suggesting a certain degree of specificity of bacteria-protist interactions. In comparison, the CT-2 network had fewer protist nodes with high centrality. However, we could detect more bacterial taxa, who co-occurred with more than one protist (40%). The most abundant bacterial OTUs included in both networks were members of the Proteobacteria, the Planctomycetes, and the Bacteroidetes.

Remarkably, 15.6% of all predicted bacteria-protist relationships in the CT-1 network and 24% in the CT-2 network were comprised of bacterial genera with members known to associate with protists (Fig. 8). This included several well-known bacterial symbionts of free-living amoebae or ciliates, such as members of the genera *Protochlamydia* (Collingro et al., 2005; Fokin, 2004; Horn, 2008; Vannini et al., 2004) and the family Rickettsiaceae (Horn et al., 1999). Bacterial symbionts of protists are ubiquitous in ciliates and amoeba and in some cases represent ancient relationships originating millions of years ago (Görtz, 2001; Horn, 2008; Molmeret et al., 2005). These bacteria generally show reduced metabolic capabilities and are thus dependent on their protist hosts (Klein et al., 2012; McCutcheon and Moran, 2012; Schmitz-Esser et al., 2010). Some act as defensive symbionts or confer resistance to environmental stress (Dziallas et al., 2012; Fujishima et al., 2005; Grosser et al., 2018; Hori and Fujishima, 2003), but the relevance for the protist host is often unclear.

The networks also included several bacterial groups transiently thriving in protists, such as members of the genera *Legionella, Mycobacterium*, *Pseudomonas, Nordella, Variovorax, Bosea, Acidovorax, Cupriavidus, Brevundimonas,* and *Flavobacterium* (Corsaro et al., 2010; Drancourt, 2014; Evstigneeva et al., 2009; Fields et al., 2002; Greub and Raoult, 2004; La Scola et al., 2004; Michel et al., 1995; Thomas et al., 2008; Zeybek and Binay, 2014). Interestingly, our co-occurrence analysis also predicted members of the groups Parcubacteria (OD1; Corsaro et al., 2010) and Depedentiae (TM6) as interaction partner of protists. This observation further supports a symbiotic lifestyle for Parcubacteria as suggested by (meta-)genomic analysis (Gong et al., 2014; Nelson and Stegen, 2015) and provides additional support for the endosymbiotic lifestyle of Depedentiae members (Delafont et al., 2015). Our analysis underscores the potential of protists in cooling tower water microbiomes as reservoirs for intracellular bacteria, including relatives of opportunistic pathogens (Greub and Raoult, 2004).

Of note, the protist-associated bacteria in our network generally show interactions with several different protist OTUs. This indicates that in a natural setting, bacterial symbionts are more promiscuous and thrive in a wide range of different protists. Among the taxa included in the networks, there are several protists recognized previously as hosts for intracellular bacteria: The amoeba *Vexillifera* is a known host of the environmental chlamydia '*Candidatus* Neptunochlamydia vexilliferae' (Pizzetti et al., 2016). Hartmannella (Vermamoeba) is the natural host of an intranuclear symbiont, '*Candidatus* Nucleicultrix amoebiphila' and the environmental chlamydia *Rubidus massiliensis* (Bou Khalil et al., 2016; Horn, 2008; Schulz et al., 2014). *Echinamoeba* is a known host for *Pseudomonas aeruginosa* (Michel et al., 1995). In addition, *Allovahlkampfia* which have been suggested to act as hosts for pathogenic bacteria (Mohamed and Huseein, 2016), were also represented in our networks. In this respect, we have recently isolated from the water sample of the tower CT-1 the first testate amoeba, *Cochliopodium minus,* containing a bacterial symbiont (Tsao et al., 2017). Taken together, our co-presence network analysis recovered a large number of known bacteria-protist interactions and suggested that such relationships are much more diverse than recognized currently.

4. Conclusions

- Bacterial and protist communities in cooling towers are broadly similar at the class level to those in natural freshwater and drinking water systems. Although community structure can be highly dynamic, the presence of core taxa suggests that cooling tower basins select for biofilm-forming and biofilm-associated microbes.
- Amplicon sequencing techniques that target ribosomal RNA genes is a useful tool to study microbial community dynamics in cooling tower water basins. Longer-term longitudinal studies including outbreak situations are required to better understand the role of the microbial community for propagation and spread of opportunistic bacterial pathogens that have been associated with disease outbreaks. The implementation of upcoming real-time sequencing technologies might facilitate online monitoring of cooling tower communities to predict biofilm formation and colonization with opportunistic pathogens.
- Co-occurrence network analysis is a helpful approach to predict bacteria-protists interactions using amplicon sequencing data. The high number of nodes and degree of connectedness in the networks obtained here highlights the importance of inter-kingdom relationships in cooling tower water systems.
- A design of cooling tower water basins and other measures that prevent biofilm formation should help to decrease the public health risk of cooling towers as a source for bacterial pathogens.

Author contributions

HFT, US, AI, JW, and MH conceived the study. HFT and US performed the sampling and DNA isolation; HFT performed PCR and prepared sequencing reactions. CH performed the demultiplexing, filtering and clustering of the reads. HFT performed the analysis of the data and prepared the figures; HFT and MH wrote the manuscript. All authors reviewed and edited the manuscript.

Declaration of interests

All authors declare that there is no conflict of interest.

Acknowledgement

This work was funded by Austrian Science Fund Project TRP209-B20. MH was supported by European Research Grant project EVOCHLAMY (ERC StG grant no. 281633). We thank Alexander Loy, Bela Hausmann, Claus Pelikan, Roey Angel, and Peter Hufnagl for their help and support.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.watres.2019.04.028.

References

Abd, H., Johansson, T., Golovliov, I., Sandström, G., Forsman, M., 2003. Survival and growth of *Francisella tularensis* in *Acanthamoeba castellanii*. Appl. Environ. Microbiol. 69, 600–606. https://doi.org/10.1128/AEM.69.1.600-606.2003.

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

Albertsen, M., Karst, S.M., Ziegler, A.S., Kirkegaard, R.H., Nielsen, P.H., 2015. Back to basics - the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. PLoS One 10. https://doi.org/10.1371/journal.pone.0132783.

Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., 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, e6372. https://doi.org/10.1371/journal.pone.0006372.

Barbaree, J.M., Fields, B.S., Feeley, J.C., Gorman, G.W., Martin, W.T., 1986. Isolation of protozoa from water associated with legionellosis outbreak and demonstration of intracellular multiplication of *Legionella pneumophila*. Appl. Environ. Microbiol. 51, 422–424.

Bereschenko, L.A., Stams, A.J.M., Euverink, G.J.W., van Loosdrecht, M.C.M., 2010. Biofilm formation on reverse osmosis membranes is initiated and dominated by *Sphingomonas* spp. Appl. Environ. Microbiol. 76, 2623–2632. https://doi.org/10.1128/AEM.01998-09.

Berk, S.G., Gunderson, J.H., Newsome, A.L., Farone, A.L., Hayes, B.J., Redding, K.S., Uddin, N., Williams, E.L., Johnson, R.A., Farsian, M., Reid, A., Skimmyhorn, J., Farone, M.B., 2006. Occurrence of infected amoebae in cooling towers compared with natural aquatic environments: implications for emerging pathogens. Environ. Sci. Technol. 40, 7440–7444. https://doi.org/10.1021/es0604257.

Berry, D., Xi, C., Raskin, L., 2006. Microbial ecology of drinking water distribution systems. Curr. Opin. Biotechnol. 17, 297–302. https://doi.org/10.1016/j.copbio.2006.05.007.

Bou Khalil, J.Y., Benamar, S., Baudoin, J.-P., Croce, O., Blanc-Tailleur, C., Pagnier, I., Raoult, D., La Scola, B., 2016. Developmental cycle and genome analysis of "*Rubidus massiliensis*," a new *Vermamoeba vermiformis* pathogen. Front. Cell. Infect. Microbiol. 6 https://doi.org/10.3389/fcimb.2016.00031.

Boucher, D., Jardillier, L., Debroas, D., 2006. Succession of bacterial community composition over two consecutive years in two aquatic systems: a natural lake and a lake-reservoir. FEMS Microbiol. Ecol. 55, 79–97. https://doi.org/10.1111/j.1574-6941.2005.00011.x.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., Wilkins, M.J., Wrighton, K.C., Williams, K.H., Banfield, J.F., 2015. Unusual biology across a group comprising more than 15% of domain Bacteria. Nature 523, 208–211. https://doi.org/10.1038/nature14486.

Buse, H.Y., Lu, J., Lu, X., Mou, X., Ashbolt, N.J., 2014. Microbial diversities (16S and 18S rRNA gene pyrosequencing) and environmental pathogens within drinking water biofilms grown on the common premise plumbing materials unplasticized polyvinylchloride and copper. FEMS Microbiol. Ecol. 88, 280–295. https://doi.org/10.1111/1574-6941.12294.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. https://doi.org/10.1038/nmeth.f.303.

Ceyhan, N., Ozdemir, G., 2008. Extracellular polysaccharides produced by cooling water tower biofilm bacteria and their possible degradation. Biofouling 24, 129–135. https://doi.org/10.1080/08927010801911316.

Chao, Y., Mao, Y., Wang, Z., Zhang, T., 2015. Diversity and functions of bacterial community in drinking water biofilms revealed by high-throughput sequencing. Sci. Rep. 5 https://doi.org/10.1038/srep10044.

Cirillo, J.D., Falkow, S., Tompkins, L.S., 1994. Growth of *Legionella pneumophila* in *Acanthamoeba castellanii* enhances invasion. Infect. Immun. 62, 3254–3261. https://doi.org/10.1016/j.ijheh.2010.03.002.

Cirillo, J.D., Falkow, S., Tompkins, L.S., Bermudez, L.E., 1997. Interaction of *Mycobacterium avium* with environmental amoebae enhances virulence. Infect. Immun. 65, 3759–3767.

Collingro, A., Toenshoff, E.R., Taylor, M.W., Fritsche, T.R., Wagner, M., Horn, M., 2005. "*Candidatus* Protochlamydia amoebophila", an endosymbiont of *Acanthamoeba* spp. Int. J. Syst. Evol. Microbiol. 55, 1863–1866. https://doi.org/10.1099/ijs.0.63572-0.

Corsaro, D., Pages, G.S., Catalan, V., Loret, J.-F., Greub, G., 2010. Biodiversity of amoebae and amoeba-associated bacteria in water treatment plants. Int. J. Hyg Environ. Health 213, 158–166. https://doi.org/10.1016/j.ijheh.2010.03.002.

Costerton, J.W., Stewart, P.S., Greenberg, E.P., 1999. Bacterial biofilms: a common cause of persistent infections. Science 80. https://doi.org/10.1126/science.284.5418.1318.

Critchley, M., Bentham, R., 2009. The efficacy of biocides and other chemical additives in cooling water systems in the control of amoebae. J. Appl. Microbiol. 106, 784–789. https://doi.org/10.1111/j.1365-2672.2008.04044.x.

Crump, B.C., Hobbie, J.E., 2005. Synchrony and seasonality in bacterioplankton communities of two temperate rivers. Limnol. Oceanogr. 50, 1718–1729. https://doi.org/10.4319/lo.2005.50.6.1718.

Delafont, V., Mougari, F., Cambau, E., Joyeux, M., Bouchon, D., Héchard, Y., Moulin, L., 2014. First evidence of amoebae-mycobacteria association in drinking water network. Environ. Sci. Technol. 48, 11872–11882. https://doi.org/10.1021/es5036255.

Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L., Héchard, Y., 2015. Shedding light on microbial dark matter: a TM6 bacterium as natural endosymbiont of a free-living amoeba. Environ. Microbiol. Rep. 7, 970–978. https://doi.org/10.1111/1758-2229.12343.

Delafont, V., Bouchon, D., Héchard, Y., Moulin, L., 2016. Environmental factors shaping cultured free-living amoebae and their associated bacterial community within drinking water network. Water Res. 100, 382–392. https://doi.org/10.1016/j.watres.2016.05.044.

Douterelo, I., Sharpe, R., Boxall, J., 2014. Bacterial community dynamics during the early stages of biofilm formation in a chlorinated experimental drinking water distribution system: implications for drinking water discolouration. J. Appl. Microbiol. 117, 286–301. https://doi.org/10.1111/jam.12516.

Drancourt, M., 2014. Looking in amoebae as a source of mycobacteria. Microb. Pathog. 77, 119–124. https://doi.org/10.1016/j.micpath.2014.07.001.

Dziallas, C., Allgaier, M., Monaghan, M.T., Grossart, H.-P., 2012. Act together—implications of symbioses in aquatic ciliates. Front. Microbiol. 3 https://doi.org/10.3389/fmicb.2012.00288.

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998. https://doi.org/10.1038/nmeth.2604.

Evstigneeva, A., Raoult, D., Karpachevskiy, L., La Scola, B., 2009. Amoeba co-culture of soil specimens recovered 33 different bacteria, including four new species and *Streptococcus pneumoniae*. Microbiology 155, 657–664. https://doi.org/10.1099/mic.0.022970-0.

Farenhorst, A., Li, R., Jahan, M., Tun, H.M., Mi, R., Amarakoon, I., Kumar, A., Khafipour, E., 2017. Bacteria in drinking water sources of a First Nation reserve in Canada. Sci. Total Environ. 575, 813–819. https://doi.org/10.1016/j.scitotenv.2016.09.138.

Faust, K., Raes, J., 2016. CoNet app: inference of biological association networks using Cytoscape. F1000Research 5, 1519. https://doi.org/10.12688/f1000research.9050.1.

Fields, B.S., Benson, R.F., Besser, R.E., 2002. *Legionella* and Legionnaires' disease: 25 years of investigation. Clin. Microbiol. Rev. 15, 506–526. https://doi.org/10.1128/cmr.15.3.506-526.2002.

Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S.A., Kjelleberg, S., 2016. Biofilms: an emergent form of bacterial life. Nat. Rev. Microbiol. 14, 563–575. https://doi.org/10.1038/nrmicro.2016.94.

Fokin, S.I., 2004. Bacterial endocytobionts of *Ciliophora* and their interactions with the host cell. Int. Rev. Cytol. https://doi.org/10.1016/S0074-7696(04)36005-5.

Foster, Z.S.L., Sharpton, T.J., Grünwald, N.J., 2017. Metacoder: an R package for visualization and manipulation of community taxonomic diversity data. PLoS Comput. Biol. 13, e1005404 https://doi.org/10.1371/journal.pcbi.1005404.

Fouque, E., Trouilhé, M.C., Thomas, V., Hartemann, P., Rodier, M.H., Hécharda, Y., 2012. Cellular, biochemical, and molecular changes during encystment of free-living amoebae. Eukaryot. Cell 11, 382–387. https://doi.org/10.1128/EC.05301-11.

Fritsche, T.R., Sobek, D., Gautom, R.K., 1998. Enhancement of in vitro cytopathogenicity by *Acanthamoeba* spp. following acquisition of bacterial endosymbionts. FEMS Microbiol. Lett. 166, 231–236. https://doi.org/10.1111/j.1574-6968.1998.tb13895.x.

Fujishima, M., Kawai, M., Yamamoto, R., 2005. *Paramecium caudatum* acquires heat-shock resistance in ciliary movement by infection with the endonuclear symbiotic bacterium *Holospora obtusa*. FEMS Microbiol. Lett. 243, 101–105. https://doi.org/10.1016/j.femsle.2004.11.053.

Garcia, A., Goñi, P., Cieloszyk, J., Fernandez, M.T., Calvo-Beguería, L., Rubio, E., Fillat, M.F., Peleato, M.L., Clavel, A., 2013. Identification of free-living amoebae and amoeba-associated bacteria from reservoirs and water treatment plants by molecular techniques. Environ. Sci. Technol. 47, 3132–3140. https://doi.org/10.1021/es400160k.

Gilbert, J.A., Jansson, J.K., Knight, R., 2014. The Earth Microbiome project: successes and aspirations. BMC Biol. 12, 69. https://doi.org/10.1186/s12915-014-0069-1.

Goertz, H.D., 2001. Intracellular bacteria in ciliates. Int. Microbiol. 4, 143–150. https://doi.org/10.1007/s10123-001-0029-9.

Goetz, A., Yu, V.L., Hanchett, J.E., Rihs, J.D., 1983. *Pseudomonas stutzeri* bacteremia associated with Hemodialysis. Arch. Intern. Med. 143, 1909–1912. https://doi.org/10.1001/archinte.1983.00350100073018.

Gong, J., Qing, Y., Guo, X., Warren, A., 2014. "*Candidatus* Sonnebornia yantaiensis", a member of candidate division OD1, as intracellular bacteria of the ciliated protist *Paramecium bursaria* (Ciliophora, Oligohymenophorea). Syst. Appl. Microbiol. 37, 35–41. https://doi.org/10.1016/j.syapm.2013.08.007.

Greub, G., Raoult, D., 2004. Microorganisms resistant to free-living amoebae. Clin. Microbiol. Rev. 17, 413–433. https://doi.org/10.1128/CMR.17.2.413-433.2004.

Grosser, K., Ramasamy, P., Amirabad, A.D., Schulz, M.H., Gasparoni, G., Simon, M., Schrallhammer, M., 2018. More than the "killer trait": infection with the bacterial endosymbiont *Caedibacter taeniospiralis* causes transcriptomic modulation in paramecium host. Genome Biol. Evol. 10, 646–656. https://doi.org/10.1093/gbe/evy024.

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G., De Vargas, C., Decelle, J., Del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Holzmann, M., Kooistra, W.H.C.F., Lara, E., Le Bescot, N., Logares, R., Mahé, F., Massana, R., Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R., Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference database (PR 2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res. 41, D597–D604. https://doi.org/10.1093/nar/gks1160.

Hamilton, K.A., Hamilton, M.T., Johnson, W., Jjemba, P., Bukhari, Z., LeChevallier, M., Haas, C.N., 2018. Health risks from exposure to *Legionella* in reclaimed water aerosols: toilet flushing, spray irrigation, and cooling towers. Water Res. 134, 261–279. https://doi.org/10.1016/j.watres.2017.12.022.

He, X., McLean, J.S., Edlund, A., Yooseph, S., Hall, A.P., Liu, S.-Y., Dorrestein, P.C., Esquenazi, E., Hunter, R.C., Cheng, G., Nelson, K.E., Lux, R., Shi, W., 2014. Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. Proc. Natl. Acad. Sci. Unit. States Am. 112, 244–249. https://doi.org/10.1073/pnas.1419038112.

Herbold, C.W., Pelikan, C., Kuzyk, O., Hausmann, B., Angel, R., Berry, D., Loy, A., 2015. A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. Front. Microbiol. 6 https://doi.org/10.3389/fmicb.2015.00731.

Herlemann, D.P.R., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J. 5, 1571–1579. https://doi.org/10.1038/ismej.2011.41.

Hess, S., 2017. Description of *Hyalodiscus flabellus* sp. nov. (Vampyrellida, Rhizaria) and identification of its bacterial endosymbiont, "*Candidatus* megaira polyxenophila" (Rickettsiales, Alphaproteobacteria). Protist 168, 109–133. https://doi.org/10.1016/j.protis.2016.11.003.

Hess, S., Suthaus, A., Melkonian, M., 2016. "*Candidatus* Finniella" (Rickettsiales, Alphaproteobacteria), novel endosymbionts of viridiraptorid amoeboflagellates (Cercozoa, rhizaria). Appl. Environ. Microbiol. 82, 659–670. https://doi.org/10.1128/AEM.02680-15.

Hilborn, E.D., Covert, T.C., Yakrus, M.A., Harris, S.I., Donnelly, S.F., Rice, E.W., Toney, S., Bailey, S.A., Stelma, G.N., 2006. Persistence of nontuberculous *Mycobacteria* in a drinking water system after addition of filtration treatment. Appl. Environ. Microbiol. 72, 5864–5869. https://doi.org/10.1128/AEM.00759-06.

Hong, P.Y., Hwang, C., Ling, F., Andersen, G.L., LeChevallier, M.W., Liu, W.T., 2010. Pyrosequencing analysis of bacterial biofilm communities in water meters of a drinking water distribution system. Appl. Environ. Microbiol. 76, 5631–5635. https://doi.org/10.1128/AEM.00281-10.

Hori, M., Fujishima, M., 2003. The endosymbiotic bacterium *Holospora obtusa* enhances heat-shock gene expression of the host *Paramecium caudatum*. J. Eukaryot. Microbiol. 50, 293–298. https://doi.org/10.1111/j.1550-7408.2003.tb00137.x.

Horn, M., 2008. Chlamydiae as symbionts in eukaryotes. Annu. Rev. Microbiol. 62, 113–131. https://doi.org/10.1146/annurev.micro.62.081307.162818.

Horn, M., Fritsche, T.R., Gautom, R.K., Schleifer, K.H., Wagner, M., 1999. Novel bacterial endosymbionts of *Acanthamoeba* spp. related to the *Paramecium caudatum* symbiont *Caedibacter caryophilus*. Environ. Microbiol. 1, 357–367. https://doi.org/10.1046/j.1462-2920.1999.00045.x.

Inoue, H., Fujimura, R., Agata, K., Ohta, H., 2015. Molecular Characterization of viable *Legionella* spp. in cooling tower water samples by combined use of ethidium monoazide and PCR. Microbes Environ. Environ 30, 108–112. https://doi.org/10.1264/jsme2.me14115.

Jensen, P.Ø., Givskov, M., Bjarnsholt, T., Moser, C., 2010. The immune system vs. *Pseudomonas aeruginosa* biofilms. FEMS Immunol. Med. Microbiol. 59, 292–305. https://doi.org/10.1111/j.1574-695X.2010.00706.x.

Kalmbach, S., Manz, W., Szewzyk, U., 1997. Isolation of new bacterial species from drinking water biofilms and proof of their in situ dominance with highly specific 16S rRNA probes. Appl. Environ. Microbiol. 63, 4164–4170.

Kalyuzhnaya, M.G., De Marco, P., Bowerman, S., Pacheco, C.C., Lara, J.C., Lidstrom, M.E., Chistoserdova, L., 2006. *Methyloversatilis universalis* gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates. Int. J. Syst. Evol. Microbiol. 56, 2517–2522. https://doi.org/10.1099/ijs.0.64422-0.

Katharios-Lanwermeyer, S., Xi, C., Jakubovics, N.S., Rickard, A.H., 2014. Mini-review: microbial coaggregation: ubiquity and implications for biofilm development. Biofouling 30, 1235–1251. https://doi.org/10.1080/08927014.2014.976206.

Keys, T.F., Melton, L.J., Maker, M.D., Ilstrup, D.M., 1983. A suspected hospital outbreak of pseudobacteremia due to *Pseudomonas stutzeri*. J. Infect. Dis. 147, 489–493. https://doi.org/10.1093/infdis/147.3.489.

Klein, C.C., Cottret, L., Kielbassa, J., Charles, H., Gautier, C., Ribeiro de Vasconcelos, A.T., Lacroix, V., Sagot, M.F., 2012. Exploration of the core metabolism of symbiotic bacteria. BMC Genomics 13, 438. https://doi.org/10.1186/1471-2164-13-438.

Kurtz, J.B., Bartlett, C.L.R., Newton, U.A., White, R.A., Jones, N.L., 1982. *Legionella pneumophila* in cooling water systems: report of a survey of cooling towers in London and a pilot trial of selected biocides. J. Hyg. 88, 369–381. https://doi.org/10.1017/S0022172400070248.

La Scola, B., Raoult, D., 2001. Survival of *Coxiella burnetii* within free-living amoeba *Acanthamoeba castellanii*. Clin. Microbiol. Infect. 7, 75–79. https://doi.org/10.1046/j.1469-0691.2001.00193.x.

La Scola, B., Barrassi, L., Raoult, D., 2004. A novel alpha-Proteobacterium, *Nordella oligomobilis* gen. nov., sp. nov., isolated by using amoebal co-cultures. Res. Microbiol. 155, 47–51. https://doi.org/10.1016/j.resmic.2003.09.012.

Li, L., Qin, T., Li, Y., Zhou, H., Song, H., Ren, H., Li, L., Li, Y., Zhao, D., 2015. Prevalence and molecular characteristics of waterborne pathogen *Legionella* in industrial cooling tower environments. Int. J. Environ. Res. Public Health 12, 12605–12617. https://doi.org/10.3390/ijerph121012605.

Liu, Y., Zhang, W., Sileika, T., Warta, R., Cianciotto, N.P., Packman, A., 2009. Role of bacterial adhesion in the microbial ecology of biofilms in cooling tower systems. Biofouling 25, 241–253. https://doi.org/10.1080/08927010802713414.

Liu, Y., Zhang, W., Sileika, T., Warta, R., Cianciotto, N.P., Packman, A.I., 2011. Disinfection of bacterial biofilms in pilot-scale cooling tower systems. Biofouling 27, 393–402. https://doi.org/10.1080/08927014.2011.577525.

Liu, R., Yu, Z., Guo, H., Liu, M., Zhang, H., Yang, M., 2012. Pyrosequencing analysis of eukaryotic and bacterial communities in faucet biofilms. Sci. Total Environ. 435–436, 124–131. https://doi.org/10.1016/j.scitotenv.2012.07.022.

Llewellyn, A.C., Lucas, C.E., Roberts, S.E., Brown, E.W., Nayak, B.S., Raphael, B.H., Winchell, J.M., 2017. Distribution of *Legionella* and bacterial community composition among regionally diverse US cooling towers. PLoS One 12, e0189937. https://doi.org/10.1371/journal.pone.0189937.

Lozupone, C., Knight, R., 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. 71, 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005.

Luef, B., Frischkorn, K.R., Wrighton, K.C., Holman, H.-Y.N., Birarda, G., Thomas, B.C., Singh, A., Williams, K.H., Siegerist, C.E., Tringe, S.G., Downing, K.H., Comolli, L.R., Banfield, J.F., 2015. Diverse uncultivated ultra-small bacterial cells in groundwater. Nat. Commun 6, 6372. https://doi.org/10.1038/ncomms7372.

Marrie, T.J., Raoult, D., La Scola, B., Birtles, R.J., De Carolis, E., Duperval, R., Field, S., Louie, T., Houston, S., Gribble, M., Williams, K., Nicolle, L., Grossman, R., Salit, I., Saginur, R., Gregson, D., Laverdiere, M., Joly, J., Marrie, T., Hutchinson, J., 2001. *Legionella*-like and other amoebal pathogens as agents of community-acquired pneumonia. Emerg. Infect. Dis. 7, 1026–1029. https://doi.org/10.3201/eid0706.010619.

Martiny, A.C., Albrechtsen, H.J., Arvin, E., Molin, S., 2005. Identification of bacteria in biofilm and bulk water samples from a nonchlorinated model drinking water distribution system: detection of a large nitrite-oxidizing population associated with *Nitrospira* spp. Appl. Environ. Microbiol. 71, 8611–8617. https://doi.org/10.1128/AEM.71.12.8611-8617.2005.

Matz, C., McDougald, D., Moreno, A.M., Yung, P.Y., Yildiz, F.H., Kjelleberg, S., 2005. Biofilm formation and phenotypic variation enhance predation-driven persistence of *Vibrio cholerae*. Proc. Natl. Acad. Sci. Unit. States Am. 102, 16819–16824. https://doi.org/10.1073/pnas.0505350102.

McCoy, S.T., VanBriesen, J.M., 2012. Temporal variability of bacterial diversity in a chlorinated drinking water distribution system. J. Environ. Eng. 138, 786–795. https://doi.org/10.1061/(asce)ee.1943-7870.0000539.

McCutcheon, J.P., Moran, N.A., 2012. Extreme genome reduction in symbiotic bacteria. Nat. Rev. Microbiol. 10, 13–26. https://doi.org/10.1038/nrmicro2670.

McMurdie, P.J., Holmes, S., 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome Census data. PLoS One 8, e61217. https://doi.org/10.1371/journal.pone.0061217.

Michel, R., Burghardt, H., Bergmann, H., 1995. Acanthamoeba, naturally intracellularly infected with *Pseudomonas aeruginosa*, after their isolation from a microbiologically contaminated drinking water system in a hospital. Zentralbl. Hyg. Umweltmed. 196, 532–544.

Miltner, E.C., Bermudez, L.E., 2000. *Mycobacterium avium* grown in *Acanthamoeba castellanii* is protected from the effects of antimicrobials. Antimicrob. Agents Chemother. 44, 1990–1994. https://doi.org/10.1128/AAC.44.7.1990-1994.2000.

Mohamed, M.E., Huseein, E.A., 2016. *Allovahlkampfia spelaea* is a potential environmental host for pathogenic bacteria. J. Bacteriol. Parasitol. 07 https://doi.org/10.4172/2155-9597.1000255.

Molmeret, M., Horn, M., Wagner, M., Santic, M., Abu Kwaik, Y., 2005. Amoebae as training grounds for intracellular bacterial pathogens. Appl. Environ. Microbiol. 71, 20–28. https://doi.org/10.1128/AEM.71.1.20-28.2005.

Muchesa, P., Leifels, M., Jurzik, L., Hoorzook, K.B., Barnard, T.G., Bartie, C., 2017. Coexistence of free-living amoebae and bacteria in selected South African hospital water distribution systems. Parasitol. Res. 116, 155–165. https://doi.org/10.1007/s00436-016-5271-3.

Mueller-Spitz, S.R., Goetz, G.W., McLellan, S.L., 2009. Temporal and spatial variability in nearshore bacterioplankton communities of Lake Michigan. FEMS Microbiol. Ecol. 67, 511–522. https://doi.org/10.1111/j.1574-6941.2008.00639.x.

Nelson, W.C., Stegen, J.C., 2015. The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. Front. Microbiol. 6 https://doi.org/10.3389/fmicb.2015.00713.

Newsome, A.L., Scott, T.M., Benson, R.F., Fields, B.S., 1988. Isolation of an Amoeba naturally harboring a distinctive *Legionella* species. Appl. Environ. Microbiol. 1688–1693.
Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., Bertilsson, S., 2011. A guide to the natural history of freshwater lake bacteria. Microbiol. Mol. Biol. Rev. 75, 14–49. https://doi.org/10.1128/MMBR.00028-10.
Nhu Nguyen, T.M., Ilef, D., Jarraud, S., Rouil, L., Campese, C., Che, D., Haeghebaert, S., Ganiayre, F., Marcel, F., Etienne, J., Desenclos, J., 2005. A community-wide outbreak of legionnaires disease linked to industrial cooling towers—How far can contaminated aerosols spread? J. Infect. Dis. 193, 102–111. https://doi.org/10.1086/498575.
Pagnier, I., Merchat, M., La Scola, B., 2009. Potentially pathogenic amoeba-associated microorganisms in cooling towers and their control. Future Microbiol. 4, 615–629. https://doi.org/10.2217/fmb.09.25.
Parker, B.C., Ford, M.A., Gruft, H., Falkinham, J.O., 1983. Epidemiology of infection by nontuberculous *mycobacteria*. IV. Preferential Aerosolization of *Mycobacterium intracellulare* from natural waters. Am. Rev. Respir. Dis. 128, 652–656. https://doi.org/10.1164/arrd.1980.122.2.259.
Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., Hugenholtz, P., 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat. Biotechnol 36, 996–1004. https://doi.org/10.1038/nbt.4229.
Parry, J.D., 2004. Protozoan grazing of freshwater biofilms. Adv. Appl. Microbiol. https://doi.org/10.1016/S0065-2164(04)54007-8.
Pereira, R.P.A., Peplies, J., Höfle, M.G., Brettar, I., 2017. Bacterial community dynamics in a cooling tower with emphasis on pathogenic bacteria and *Legionella* species using universal and genus-specific deep sequencing. Water Res. 122, 363–376. https://doi.org/10.1016/j.watres.2017.06.011.
Pizzetti, I., Schulz, F., Tyml, T., Fuchs, B.M., Amann, R., Horn, M., Fazi, S., 2016. Chlamydial seasonal dynamics and isolation of "*Candidatus* Neptunochlamydia vexilliferae" from a Tyrrhenian coastal lake. Environ. Microbiol. 18, 2405–2417. https://doi.org/10.1111/1462-2920.13111.
Potvliege, C., Jonckheer, J., Lenclud, C., Hansen, W., 1987. *Pseudomonas stutzeri* pneumonia and septicemia in a patient with multiple myeloma. J. Clin. Microbiol. 25, 458–459.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41, 590–596. https://doi.org/10.1093/nar/gks1219.
Revetta, R.P., Pemberton, A., Lamendella, R., Iker, B., Santo Domingo, J.W., 2010. Identification of bacterial populations in drinking water using 16S rRNA-based sequence analyses. Water Res. 44, 1353–1360. https://doi.org/10.1016/j.watres.2009.11.008.
Rickard, A.H., Leach, S.A., Hall, L.S., Buswell, C.M., High, N.J., Handley, P.S., 2002. Phylogenetic relationships and coaggregation ability of freshwater biofilm bacteria. Appl. Environ. Microbiol. 68, 3644–3650. https://doi.org/10.1128/AEM.68.7.3644-3650.2002.
Rickard, A.H., Gilbert, P., Handley, P.S., 2004. Influence of growth environment on coaggregation between freshwater biofilm bacteria. J. Appl. Microbiol. 96, 1367–1373. https://doi.org/10.1111/j.1365-2672.2004.02297.x.
Rożej, A., Cydzik-Kwiatkowska, A., Kowalska, B., Kowalski, D., 2015. Structure and microbial diversity of biofilms on different pipe materials of a model drinking water distribution systems. World J. Microbiol. Biotechnol. 31, 37–47. https://doi.org/10.1007/s11274-014-1761-6.
Sandström, G., Saeed, A., Abd, H., 2010. Acanthamoeba polyphaga is a possible host for Vibrio cholerae in aquatic environments. Exp. Parasitol 126, 65–68. https://doi.org/10.1016/j.exppara.2009.09.021.
Sauer, K., Camper, A.K., Ehrlich, G.D., Costerton, J.W., Davies, D.G., 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. J. Bacteriol. 184, 1140–1154. https://doi.org/10.1128/jb.184.4.1140-1154.2002.
Scheikl, U., Tsao, H.F., Horn, M., Indra, A., Walochnik, J., 2016. Free-living amoebae and their associated bacteria in Austrian cooling towers: a 1-year routine screening. Parasitol. Res. 115, 3365–3374. https://doi.org/10.1007/s00436-016-5097-z.
Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541. https://doi.org/10.1128/AEM.01541-09.
Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., Horn, M., 2010. The genome of the amoeba symbiont "Candidatus Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. J. Bacteriol 192, 1045–1057.
Schulz, F., Lagkouvardos, I., Wascher, F., Aistleitner, K., Kostanjšek, R., Horn, M., 2014. Life in an unusual intracellular niche: a bacterial symbiont infecting the nucleus of amoebae. ISME J. 8, 1634–1644. https://doi.org/10.1038/ismej.2014.5.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504. https://doi.org/10.1101/gr.1239303.
Shpigel, N.Y., Pasternak, Z., Factor, G., Gottlieb, Y., 2015. Diversity of bacterial biofilm communities on sprinklers from dairy farm cooling systems in Israel. PLoS One 10, e0139111. https://doi.org/10.1371/journal.pone.0139111.
Singh, R., Williams, H.N., Spitznagel, J.K., Smith, D.L., Stine, O.C., Labib, M.E., 2003. Microbial diversity of biofilms in dental unit water systems. Appl. Environ. Microbiol. 69, 3412–3420. https://doi.org/10.1128/aem.69.6.3412-3420.2003.
Soto-Giron, M.J., Rodriguez-R, L.M., Luo, C., Elk, M., Ryu, H., Hoelle, J., Santo Domingo, J.W., Konstantinidis, K.T., 2016. Biofilms on hospital shower Hoses: Characterization and implications for nosocomial infections. Appl. Environ. Microbiol. 82, 2872–2883. https://doi.org/10.1128/aem.03529-15.
States, S.J., Conley, L.F., Towner, S.G., Wolford, R.S., Stephenson, T.E., McNamara, A.M., Wadowsky, R.M., Yee, R.B., 1987. An alkaline approach to treating cooling towers for control of *Legionella pneumophila*. Appl. Environ. Microbiol. 53, 1775–1779 .
Tang, X., Xie, G., Shao, K., Dai, J., Chen, Y., Xu, Q., Gao, G., 2015. Bacterial community composition in oligosaline lake bosten: low overlap of *betaproteobacteria* and *bacteroidetes* with freshwater ecosystems. Microb. Environ. 30, 180–188. https://doi.org/10.1264/jsme2.me14177.
Taylor-Brown, A., Vaughan, L., Greub, G., Timms, P., Polkinghorne, A., 2014. Twenty years of research into Chlamydia-like organisms: a revolution in our understanding of the biology and pathogenicity of members of the phylum Chlamydiae. Pathog. Dis. 73, 1–15. https://doi.org/10.1093/femspd/ftu009.
Thom, S., Warhurst, D., Drasar, B.S., 1992. Association of Vibrio cholerae with fresh water amoebae. J. Med. Microbiol. 36, 303–306. https://doi.org/10.1099/00222615-36-5-303.
Thomas, V., McDonnell, G., 2007. Relationship between mycobacteria and amoebae: ecological and epidemiological concerns. Lett. Appl. Microbiol. 45, 349–357. https://doi.org/10.1111/j.1472-765x.2007.02206.x.
Thomas, V., Loret, J.-F., Jousset, M., Greub, G., 2008. Biodiversity of amoebae and amoebae-resisting bacteria in a drinking water treatment plant. Environ. Microbiol. 10, 2728–2745. https://doi.org/10.1111/j.1462-2920.2008.01693.x.
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., Kosciolek, 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., 2017. A communal catalogue reveals Earth's multiscale microbial diversity. Nature 551, 457–463. https://doi.org/10.1038/nature24621.
Tokajian, S.T., Hashwa, F.A., Hancock, I.C., Zalloua, P.A., 2005. Phylogenetic assessment of heterotrophic bacteria from a water distribution system using 16S rDNA sequencing. Can. J. Microbiol. 51, 325–335. https://doi.org/10.1139/w05-007.
Torvinen, E., Suomalainen, S., Paulin, L., Kusnetsov, J., 2013. Mycobacteria in Finnish cooling tower waters. APMIS 122, 353–358. https://doi.org/10.1111/apm.12153.
Tsao, H.-F., Scheikl, U., Volland, J.-M., Köhsler, M., Bright, M., Walochnik, J., Horn, M., 2017. '*Candidatus* Cochliophilus cryoturris' (Coxiellaceae), a symbiont of the testate amoeba *Cochliopodium minus*. Sci. Rep. 7 https://doi.org/10.1038/s41598-017-03642-8.
Türetgen, I., Cotuk, A., 2007. Monitoring of biofilm-associated *Legionella pneumophila* on different Substrata in model cooling tower system. Environ. Monit. Assess. 125, 271–279. https://doi.org/10.1007/s10661-006-9519-8.
Ulleryd, P., Hugosson, A., Allestam, G., Bernander, S., Claesson, B.E., Eilertz, I., Hagaeus, A.-C., Hjorth, M., Johansson, A., de Jong, B., Lindqvist, A., Nolskog, P., Svensson, N., 2012. Legionnaires' disease from a cooling tower in a community outbreak in Lidköping, Sweden- epidemiological, environmental and microbiological investigation supported by meteorological modelling. BMC Infect. Dis. 12, 313. https://doi.org/10.1186/1471-2334-12-313.
Vaerewijck, M.J.M., Huys, G., Palomino, J.C., Swings, J., Portaels, F., 2005. *Mycobacteria* in drinking water distribution systems: ecology and significance for human health. FEMS Microbiol. Rev. 29, 911–934. https://doi.org/10.1016/j.femsre.2005.02.001.
Van der Henst, C., Scrignari, T., Maclachlan, C., Blokesch, M., 2016. An intracellular replication niche for *Vibrio cholerae* in the amoeba *Acanthamoeba castellanii*. ISME J. 10, 897–910. https://doi.org/10.1038/ismej.2015.165.
van der Kooij, D., Bakker, G.L., Italiaander, R., Veenendaal, H.R., Wullings, B.A., 2017. Biofilm composition and threshold concentration for growth of *Legionella pneumophila* on surfaces exposed to flowing warm tap water without disinfectant. Appl. Environ. Microbiol. 83 https://doi.org/10.1128/aem.02737-16.
Vannini, C., 2004. Identification of the bacterial endosymbionts of the marine ciliate *Euplotes magnicirratus* (Ciliophora, Hypotrichia) and proposal of "*Candidatus* Devosia euplotis." Int. J. Syst. Evol. Microbiol. 54, 1151–1156. https://doi.org/10.1099/ijs.0.02759-0.
Vaz-Moreira, I., Nunes, O.C., Manaia, C.M., 2011. Diversity and Antibiotic resistance patterns of Sphingomonadaceae isolates from drinking water. Appl. Environ. Microbiol. 77, 5697–5706. https://doi.org/10.1128/aem.00579-11.
Vaz-Moreira, I., Egas, C., Nunes, O.C., Manaia, C.M., 2012. Bacterial diversity from the source to the tap: a comparative study based on 16S rRNA gene-DGGE and culture-dependent methods. FEMS Microbiol. Ecol. 83, 361–374. https://doi.org/10.1111/1574-6941.12002.
Wang, J., Liu, M., Xiao, H., Wu, W., Xie, M., Sun, M., Zhu, C., Li, P., 2013. Bacterial community structure in cooling water and biofilm in an industrial recirculating cooling water system. Water Sci. Technol. 68, 940–947. https://doi.org/10.2166/wst.2013.334.
Walser, S.M., Gerstner, D.G., Brenner, B., Höller, C., Liebl, B., Herr, C.E.W., 2014. Assessing the environmental health relevance of cooling towers – A systematic review of legionellosis outbreaks. Int. J. Hyg. Environ. Health 217, 145–154.

https://doi.org/10.1016/j.ijheh.2013.08.002.
Williams, M.M., Domingo, J.W.S., Meckes, M.C., Kelty, C.A., Rochon, H.S., 2004. Phylogenetic diversity of drinking water bacteria in a distribution system simulator. J. Appl. Microbiol. 96, 954–964. https://doi.org/10.1111/j.1365-2672.2004.02229.x.
Wingender, J., Flemming, H.-C., 2011. Biofilms in drinking water and their role as reservoir for pathogens. Int. J. Hyg Environ. Health 214, 417–423. https://doi.org/10.1016/j.ijheh.2011.05.009.
Yamamoto, H., Sugiura, M., Kusunoki, S., Ezaki, T., Ikedo, M., Yabuuchi, E., 1992. Factors stimulating propagation of *Legionellae* in cooling tower water. Appl. Environ. Microbiol. 58, 1394–1397.
Zacheus, O.M., Iivanainen, E.K., Nissinen, T.K., Lehtola, M.J., Martikainen, P.J., 2000. Bacterial biofilm formation on polyvinyl chloride, polyethylene and stainless steel exposed to ozonated water. Water Res. 34, 63–70. https://doi.org/10.1016/S0043-1354(99)00113-X.
Zeybek, Z., Binay, A.R., 2014. Growth ability of Gram negative bacteria in free-living amoebae. Exp. Parasitol. 145, S121–S126. https://doi.org/10.1016/j.exppara.2014.06.009.
Zhang, M., Liu, W., Nie, X., Li, C., Gu, J., Zhang, C., 2012. Molecular analysis of bacterial communities in biofilms of a drinking water Clearwell. Microb. Environ. 27, 443–448. https://doi.org/10.1264/jsme2.me12035.
Zwart, G., Crump, B., Kamst-van Agterveld, M., Hagen, F., Han, S., 2002. Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. Aquat. Microb. Ecol. 28, 141–155. https://doi.org/10.3354/ame028141.

Chapter IV

'Candidatus Cochliophilus cryoturris' (Coxiellaceae), a symbiont of the testate amoeba Cochliopodium minus

SCIENTIFIC REPORTS

OPEN

'*Candidatus* Cochliophilus cryoturris' (Coxiellaceae), a symbiont of the testate amoeba *Cochliopodium minus*

Received: 8 March 2017
Accepted: 2 May 2017
Published online: 13 June 2017

Han-Fei Tsao[1], Ute Scheikl[2], Jean-Marie Volland[3], Martina Köhsler[2], Monika Bright[3], Julia Walochnik[2] & Matthias Horn[1]

Free-living amoebae are well known for their role in controlling microbial community composition through grazing, but some groups, namely *Acanthamoeba* species, also frequently serve as hosts for bacterial symbionts. Here we report the first identification of a bacterial symbiont in the testate amoeba *Cochliopodium*. The amoeba was isolated from a cooling tower water sample and identified as *C. minus*. Fluorescence *in situ* hybridization and transmission electron microscopy revealed intracellular symbionts located in vacuoles. 16S rRNA-based phylogenetic analysis identified the endosymbiont as member of a monophyletic group within the family Coxiellaceae (Gammaproteobacteria; Legionellales), only moderately related to known amoeba symbionts. We propose to tentatively classify these bacteria as '*Candidatus* Cochliophilus cryoturris'. Our findings add both, a novel group of amoeba and a novel group of symbionts, to the growing list of bacteria-amoeba relationships.

Free-living amoebae are ubiquitous unicellular eukaryotes found in a wide range of habitats ranging from soil and aquatic environments to dust and air[1,2]. Grazing upon other microbes, they are important predators shaping microbial communities and affecting ecosystem functioning including nutrient availability and mineralization[3].

Free-living amoebae are also known as hosts for diverse bacteria and giant DNA viruses[4–6]. They serve as reservoirs for a number of human pathogens such as *Legionella pneumophila*[7], *Pseudomonas aeruginosa*[8], *Francisella tularensis*[9], *Coxiella burnetii*[10], *Vibrio cholerae*[11,12], *Aeromonas hydrophila*[13], and *Mycobacterium* species[14], all of which escape the regular phagolysosomal pathway and transiently replicate within amoeba trophozoites. In addition, long-term stable associations between obligate intracellular bacteria and amoebae have been reported, with known symbionts being affiliated with the bacterial phyla Chlamydiae, Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Bacteroidetes, and the candidate phylum Dependentiae (formerly TM6)[5,15–19].

Bacteria-amoeba relationships can have diverse effects on both partners. Enhancement of bacterial virulence upon amoeba passage has been reported as well as an increased cytopathogenicity of the amoeba host cell in the presence of bacterial symbionts[20–22]. Amoeba-associated bacteria share a number of characteristic genomic features; amoebae have thus been proposed to represent "melting pots" facilitating horizontal gene transfer between bacterial symbionts[23–26]. In addition, amoebae may have served as evolutionary training grounds for bacterial pathogens by providing conditions favoring bacteria with enhanced pathogenicity[27–29].

Bacteria-amoeba relationships have been studied almost exclusively in *Acanthamoeba* and few *Vanella* and *Vermamoeba* (formerly *Hartmannella*) isolates, all of them being free-living naked amoebae. Here we analyzed an amoeba newly recovered from a cooling tower water sample and identified as belonging to the testate amoebae *Cochliopodium*. These amoebae are covered with a tectum, a dorsal scale-like carbohydrate cell coat that protects the plasma membrane[30]. About 20 species have been recognized; they are primarily found in freshwater, brackish-water and marine environments, and rarely in soil[31–35]. The genus *Cochliopodium* represents a monophyletic group within the Amoebozoa, order Himatismenida, which forms a sister clade of the Centramoebida (containing the genus *Acanthamoeba* and others)[36].

[1]Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna, Vienna, Austria. [2]Institute of Specific Prophylaxis and Tropical Medicine, Medical University of Vienna, Vienna, Austria. [3]Department of Limnology and Oceanography, University of Vienna, Vienna, Austria. Correspondence and requests for materials should be addressed to M.H. (email: horn@microbial-ecology.net)

Here we report on the first characterization of a bacterial symbiont found in the testate amoeba *Cochliopodium*. The rod-shaped, Gram-negative bacterial symbiont replicates in host-derived vacuoles within its amoeba host cell and represents a distinct, yet uncharacterized lineage within the family Coxiellaceae (Gammaproteobacteria, Legionellales).

Results and Discussion

***Cochliopodium minus* from a cooling tower water sample.** During a cooling tower water screening study[37], an amoeba was isolated that could be readily propagated on non-nutrient agar plates coated with *E. coli*. Our attempts to establish an axenic culture using various media and a hypersensitive *E. coli* mutant[38] failed, and the amoeba was thus maintained routinely on agar plates. Morphological analyses, along with molecular identification based on 18S rRNA gene sequencing, confirmed the classification of this isolate as *Cochliopodium minus*, a testate amoeba found in diverse marine and fresh-water environments[31–35]. Highest 18S rRNA sequence similarity (>99%) was observed with *Cochliopodium* sp. F-117 (ATCC® 30936™) and various *Cochliopodium minus* strains. Characteristic for members of the genus, the trophozoites of this new isolate are covered with a dorsal monolayer of scale-like structures, the so called tectum (Fig. 1A,E,I,J). Lightmicrographs show the thin, scaled-covered hyaloplasmic sheet surrounding the granuloplasm, as well as the presence of subpseudopodiae (Fig. 1A,B). Contractile vacuoles of various stages, able to undergo fusion, are present in the cytoplasm (Fig. 1B,C). Occasionally, we could observe various encystation stages, including rounded trophozoites, the beginning of the encystation process (Fig. 1D). Transmission electron microscopy demonstrated that arrangement and fine structure of the scales are consistent with those described for other *Cochliopodium minus* isolates[39] (Fig. 1I,J). We thus refer to the novel isolate as *C. minus* strain 9B. It is interesting to note that another *C. minus* isolate was previously reported to contain bacterial symbionts, which could not be further characterized at the time[40].

Intracellular bacteria in *Cochliopodium minus* 9B Staining of *C. minus* 9B trophozoites with the DNA dye DAPI readily revealed small, rod-shaped bacteria within the amoeba cytoplasm that differed in fluorescence intensity, quantity and size from *E. coli* cells, which were primarily observed outside of the trophozoites (data not shown). The presence of bacteria other than *E. coli* was further confirmed by fluorescence *in situ* hybridization (FISH), and 16S rRNA gene sequencing recovered a sequence with highest similarity to members of the bacterial order Legionellales (Gammaproteobacteria). We designed an oligonucleotide probe for the specific detection of this sequence; its application in FISH together with general bacterial and eukaryotic probes demonstrated unambiguously the presence of bacterial endosymbionts in *C. minus* (Fig. 2). All analyzed trophozoites were infected, and the bacteria were always located in the amoeba cytoplasm, being notably absent in the nucleus[19,41]. The number of symbionts per amoeba cell varied and ranged from a few to over 100. The infection did not compromise the host's capability to encyst as described for some other symbionts[16,42], nor did we observe pronounced lysis of the amoeba at room temperature or at 28 °C. Over a period of two years, the amoebae remained infected, demonstrating that this symbiont-amoeba relationship is a stable long-term association. Failed attempts at host-free cultivation in diverse nutrient-rich and complex media under oxic and micro-oxic conditions indicate that the symbiont is dependent on its amoeba host and should therefore be considered obligate intracellular.

A novel clade of endosymbionts in the Coxiellaceae. The near full length 16S rRNA sequence (1,506 bp) of the *C. minus* 9B symbiont showed highest sequence similarity to a clone sequence from a soil sample (91%; accession number GQ263960.1). The most similar cultivated representative was *Coxiella burnetii* RSA 331 (CP000890.1), with only moderate sequence similarity (86%). Phylogenetic analysis confirmed that the symbiont is affiliated with the order Legionellales, in which it forms a well-supported monophyletic group together with a number of uncultured microbes predominantly from diverse marine environments (Fig. 3). This yet uncharacterized group represents a sister clade of the *Rickettsiella/Diplorickettsiella/Aquicella* group, three genera in the family Coxiellaceae. Notably, the bacterial symbiont of *C. minus* 9B is not closely affiliated with any known amoeba endosymbiont. However, its moderate relationship with members of the Legionellales is intriguing, as this order comprises a number of bacterial taxa associated with eukaryotes, including human and animal pathogens as well as parasites of amoebae (Fig. 3).

Members of the genera *Rickettsiella* and *Diplorickettsia* are parasites and symbionts of arthropods, including insects, crustaceans, and arachnids[43]. *Aquicella* species were first isolated from borehole and spa water samples and later shown to be able to thrive in co-culture with *Vermamoeba vermiformis*[44]. The genus *Coxiella* currently includes a single recognized species, *C. burnetii*, with numerous pathovars; these obligate intracellular bacteria are associated with insects and can cause severe infections in humans (aka Q fever)[45]; they might also be able to thrive in *Acanthamoeba castellanii*[10]. *Berkiella* species have recently been identified as intranuclear symbionts of *Acanthamoeba polyphaga*[41]. Furthermore, the Legionellaceae comprise a large number of facultative intracellular bacteria able to infect protists and animals including humans[46,47].

Taking into account current thresholds for the delineation of bacterial genera and families[48], the *C. minus* symbiont identified in this study represents a novel genus. We thus propose to tentatively classify this microbe as *Cochliophilus cryoturris* PDD8 (Cochliophilus, pertaining to the obligate intracellular lifestyle in its native host *Cochliopodium minus*; cryoturris, pertaining to the origin of the water sample, a cooling tower, from which the amoeba host was isolated). Currently, the novel genus is placed within the Coxiellaceae, although we noted that the sequence similarity of *C. cryoturris* and its relatives to other members of the Coxiellaceae is below the commonly used family level threshold of 86.5% (Fig. 3)[48]. *C. cryoturris* is currently the sole isolated representative of this novel genus; whether the other uncultured members of this clade are also naturally associated with protists is still unclear.

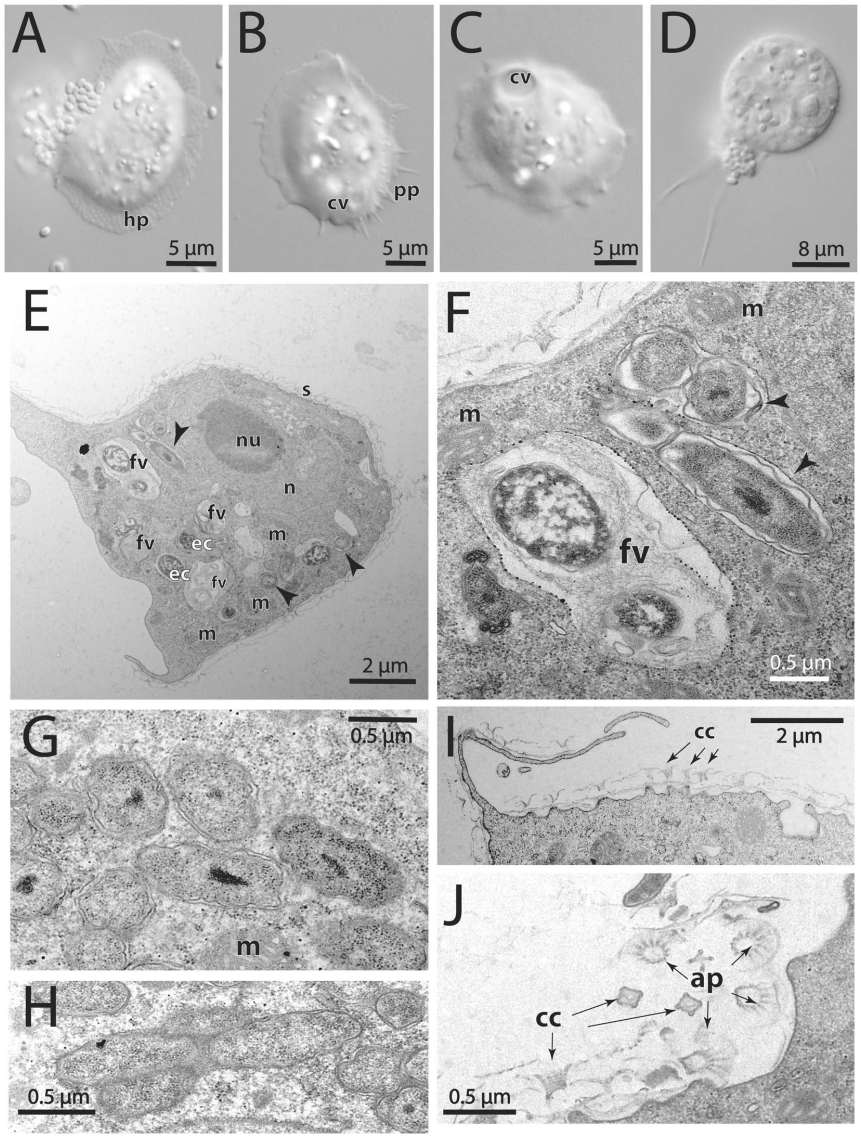


Figure 1. Intracellular location and morphology of '*Candidatus* Cochliophilus cryoturris' PDD8 in *Cochliopodium minus* 9B. (**A**–**C**) Lightmicrographs of *Cochliopodium minus* trophozoites with the characteristic scaled hyaloplasm (hp); the contractile vacuole (cv) involved in osmoregulation and subpseudopodia (pp) are readily visible. (**D**) Early stage of the encystation process showing the onset of cytoplasm condensation. (**E**) Electronmicrograph of a trophozoite containing several *C. cryoturris* symbionts (arrow heads) and food vacuoles (fv) with *E. coli* (ec; note the evidence for degradation). The host nucleus (n), the nucleolus (nu), and the dorsal scale cover (s) can be recognized. (**F**–**H**) *C. cryoturris* is located in membrane-bound compartments (arrow heads); the bacteria show a Gram-negative type cell envelope, with a partially widened periplasmic space; an electron-dense central area indicating condensed cytoplasmic components is present in many of the symbionts; mitochondria (m) can be seen in the vicinity of symbiont-containing vacuoles. (**I**) The characteristic scales of the *Cochliopodium minus* host are shown in a cross section including the funnel-shaped central column (cc). (**J**) The apical part (ap) of the scales as well as vertical and tangential sections of central columns (cc) are visible.

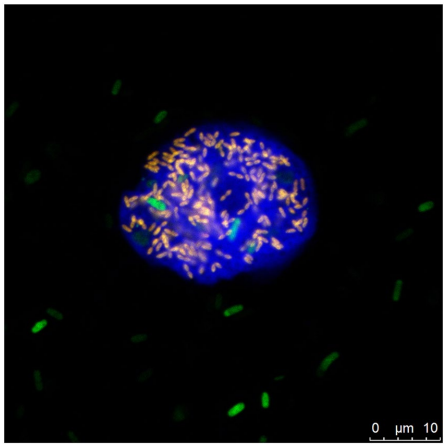


Figure 2. Identification of '*Candidatus* Cochliophilus cryoturris' PDD8 in its native host *Cochliopodium minus* by fluorescence *in situ* hybridization. The bacterial endosymbionts were visualized with the specific probe PDD8-644 labeled with Cy3 (red) and general bacterial probes (EUB-mix) labeled with Fluos (green; the overlap appears yellow). *E. coli* cells added as amoeba food source are visible in green; the *Cochliopodium minus* trophozoite was counterstained using the eukaryotic probe EUK516 labeled with Cy5 (blue).

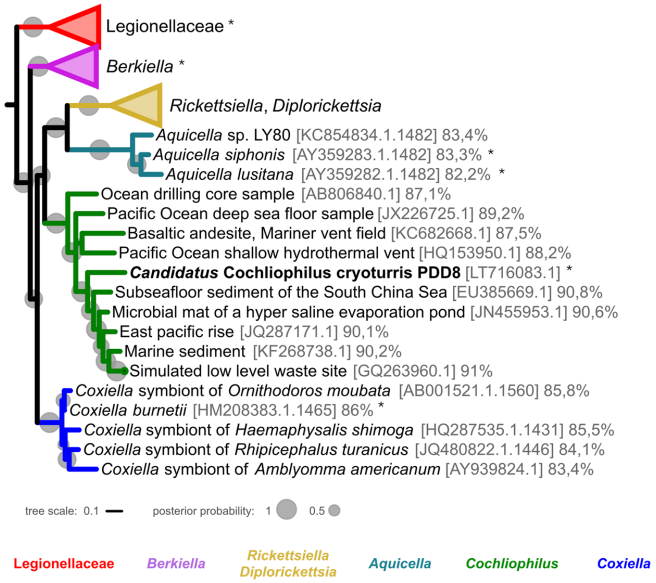


Figure 3. Relationship of '*Candidatus* Cochliophilus cryoturris' PDD8 with other members of the Legionellales. A 16S rRNA tree based on PhyloBayes using the CAT model and GTR exchange rates is shown. Posterior probability values are indicated as grey circles at the nodes. Accession numbers and sequence similarity values to *C. cryoturris* are provided. Taxa with reported amoeba association are labeled with an asterisk. *C. cryoturris* together with a number of sequences obtained from various aquatic samples forms a novel sister clade of the *Rickettsiella/Diplorickettsiella/Aquicella* group in the family Coxiellaceae.

Vacuolar location of *C. cryoturris*. The appearance and host-associated intracellular lifestyle of *Cochliophilus cryoturris* PDD8 is reminiscent of that of many of its relatives in the Legionellales. *C. cryoturris* cells are small and show a short rod-shaped morphology, measuring 0.5 ± 0.1 µm in width and around 1 ± 0.2 µm in length; they show a Gram-negative type cell envelope and frequently a condensed, electron-dense central region in the cytoplasm (Fig. 1F,G). The symbionts are not located directly within the amoeba cytoplasm but in membrane-bound compartments (Fig. 1E,F). These clearly differ from the food vacuoles containing (degraded) *E. coli* cells, which can also be seen in the amoeba cytoplasm (Fig. 1E,F). Mitochondria are frequently located in the vicinity of the symbiont-containing vacuoles (Fig. 1E–H).

Bacterial strategies for escaping phagolysosomal degradation differ. *Coxiella burnetii*, the closest cultured relative of *C. cryoturris*, is able to resist the harsh conditions after fusion of the phagosome with lysosomes and modifies the phagolysosome to interact with the autophagic pathway, promoting metabolic activity and replication[49,50]. *Legionella pneumophila* takes an alternative route and prevents lysosomal fusion to establish a heavily modified vacuolar compartment resembling endoplasmic reticulum (ER) membranes[51,52]. How *C. cryoturris* establishes its intracellular niche is currently unknown. We were, however, unable to infect different *Acanthamoeba* species, which are otherwise permissive to an array of phylogenetically diverse intracellular bacteria. This suggests that *C. cryoturris* is well adapted to infection of its natural *Cochliopodium* host but might have a limited host range.

In conclusion, we discovered and identified the first naturally occurring *Cochliopodium* endosymbiont together with its amoeba host. The symbiont is a representative of a hitherto uncharacterized clade of microbes found in diverse aquatic environments and related to other intracellular bacteria in the family Coxiellaceae. Together this indicates that relationships between free-living amoebae and bacterial symbionts are more widespread than currently recognized.

Description of '*Candidatus* Cochliophilus cryoturris' PDD8. *Cochliophilus cryoturris (*Cochliophilus, pertaining to the obligate intracellular lifestyle of the bacteria in the native host *Cochliopodium minus;* cryoturris, pertaining the origin of the water sample from which the amoeba host was recovered, a cooling tower). Phylogenetic position: *Proteobacteria*; *Gammaproteobacteria*; *Legionellales*; *Coxiellaceae*. Rod-shaped Gram-negative bacteria; 0.5 µm in width and 1 µm in length. Obligate intracellular symbiont of *Cochliopodium minus* 9B; residing in membrane-bound compartments. The amoeba host was isolated from water samples of a cooling tower in Vienna, Austria (18S rRNA gene sequence accession number at Genbank/ENA/DDBJ KU215597). Basis of assignment 16S rRNA gene (Genbank/ENA/DDBJ accession number LT716083) and oligonucleotide probe PDD8-644 (5′-TCTTCGACTCCAGCCGCAC-3′; http://probebase.net/pb_report/probe/4042).

Material and Methods

Amoeba isolation and cultivation. A water sample was collected from the tank-bottom of a cooling tower and stored at 4 °C. 250 µl were filtered onto a cellulose-nitrate filter (Sartorius Lab Instruments GmbH & Co. KG, Göttingen, Germany; pore-size 0.2 µm). The filter was cut into two pieces, which were transferred onto non-nutrient agar plates (PAS; 0.12 g l^{-1} NaCl, 0.004 g l^{-1} $MgSO_4$ *7 H_2O, 0.004 g l^{-1} $CaCl_2$*2H_2O, 0.142 g l^{-1} Na_2 HPO_4, 0.136 g l^{-1} KH_2PO_4, 1.5 g l^{-1} agar) coated with *Escherichia coli* and stored at room temperature. Detected amoebae were cloned by daily serial sub-culturing of single cells onto fresh *E. coli*-coated agar plates using a sterile inoculation loop. Clonal cultures were maintained by weekly sub-culturing and morphological identification of the amoeba was accomplished by inverted phase contrast and bright field microscopy (Nikon Eclipse E800) using the identification keys of Page and Smirnov[53,54]. Amoebae thriving on the agar surface were repeatedly transferred to fresh agar plates containing *E. coli* JW5503-1 ΔtolC732::kan as food source to facilitate axenization as described[38,55]. Amoeba were routinely maintained on agar plates covered with *E. coli*. Fresh bacteria suspended in PAS were added to the agar plate once per week, and once per month an agar piece was transferred to a new plate.

To facilitate microscopic and molecular analysis, an agar piece containing amoebae was transferred to a culture flask (Nunclon delta-surface, Thermo Scientific, St. Leon-Rot, Germany) containing 6 ml PAS. Amoebae were allowed to attach for 12 h and washed twice with PAS to remove *E. coli*. Amoeba cells were collected for further analysis by detaching through vigorous shaking.

Infection of *Acanthamoeba*. Amoebae were harvested from cell culture flasks, and the cell suspension was transferred to a Dounce tissue grinder (Sigma-Aldrich Handels GmbH, Vienna, Austria) using the tight pestle for 15 times to break up the cells and release the endosymbionts. The suspension was filtered (Macherey-Nagel, Düren, Germany; pore size 5 µm) twice to remove remaining intact amoebae. The symbiont suspension was added to cultures of *Acanthamoeba castellanii* Neff (ATCC 30010), *Acanthamoeba* sp. 5a2[18], and *Acanthamoeba* sp. UWC12[56]. The outcome of the infection progress was monitored by fluorescence *in situ* hybridization.

Host-free cultivation attempts. Host free growth of purified symbionts was tested using the following media: peptone-yeast-glucose (PYG) broth, trypticase-soy broth with yeast extract (TSY), peptone-yeast-nucleic acids-folic acids-hemin (PYNFH) broth, buffered charcoal yeast extract (BCYE) medium with and without supplements. Micro-oxic conditions were applied using the CampyGen Compact system (Thermo Scientific, St. Leon-Rot, Germany).

Transmission electron microscopy. Amoebae were fixed (2.5% glutaraldehyde in 3 mM cacodylate buffer containing 0.1 M sucrose, pH 6.5) in culture flasks for one hour, detached using a cell-scraper, and concentrated by centrifugation (2900 rcf, 6 min). The pelleted cells were washed with 0.1 M cacodylate-sucrose buffer (pH 7.2–7.4) for three times and then resuspended in 40 µl 1% agarose (Low melting point agarose; Promega, Mannheim, Deutschland). The agar pellet was solidified on ice for 45 min and then cut into smaller pieces with 1 mm thickness, which were fixed in 1% OsO_4 for 1 h and dehydrated in an increasing ethanol series. Agar blocks were embedded in Low Viscosity resin (Agar Scientific®) and polymerized for 48 h at 60 °C. Ultrathin sections placed on Formvar® coated slot grids were stained with 0.5% uranyl acetate and 3% lead citrate prior to imaging with a Zeiss® Libra 120 transmission electron microscope.

Fluorescence *in situ* hybridization. An agar piece containing amoebae was placed upside-down on a microscope slide well covered with 10 µl PAS. Amoebae were allowed to attach for 30 min, and the well was washed once with PAS, followed by fixation with 4% formaldehyde (12 min at room temperature). An endosymbiont-specific probe (PDD8-644, 5′-TCTTCGACTCCAGCCGCAC-3′) was designed using the ARB software package and validated with probeCheck and Silva testProbe 3.0[57, 58]. The probe sequence was deposited at probeBase (http://probebase.net/pb_report/probe/4042)[59]. All probes were synthesized by Thermo Fisher Scientific (St. Leon-Rot, Germany). Hybridizations were performed by combining the specific probe PDD8-644, the eukaryotic probe EUK516 (5′-ACCAGACTTGCCC TCC-3′)[60], and the bacterial probe set EUB338 I-III (5′-GCTGCCTCCCGTAGGAGT-3′, 5′-GCAGCCACCCGTAGGTGT-3′ and 5′-GCTGCCACCCGTAGGTGT-3′)[60, 61]. Hybridization was carried out for 2 h at 46 °C with 20% formamide using standard hybridization and washing buffers[62]. Slides were embedded in Citifluor prior to examination with a confocal laser scanning microscope (LSM 510 Meta, Zeiss, Oberkochen, Germany) or an epifluorescence microscope (Axioplan 2 Imaging, Zeiss) equipped with a CCD camera (AxioCam HRc, Zeiss).

DNA extraction, PCR, cloning and sequencing. Amplification and sequencing of the amoebal 18S rDNA was carried out as described earlier[37]. Briefly, trophozoites from clonal cultures were harvested from plates with cotton swabs and re-suspended in 15 ml centrifuge tubes filled with 5 ml 0.9% sodium chloride (NaCl). The samples were centrifuged for 10 min at 800 × g, the supernatants were discarded and the pellets were re-suspended in 200 µl 0.9% NaCl. Total genomic DNA was extracted from the cells using the QIAmp® DNA Mini Kit (Qiagen, Hilden, Germany). The entire 18S rDNA gene was amplified using the newly designed Cochlio-1 primer (5′-CCTGGTTGATCCTGCCAG-3′) and the SSU2 (5′-TCCTGATCCTTCTGCAGGTTCAC-3′)[63]. Gel bands were extracted with the GFX PCR DNA and Gel Band Purification Kit (GE Healthcare, UK) and overlapping fragments were directly sequenced in both directions with the ABI PRISM® BigDye sequencing kit, using the internal primers P1fw 5′-CAAGTCTGGTGCCAGCAGC-3′, P1rev 5′-GCTGCTGGCACCAGACTTG-3′, P2fw 5′-GATCAGATACCGTCGTAGTC-3′, P2rev 5′-GACTACGACGGTATCTGATC-3′, P3fw 5′-CAGGTCTGTGATGCCCTTAG-3′ and P3rev 5′-CTAAGGGCATCACAGACCTG-3′[64] and an ABI PRISM 310® automated sequencer (PE Applied Biosystems, Germany). A consensus sequence was built with the GeneDoc sequence editor[65].

In order to sequence the bacterial 16S rRNA gene, amoebae were collected from an *E. coli*-depleted culture flask, and 2 ml of the suspension were transferred to a 2 ml microcentrifuge tube. DNA extraction of the amoeba culture was carried out using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Amplification of bacterial 16S rRNA genes was performed using primers 616F (5′-AGAGTTTGATYMTGGCTCAG-3′) and 1492R (5′-GGYTACCTTGTTACGACTT-3′) at 52 °C annealing temperature and 35 cycles[66, 67]. PCR reactions contained 100 ng template DNA, 50 pmol/µl of each primer, 1 unit of Taq DNA polymerase (Fermentas, St. Leon-Rot, Germany), 10x Taq buffer with KCl and 2 mM $MgCl_2$, and 0.2 mM of each deoxynucleotide in a total volume of 50 µl. PCR products were purified using the PCR Purification Kit (Qiagen) and subsequently cloned using the TOPO XL Cloning Kit (Invitrogen, Darmstadt, Germany) per manufacturer's recommendations. Sanger sequencing of four clones was performed by Microsynth Austria.

Phylogenetic analysis. The obtained 16S rRNA gene sequence was subjected to sequence homology search against the nr/nt database using the BLASTn service available at the NCBI website[68]. The top ten high scoring sequences with a minimum length of 1,400 nt were downloaded, and phylogenetic analysis was performed together with a selection of related taxa retrieved from the SILVA ribosomal RNA gene database[69]. The SINA aligner[70] with standard settings and variability set to "Bacteria" was used for sequence alignment. The alignment was trimmed at both ends to only include positions covered in all sequences. Pairwise sequence similarity was calculated using ARB[57]. Phylogenetic trees were calculated using PhyloBayes[71] with the CAT model[72] and GTR exchange rates. Ten independent chains were calculated with 210 generations each. For the final converged tree, all 10 chains were taken into account whereas the first 20 generations trees were removed. iTOL v3[73] was used to edit and label the tree.

Data availability. DNA sequences determined in this study were deposited at Genbank/ENA/DDBJ under accession numbers KU215597 (18S rRNA gene sequence of *C. minus* 9B) and LT716083 (16S rRNA gene sequence of '*Candidatus* Cochliophilus cryoturris' PDD8).

References

1. Rodríguez-Zaragoza, S. Ecology of free-living amoebae. *Crit. Rev. Microbiol.* **20**, 225–41 (1994).
2. Khan, N. A. Acanthamoeba: Biology and increasing importance in human health. *FEMS Microbiol. Rev.* **30**, 564–595 (2006).
3. Koller, R., Robin, C., Bonkowski, M., Ruess, L. & Scheu, S. Litter quality as driving factor for plant nutrition via grazing of protozoa on soil microorganisms. *FEMS Microbiol. Ecol.* **85**, 241–250 (2013).
4. Aherfi, S., Colson, P., La Scola, B. & Raoult, D. Giant Viruses of Amoebas: An Update. *Front. Microbiol.* **7**, 1–14 (2016).
5. Horn, M. & Wagner, M. Bacterial endosymbionts of free-living amoebae. *J. Eukaryot. Microbiol.* **51**, 509–14 (2004).
6. Greub, G. & Raoult, D. Microorganisms resistant to free-living amoebae. *Clin. Microbiol. Rev.* **17**, 413–33 (2004).
7. Newsome, A. L., Baker, R. L., Miller, R. D. & Arnold, R. R. Interactions between Naegleria fowleri and Legionella pneumophila. *Infect. Immun.* **50**, 449–452 (1985).
8. Michel, R., Burghardt, H. & Bergmann, H. Acanthamoeba, naturally intracellularly infected with Pseudomonas aeruginosa, after their isolation from a microbiologically contaminated drinking water system in a hospital. *Zentralbl Hyg Umweltmed* **196**, 532 (1995).
9. Abd, H., Johansson, T., Golovliov, I., Sandstrom, G. & Forsman, M. Survival and Growth of Francisella tularensis in Acanthamoeba castellanii. *Appl. Environ. Microbiol.* **69**, 600–606 (2003).
10. La Scola, B. & Raoult, D. Survival of Coxiella burnetii within free-living amoeba Acanthamoeba castellanii. *Clin. Microbiol. Infect.* **7**, 75–9 (2001).

11. Thom, S., Warhurst, D. & Drasar, B. S. Association of Vibrio cholerae with fresh water amoebae. *J. Med. Microbiol.* **36**, 303–306 (1992).
12. Sandström, G., Saeed, A. & Abd, H. Acanthamoeba polyphaga is a possible host for Vibrio cholerae in aquatic environments. *Exp. Parasitol.* **126**, 65–8 (2010).
13. Yousuf, F. A., Siddiqui, R. & Khan, N. A. Acanthamoeba castellanii of the T4 genotype is a potential environmental host for Enterobacter aerogenes and Aeromonas hydrophila. *Parasit. Vectors* **6**, 169 (2013).
14. Thomas, V. & McDonnell, G. Relationship between mycobacteria and amoebae: Ecological and epidemiological concerns. *Lett. Appl. Microbiol.* **45**, 349–357 (2007).
15. Pagnier, I. *et al*. Babela massiliensis, a representative of a widespread bacterial phylum with unusual adaptations to parasitism in amoebae. *Biol. Direct* **10**, 13 (2015).
16. Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L. & Héchard, Y. Shedding light on microbial dark matter: a TM6 bacterium as natural endosymbiont of a free-living amoeba. *Environ. Microbiol. Rep.* **7**, 970–978 (2015).
17. Collingro, A. *et al*. Recovery of an environmental Chlamydia strain from activated sludge by co-cultivation with Acanthamoeba sp. *Microbiology* **151**, 301–9 (2005).
18. Schmitz-Esser, S. *et al*. Diversity of bacterial endosymbionts of environmental acanthamoeba isolates. *Appl. Environ. Microbiol.* **74**, 5822–31 (2008).
19. Schulz, F. *et al*. Life in an unusual intracellular niche: a bacterial symbiont infecting the nucleus of amoebae. *ISME J.* **8**, 1634–1644 (2014).
20. Cirillo, J. D. *et al*. Interaction of Mycobacterium avium with environmental amoebae enhances virulence. *Infect. Immun.* **65**, 3759–67 (1997).
21. Cirillo, J. D., Falkow, S. & Tompkins, L. S. Growth of Legionella pneumophila in Acanthamoeba castellanii enhances invasion. *Infect. Immun.* **62**, 3254–61 (1994).
22. Fritsche, T. R., Sobek, D. & Gautom, R. K. Enhancement of *in vitro* cytopathogenicity by Acanthamoeba spp. following acquisition of bacterial endosymbionts. *FEMS Microbiol. Lett* **166**, 231–236 (1998).
23. Thomas, V. & Greub, G. Amoeba/Amoebal Symbiont Genetic Transfers: Lessons from Giant Virus Neighbours. *Intervirology* **53**, 254–267 (2010).
24. Ogata, H. *et al*. Genome sequence of Rickettsia bellii illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet.* **2**, e76 (2006).
25. Boyer, M. *et al*. Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc. Natl. Acad. Sci* **106**, 21848–21853 (2009).
26. Schmitz-Esser, S. *et al*. The genome of the amoeba symbiont 'Candidatus Amoebophilus asiaticus' reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* **192**, 1045–57 (2010).
27. Molmeret, M. *et al*. Amoebae as training grounds for intracellular bacterial pathogens. *Appl. Environ. Microbiol.* **71**, 20–28 (2005).
28. Horn, M. Illuminating the evolutionary history of chlamydiae. *Science (80-.).* **304**, 728–730 (2004).
29. Harb, O. S., Gao, L. & Kwaik, Y. A. From protozoa to mammalian cells: a new paradigm in the life cycle of intracellular bacterial pathogens. *Environ. Microbiol.* **2**, 251–265 (2000).
30. Bark, A. W. A study of the genus Cochliopodium Hertwig and Lesser 1874. *Protistologica* **9**, 119–138 (1973).
31. Kudryavtsev, A., Brown, S. & Smirnov, A. Cochliopodium barki n. sp. (Rhizopoda, Himatismenida) re-isolated from soil 30 years after its initial description. *Eur. J. Protistol.* **40**, 283–287 (2004).
32. Tekle, Y. I., Roger, A. O., Lecky, A. F. & Kelly, S. D. A new freshwater amoeba: Cochliopodium pentatrifurcatum n. sp. (Amoebozoa, Amorphea). *J. Eukaryot. Microbiol.* **60**, 342–349 (2013).
33. Garstecki, T. & Arndt, H. Seasonal abundances and community structure of benthic rhizopods in shallow lagoons of the southern Baltic Sea. *Eur. J. Protistol.* **36**, 103–115 (2000).
34. Kudryavtsev, A. & Smirnov, A. Cochliopodium gallicum n. sp. (Himatismenida), an amoeba bearing unique scales, from cyanobacterial mats in the Camargue (France). *Eur. J. Protistol.* **42**, 3–7 (2006).
35. Geisen, S., Kudryavtsev, A., Bonkowski, M. & Smirnov, A. Discrepancy between species borders at morphological and molecular levels in the genus Cochliopodium (Amoebozoa, Himatismenida), with the description of Cochliopodium plurinucleolum n. sp. *Protist* **165**, 364–383 (2014).
36. Cavalier-Smith, T., Chao, E. E. & Lewis, R. 187-gene phylogeny of protozoan phylum Amoebozoa reveals a new class (Cutosea) of deep-branching, ultrastructurally unique, enveloped marine Lobosa and clarifies amoeba evolution. *Mol. Phylogenet. Evol.* **99**, 275–296 (2016).
37. Scheikl, U., Tsao, H. F., Horn, M., Indra, A. & Walochnik, J. Free-living amoebae and their associated bacteria in Austrian cooling towers: a 1-year routine screening. *Parasitol. Res.* **115**, 3365–3374 (2016).
38. Lagkouvardos, I., Shen, J. & Horn, M. Improved axenization method reveals complexity of symbiotic associations between bacteria and acanthamoebae. *Environ. Microbiol. Rep* **6**, 383–8 (2014).
39. Kudryavtsev, A. 'Minute' species of Cochliopodium (Himatismenida): Description of three new fresh- and brackish-water species with a new diagnosis for Cochliopodium minus Page, 1976. *Eur. J. Protistol.* **42**, 77–89 (2006).
40. Dyková, I., Lom, J. & Macháčková, B. Cochliopodium minus, a scale-bearing amoeba isolated from organs of perch Perca fluviatilis. *Dis. Aquat. Organ* **34**, 205–210 (1998).
41. Mehari, Y. T. *et al*. Description of 'Candidatus Berkiella aquae' and 'Candidatus Berkiella cookevillensis', two intranuclear bacteria of freshwater amoebae. *Int. J. Syst. Evol. Microbiol* **66**, 536–541 (2016).
42. Leitsch, D. *et al*. Proteomic aspects of Parachlamydia acanthamoebae infection in Acanthamoeba spp. *ISME J.* **4**, 1366–1374 (2010).
43. Cordaux, R. *et al*. Molecular characterization and evolution of arthropod-pathogenic Rickettsiella bacteria. *Appl. Environ. Microbiol.* **73**, 5045–5047 (2007).
44. Santos, P. *et al*. Gamma-Proteobacteria Aquicella lusitana gen. nov., sp. nov., and Aquicella siphonis sp. nov. Infect Protozoa and Require Activated Charcoal for Growth in Laboratory Media. *Appl. Environ. Microbiol.* **69**, 6533–6540 (2003).
45. Maurin, M. & Raoult, D. Q fever. *Clin. Microbiol. Rev.* **12**, 518–53 (1999).
46. Falkow, S., Rosenberg, E., Schleifer, K.-H. & Stackebrandt, E. In *The Prokaryotes: Vol. 6: Proteobacteria: Gamma Subclass* 990, doi:10.1007/978-3-642-38954-2 (Springer Science & Business Media, 2006).
47. Burstein, D. *et al*. Genomic analysis of 38 Legionella species identifies large and diverse effector repertoires. *Nat. Genet.* **48**, 167–175 (2016).
48. Yarza, P. *et al*. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645 (2014).
49. Ghigo, E., Colombo, M. I. & Heinzen, R. A. The Coxiella burnetii parasitophorous vacuole. *Adv. Exp. Med. Biol* **984**, 141–169 (2012).
50. Jimenez, A., Chen, D. & Alto, N. M. How bacteria subvert animal cell structure and function. *Annu. Rev. Cell Dev. Biol.* **32**, 373–397 (2016).
51. Finsel, I. & Hilbi, H. Formation of a pathogen vacuole according to Legionella pneumophila: how to kill one bird with many stones. *Cell. Microbiol.* **17**, 935–950 (2015).
52. Escoll, P., Mondino, S., Rolando, M. & Buchrieser, C. Targeting of host organelles by pathogenic bacteria: a sophisticated subversion strategy. *Nat. Rev. Microbiol.* **14**, 5–19 (2015).

53. Smirnov, A. V., Chao, E., Nassonova, E. S. & Cavalier-Smith, T. A revised classification of naked lobose amoebae (Amoebozoa: Lobosa). *Protist* **162**, 545–570 (2011).
54. Page, F. C. & Siemensma, F. J. *Protozoenfauna, Bd. 2: Nackte Rhizopoda und Heliozoea*. (Spektrum Akademischer Verlag, 1991).
55. Baba, T. *et al*. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
56. Horn, M. *et al*. Obligate bacterial endosymbionts of Acanthamoeba spp. related to the beta-Proteobacteria: proposal of 'Candidatus Procabacter acanthamoebae' gen. nov., sp. nov. *Int. J. Syst. Evol. Microbiol* **52**, 599–605 (2002).
57. Ludwig, W. ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**, 1363–1371 (2004).
58. Loy, A. *et al*. probeCheck - a central resource for evaluating oligonucleotide probe coverage and specificity. *Environ. Microbiol* **10**, 2894–2898 (2008).
59. Greuter, D., Loy, A., Horn, M. & Rattei, T. probeBase—an online resource for rRNA-targeted oligonucleotide probes and primers: new features 2016. *Nucleic Acids Res.* **44**, D586–D589 (2016).
60. Amann, R. I. *et al*. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* **56**, 1919–25 (1990).
61. Daims, H., Brühl, A., Amann, R., Schleifer, K.-H. & Wagner, M. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: Development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol.* **22**, 434–444 (1999).
62. Daims, H., Stoecker, K. & Wagner, M. Fluorescence *in situ* hybridization for the detection of prokaryotes. *Mol. Microb. Ecol.* 213–239 (2005).
63. Gast, R. J., Fuerst, P. A. & Byers, T. J. Discovery of group I introns in the nuclear small subunit ribosomal RNA genes of Acanthamoeba. *Nucleic Acids Res.* **22**, 592–596 (1994).
64. Walochnik, J., Michel, R. & Aspöck, H. A molecular biological approach to the phylogenetic position of the genus Hyperamoeba. *J. Eukaryote Microbiol.* **51**, 433–440 (2004).
65. Nicholas, K. B., Nicholas, H. B. & Deerfield, D. W. GeneDoc: analysis and visualization of genetic variation. *EMBnet News* **4**, 14 (1997).
66. Juretschko, S. *et al*. Combined molecular and conventional analyses of nitrifying bacterium diversity in activated sludge: Nitrosococcus mobilis and Nitrospira-like bacteria as dominant populations. *Appl. Environ. Microbiol.* **64**, 3042–3051 (1998).
67. Loy, A. *et al*. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* **68**, 5064–5081 (2002).
68. Altschul, S. F. *et al*. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389–3402 (1997).
69. Quast, C. *et al*. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* **41**, 590–596 (2013).
70. Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).
71. Lartillot, N. & Poujol, R. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* **28**, 729–744 (2011).
72. Lartillot, N. & Philippe, H. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol. Biol. Evol.* **21**, 1095–1109 (2004).
73. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).

Acknowledgements

Electron microscopy work was performed at the Core Facility Cell Imaging and Ultrastructure Research, University of Vienna. We acknowledge Dr. Martin Mrva for his support in morphological identification of the amoeba. We also want to thank Dr. Alexander Indra for establishing and organizing the water sample logistics. This work was supported by Austrian Science Fund (FWF) grants no. TRP209-B20, P24565-B22, and I1628-B22 in the framework of the ERA-NET InfectERA EUGENPATH. MH was supported by the European Research Council project EVOCHLAMY (ERC StG grant no. 281633).

Author Contributions

H.F.T., U.S., J.W., and M.H. conceived the study. H.F.T. and U.S. performed the sampling; U.S. and M.K. isolated and characterized the amoeba. H.F.T. identified and characterized the bacterial symbiont. J.M.V. and M.B. performed electron microscopy analysis. H.F.T. prepared the figures; H.F.T. and M.H. wrote the manuscript. All authors reviewed and edited the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

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

Chapter V

Better in than out: *Procabacter acanthamoebae*, an amoebae endosymbiont in the making

Han-Fei Tsao, Stephan Köstlbacher, and Matthias Horn

Draft manuscript

Abstract

Free-living amoebae are known regulators of environmental microbial communities through predatory behavior. The Acanthamoeba species is also frequently observed as hosts for bacterial symbionts. *Procabacter acanthamoebae* is a member of a novel clade of amoeba endosymbionts, classified within the class Betaproteobacteria. Here we analyzed the infection of *Procabacter* with its natural host *Acanthamoeba castellanii*, the influence on the host's fitness, and the host-free cultivation of *Procabacter* with fluorescence in situ hybridization. Further, we have sequenced four *Procabacter* spp. to elucidate their genomic repertoire and the underlying genomic basis for an intracellular lifestyle. *Procabacter's* infection cycle is completed after nine days and exhibits only a moderate impact on the host's growth speed. Under host-free, nutrient-rich conditions, we could observe dividing *Procabacter* cells and biofilm formation. Genome analysis revealed the conservation of critical metabolic pathways as well as the *de-novo* biosynthesis capability of nucleotides and cofactors. They also encode for various secretion systems (T1SS, T2SS, T3SS, and T5SS) with a variety of predicted effector proteins, including proteins with eukaryotic domains. The presence of an additionally horizontally transferred cytochrome bd oxidase might have increased Procabacter's survival capacity in environments with different oxygen levels. Ancestral genome reconstruction showed elevated gene loss rate with the simultaneous acquisition of proteins essential for host cell interaction. Taken together, we suggest that the group of Procabacter comprises novel endosymbionts currently undergoing adaptation to the intracellular lifestyle.

Introduction

Free-living amoeba (FLA) are ubiquitous unicellular eukaryotes, inhabiting various environments (Horn et al. 2002; Rodríguez-Zaragoza 1994). They are also able to influence surrounding microbial communities by serving as predators of other microbes due to their phagocytic properties. Not all bacteria end up as food, as some groups have adapted to resist the phagolysosomal pathway and use FLA's cytosol as a niche for resources and protection. There are several reports of long-term associations between FLA and intracellular bacteria (Horn et al. 2002). Until now, all described bacterial endosymbionts of FLA are members of either one of the following taxonomic groups; Chlamydiae, Bacteroidetes, Dependentiae (formerly known as TM6), Alphaproteobacteria and Gammaproteobacteria (Delafont et al. 2015; Schmitz-Esser et al. 2008; Tsao et al. 2017; Horn 2008; Schulz et al. 2015, 2014). All of them are obligate endosymbionts, unable to be cultivated outside of a host cell. Their genomes are reduced and often lack the potential to *de-novo* synthesize amino acids, cofactors, and nucleotides. Due to their high reliance on the amoeba host to provide for nutrients and a replication niche, they have developed a range of means to aid host cell interaction like the expression of secretion systems with effector proteins, substrate transporters to utilize hosts' metabolites, and other various virulence factors to evade the cellular defense mechanisms (Schulz et al. 2015; Collingro et al. 2011; Schmitz-Esser et al. 2010; Böck et al. 2017). Often within the cytosol of the host, there is gene exchange between intracellular symbionts and selection for mechanisms of more advanced invasion strategies. Thus, scientists have proposed FLA being "genomic melting pots" (Thomas and Greub 2010; Ogata et al. 2006; Boyer et al. 2009). The insights gained by studying endosymbionts can also be applied to pathogens, as the initial infection and the circumvention of cellular defense are often very similar (Elwell, Mirrashidi, and Engel 2016).

Members of the genus *Procabacter* are the only group within the class Betaproteobacteria known to infect FLAs and form a long-term relationship. These rod-shaped, Gram-negative bacteria were isolated from different *Acanthamoeba* spp. found in various environments, ranging from soil samples from the United States, lake sediments in Malaysia, to corneal scrapings of a keratitis patient (Horn et al. 2002; Heinz et al. 2007; Schmitz-Esser et al. 2008). The first described species of the genus, *Candidatus* Procabacter acanthamoebae, resides within the host's cytosol, as shown by transmission electron microscopy (TEM) and

fluorescence in situ hybridization with genus-specific probes (Horn et al. 2002). Co-infected FLA sample comprising of a *Procabacter* sp. and another amoebal endosymbiont, *Parachlamydia acanthamoebae* OEW1 (Chlamydiae), within the host's cytosol, has also been reported (Heinz et al. 2007). Initial 16S rRNA sequence similarity analysis suggested less than 90% similarity to other Betaproteobacteria sequences, indicating that *Procabacter* belongs to its distinct lineage, with members of the family Neisseriaceae being the closest relatives (Horn et al. 2002). However, no data about *Procabacter*'s infection of its host, how the infection affects the host, nor the underlying genomic repertoire responsible for the intracellular lifestyle are available. By answering these questions, we hope to better understand the mechanisms of the symbiotic interactions used by this unique group of endosymbionts. Thus, revealing potential novel host-symbiont interaction and increasing our overall knowledge of endosymbiosis in protists.

Here we first analyze and describe the infection cycle of *P. acanthamoebae* and the impact on the hosts' growth. Then we sequenced and annotated the genome of four *Procabacter* strains to describe their genomic features. Cultivation attempts with nutrient-rich medium show dividing *Procabacter* cells and biofilm formation under host-free microaerobic conditions. Phylogenetic analysis suggests that the *Procabacter*-clade are indeed different enough to be considered a unique family within the order Burkholderiales, closely related to the families Chromobacteriaceae and Neisseriaceae. *Procabacter* genomes are only moderately reduced, retaining the biosynthetic-capability for various metabolic factors. We found a distinct set of genes that are relevant for establishing an intracellular lifestyle and host-communication, including several secretion systems, pili, and effectors. Through ancestral genome reconstruction, we identified the gain of several protein families, which are potentially necessary for the host interaction. Taken together, we suggest that *Procabacter* is a symbiont "in the making", currently under adaptation to the intracellular lifestyle.

Results and discussion

1. *P. acanthamoebae* infection of its natural amoebae host and higher eukaryotes

We investigated the infection cycle of *Procabacter acanthamoebae* UWC12 in the natural host *Acanthamoeba* sp. UWC12. Here we define the term "natural host" as the host organism the endosymbiont has been first found in. We distinguished between three different infection loads, uninfected, lowly infected (up to eight endosymbionts detected in a single amoeba cell), and highly infected (more than eight endosymbionts per host). After uptake, bacteria encounter digestion by the host. Sometimes a specific infection load or multiplicity of infection is required to overcome this cellular defense. This distinction helped us to describe the infection progress in a more detailed manner.

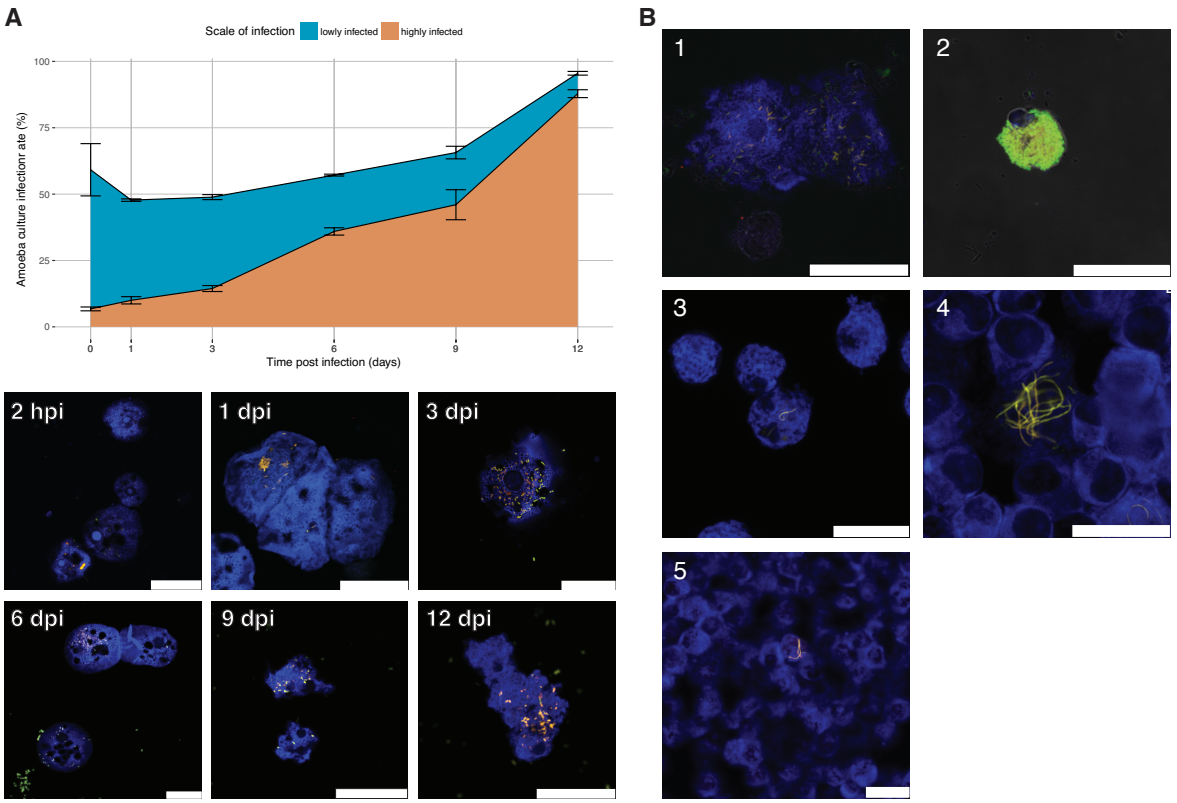


Figure 1
The (A) course of *Procabacter acanthamoebae* UWC12 infection cycle visualized as stacked plot and fluorescence in situ hybridisation pictures of the different stages. The blue part of the stacked plot resembles the lowly infected portion (less than eight *Procabacter* / cell) and the brown part represents the highly infected portion (more than eight *Procabacter* / cell). Six time points were taken: at 2 h, 1 d, 3 d, 6 d, 9 d, and 12 d post infection and host cell infection was determined using FISH staining of bacterial endosymbiont cells and general eukaryote staining for host cells. (B) FISH pictures of *Procabacter* infection of non-Acanthamoeba cells at 24 hpi. (B-1) Intact infected *Drosophila melanogaster* S2 cells; represent only the minority of eukaryotic cells. (B-2) Most of the host cells were highly infected with the endosymbiont and seem to lyse. (B-3,4,5) 24 hpi, 2 dpi and 3 dpi of a *Procabacter* infected human monocytic THP-1 cell. Note the filamentous morphology of the symbiont. The percentage of infected THP-1 cells remains low; no newly infected cells observed. Scalebar: 25 µm

Acanthamoeba castellanii is known to phagocytose extracellular particles within five minutes upon contact and reach the maximal efficiency after two hours (Avery, Harwood, and Lloyd 1995). At two hours after the initial *Procabacter* infection (2 hpi), 59.2% of the amoeba population was infected, with the majority of them (88.6% of all infected individuals) being lowly infected (Fig. 1A). However, the fluorescent signal of *Procabacter* was inhomogeneous, as some bacteria exhibited weaker FISH signals than others. Those cells could represent bacteria, which were unable to establish an infection within the amoeba. The subsequent washing step prohibited additional uptake of bacterial cells from the medium. 3D reconstruction with stacked FISH pictures (data not shown) indicates the symbionts location within vacuole-like subcompartments indicated by a bubble of faint Cy5-signal around the intracellular bacteria. At the next time point (1 dpi, Fig. 1A), the total percentage of infected cells dropped to 47.7 %, probably as a result of phagosomal digestion. This observation is consistent with the reported degradation time frame of food bacteria such as *Escherichia coli* K12, which happens within 24 h (Alsam et al. 2006). However, the fraction of lowly infected amoebae further decreased slightly at the next time point (Fig. 1A), whereas the number of highly infected cells increased slightly by 3.2 %. At this time point, *Procabacter* was sometimes observed as clusters of multiple cells (Fig. 1A; 1 dpi). The cluster formation of *Procabacter,* together with the rise of highly infected cells, points to multiplying *Procabacter* within the *Acanthamoeba* host. Between 3 dpi and six dpi, the total amount of infected cells stagnated around 50% and increased only slightly. However, the percentage of highly infected cells more than doubled. We could already observe some fully infected amoebae as well as extracellular *Procabacter*, probably released from lysed host cells. The accumulation of released *Procabacter*

will likely lead to infection of other amoebae by uptake and initiate new infection cycles, supported by a substantial increase in total infection at nine dpi, which peaked at 12 dpi with 95.5 % of the whole culture being infected.

Next, we analyzed how a *Procabacter* infection affects the host's growth (Supp. Fig .1). At 20°C, the infected amoeba population grew slower than the population without endosymbionts. However, this difference in growth rate was less pronounced when incubated at 30°C. It is unsurprising that the infection with *Procabacter* negatively impacts the host's growth, as the symbiont likely competes with the host for energy sources and metabolites. This observation agrees with the infection of other amoebal endosymbionts (Schulz et al. 2015; Benamar et al. 2017). The exact benefit of this symbiotic interaction for the host is unclear, but the advantages likely outweigh the energetic cost needed to engage in symbiosis. A recent study has shown the potential benefits of chlamydial symbionts in the presence of the pathogen *Legionella pneumophila* (König et al. 2019; Ishida et al. 2014). The defensive symbiont protects the amoeba host from lethal infection of pathogens at the cost of maintaining the symbiosis. In other symbiotic systems, such as insect-endosymbiont relationships, the symbiont commonly provides essential amino acids to the host. However, in protist-symbiont systems, this benefit has not yet been described. It seems unlikely, as free-living amoebae's diet consists mainly of prokaryotes or other smaller eukaryotes, and can easily obtain all amino acids. This stands in stark contrast to insects, which need a dietary supply, as nutrient values of their monotonous diet (e.g., plant sap) are limited (Moran 2007).

Next, we investigated potential host cells other than the original *Acanthamoeba* host. *Procabacter* cells were purified and added to *Drosophila* Schneider 2 insect cells (Fig. 1B 1-2; Macrophage-like Schneider 2; (I. Schneider 1972), HeLa cells, and human macrophage THP-1 cells (Auwerx 1991; Fig. 1B 3-4). *Procabacter* was able to infect S2 cells and, to a much lesser extent, THP−1 cells, both cell-lines with phagocytic abilities. S2 cells are very prone to *Procabacter* infection, as over 98% of the amoeba cells were fully infected after 24 h. However, the symbiosis was unstable, as most host cells began to lyse after 48 h incubation time (Fig. 1B-2). *Procabacter* has a different morphology while infecting THP-1 cells in comparison to infecting amoeba, and importantly only approximately 1 in 1000 THP-1 cells was infected (Fig. 1B 3-4). The data has shown that *Procabacter* can infect non-Acanthamoeba-cells, but a

stable long term symbiotic association could not be established. Most amoebal endosymbionts infect only within their natural host range. There have been reports of amoeba endosymbiont infection other than its original host species. However, the long-term interaction was not possible as it either results in host cell lysis or loss of endosymbionts (Schulz et al. 2014; Henning et al. 2007). Invasion of a cell and evasion of host-cell defense mechanisms are complex processes. The adaptation toward specific hosts might restrict the endosymbiont's ability to establish and survive in other host cell types (Sixt et al. 2012).

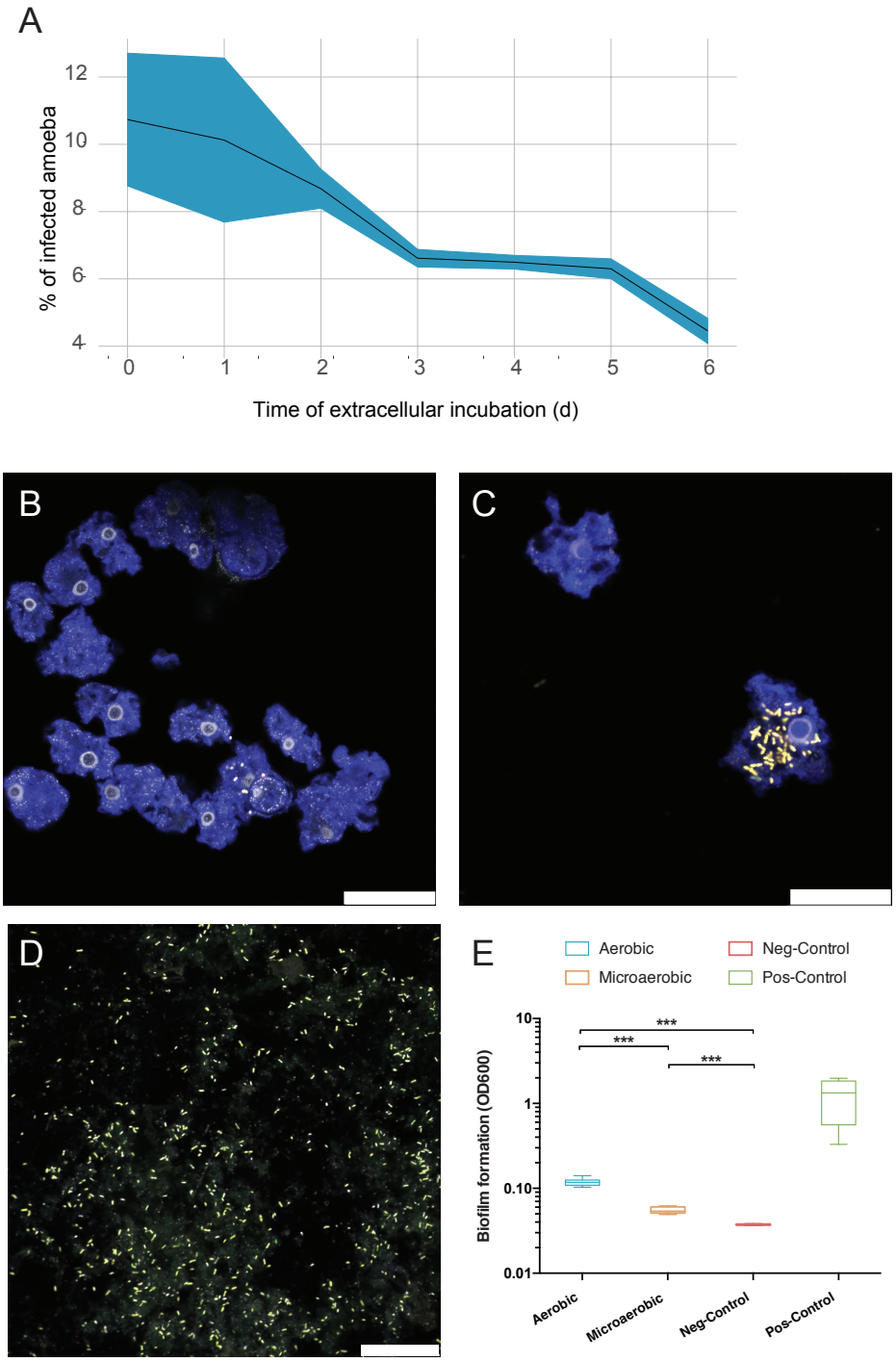


Figure 2
(A) Maintenance of infectivity of *Procabacter acanthamoebae* UWC12 in a host-cell free environment over six days. While *Procabacter* remained infectious, the infectivity-rate dropped as the time of extracellular incubation increased. (B-D) FISH pictures of *Acanthamoebae castellanii* infected with host-free incubated *Procabacter* (Euk.probe: Cy5; Bacterial probe: FITC; *Procabacter*-specific probe: Cy3). The host-free preincubation experiments were performed in Schneider's medium for 16 days under (B) aerobic and (C) microaerobic conditions. The infection rate of microaerophilic cultivated *Procabacter* was higher. (D) *Procabacter* biofilm formation under aerobic incubation conditions. The bacteria are active and dividing cells could be observed. (E) The result of the biofilm quantification assay shows significant higher biofilm formation under aerobic conditions. Scale bar: 25 µm

2. Maintenance of infectivity in a host-free environment

Obligate endosymbionts rely on their hosts' intracellular niche for replication. Upon leaving the host, the bacteria's survival depends on the encounter of another suitable host. Highly dynamic environments such as flowing aquatic habitats or disadvantageous situations (such as direct exposure to UV, unsuitable pH, and drought) either reduce the probability of meeting another host cell or has an impact on the symbionts' fitness, killing them before being able to establish another infection. Therefore, a symbiont's innate ability to survive in host-free environments directly impacts their propagation success. We analyzed *Procabacter's* ability to infect their natural hosts after prolonged incubation in a host-free and nutrient-rich environment (Fig. 2A). The infection rates are the highest at the no incubation and 24 h incubation time points and decrease continuously with progressing time points. Still after six days of host-free incubation, despite the efficiency drop to ~4.5 %, *Procabacter* remained infectious. We were unable to detect infected amoebae by fluorescence in situ hybridization (FISH) after seven days of host-free incubation (Fig. 2A).

Previous studies showed the successful axenic cultivation of endosymbionts, such as *Sodalis glossinidius* and *Coxiella burnetii,* by providing suitable conditions such as pH and oxygen-availability (C. Dale and Maudlin 1999; Omsland, Hackstadt, and Heinzen 2013). We examined the extent of *Procabacter's* infectivity maintenance without host cells by providing a medium with higher nutrient concentration (S2 medium) to reduce nutrient bottlenecks. Mimicking the low intracellular oxygen concentration of tissues, we additionally experimented with microaerobic conditions (Braun et al. 2001). After 16 incubation days, we observed *Procabacter* cell flocks under both oxygen-conditions. Both cultures kept their ability to infect their natural amoeba host (Fig. 2BC). However, under microaerobic conditions (Fig. 2C), the

cell-abundance and FISH signal intensity of the bacteria were higher, implying an overall higher activity of the cells. We observed biofilm formation and dividing *Procabacter* cells under aerobic incubation conditions (Fig. 2D), and the difference in biofilm formation between the two oxygen conditions was assessed by crystal violet staining assay. The amount of biofilm present is significantly higher (p-values < 0.001) under aerobic conditions compared to the microaerobic one (Fig. 2E). Bacterivorous protists like *Acanthamoeba* spp. are known to graze upon biofilms as these pose as an abundant food source (Hahn and Höfle 2001; Huws, McBain, and Gilbert 2005; Sherr and Sherr 2002). In the environment, the biofilm formation ability of *Procabacter* will increase its chance of coming in contact with potential host organisms and also further shape Procabacter's ability to cope with low oxygen levels.

3. Basic genomic features of *Procabacter* spp.

For better insights into *Procabacter* spp.'s interaction with the amoeba host, the molecular basis, and their underlying genomic repertoire, we created draft genome sequences for four Procabacter strains. The estimated genome completeness based on 182 conserved Proteobacteria marker genes of all *Procabacter* genomes was above 96.64% with no predicted contamination nor strain heterogeneity. The genome-sizes of *Procabacter* spp. are all in range of 2.44 Mb (Supp. Table 1), with OEW1 being the smallest (2.439 Mb) and CRCP the largest (2.478 Mb). The average G+C content is between 41-46% (Supp. Table 1). We identified a total of 42-43 tRNA genes, corresponding to all 20 amino acids. The number of coding sequences varies between 2,231 and 2,542, while the percentage of hypothetical proteins for all genomes is around 43%. Endosymbionts usually have a reduced genome in comparison to their free-living relatives, due to higher mutation rate and loss of metabolic functions and physiological capabilities (Nilsson et al. 2005). This phenomenon is well described for (primary-) endosymbionts of insects, where the genome reductions are quite extreme and usually accompanied by a low G+C content. In some cases, the genomes even undercut the 500 kb mark (McCutcheon 2010; Wernegreen 2002). The 1.2 Mb large *Babela massiliensis* genome is the smallest one among the sequenced obligate amoeba endosymbionts (Pagnier et al. 2015), whose genome size is usually between 1.8 Mb and 3 Mb (Chlamydiales, Rickettsiales, and Cytophagales). *Procabacter*s' genome sizes fall right within that range, whereas the G+C contents are higher than the genomes of other amoeba endosymbionts (< 40%). The lesser genome reduction of amoeba endosymbionts in comparison to primary endosymbionts of

insects is probably a reflection of their distinct lifecycles. At the final stages of their infection cycles, amoeba symbionts are usually released to the extracellular environment before infecting another host organism. Due to the complexity of the life stages and the need to retain certain activities in a host-free environment, the amoeba endosymbionts can only abolish a smaller number of unneeded genes.

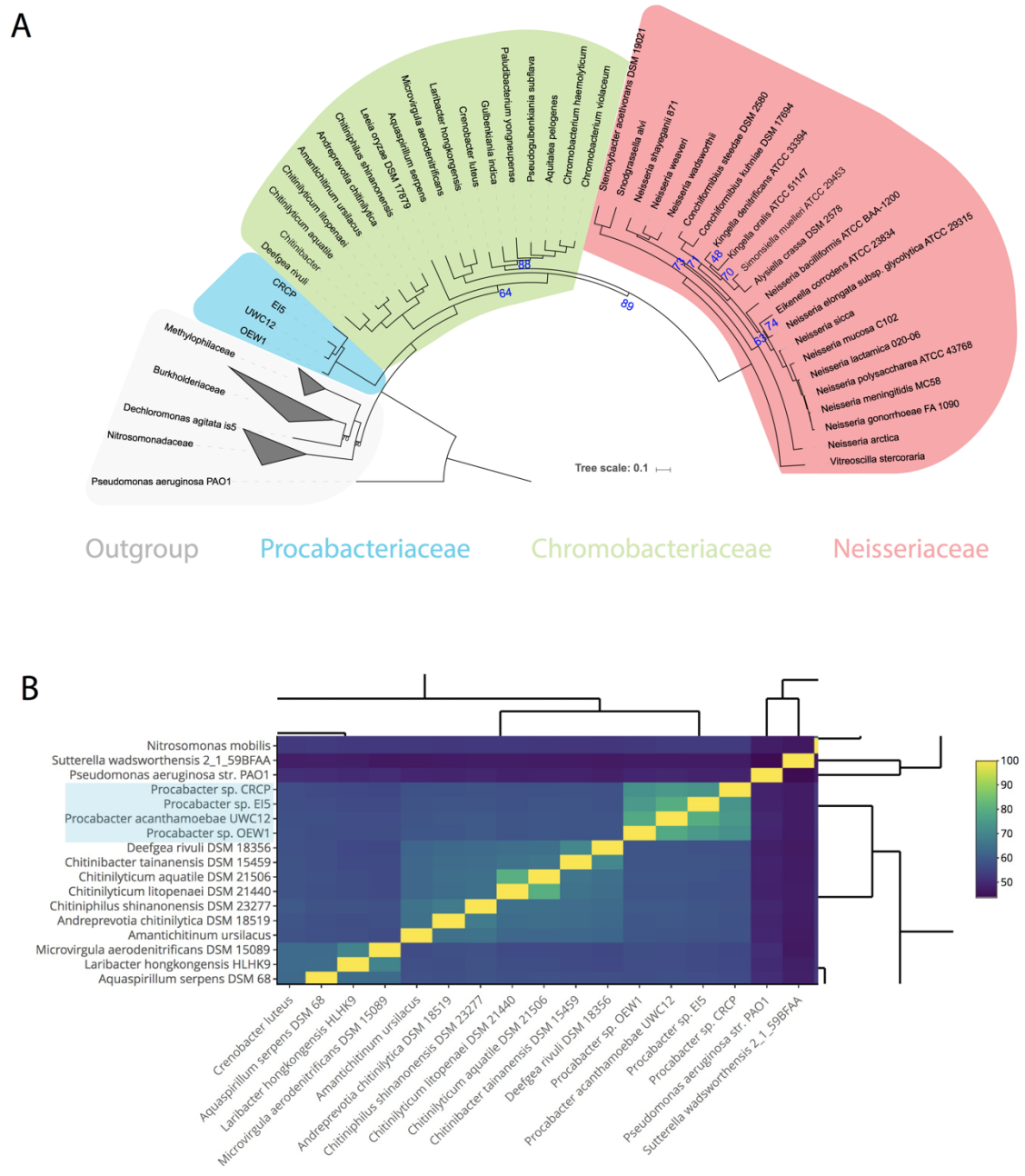


Figure 3
(A) *Procabacter* spp.'s. phylogenetic placement based on concatenated marker genes calculated with iqtree. *Procabacter* group marked in blue, Chromobacteriaceae members in green, Neisseriaceae members in red, and the outgroup in grey. Only confidence scores under 90 are

shown. *Procabacter* spp. form a monophyletic group at the root of the Chromobacteriaceae / Neisseriaceae branch. (B) Taxonomic classification based on average amino-acid identity (AAI). Higher similarity values within the *Procabacter* group (72% - 80%) in comparison to the most similar neighbouring organisms, several members of the Chromobacteriaceae with around 58% similarity. This provides further support for *Procabacter* spp. as its distinct group.

3.1 Phylogeny of *Procabacter*

Previous 16S rRNA genes based phylogenetic study has placed *Procabacter* in the own order of Procabacteriales (Horn et al. 2002). With the availability of new genomic data, we attempt to further infer *Procabacter's* phylogenetic relationship with the neighbouring groups by using a concatenated alignment of marker genes (Fig. 3A). Also, we compared the genomes based on their average amino acid identity (AAI; Hausmann et al. 2018; Fig. 3B).

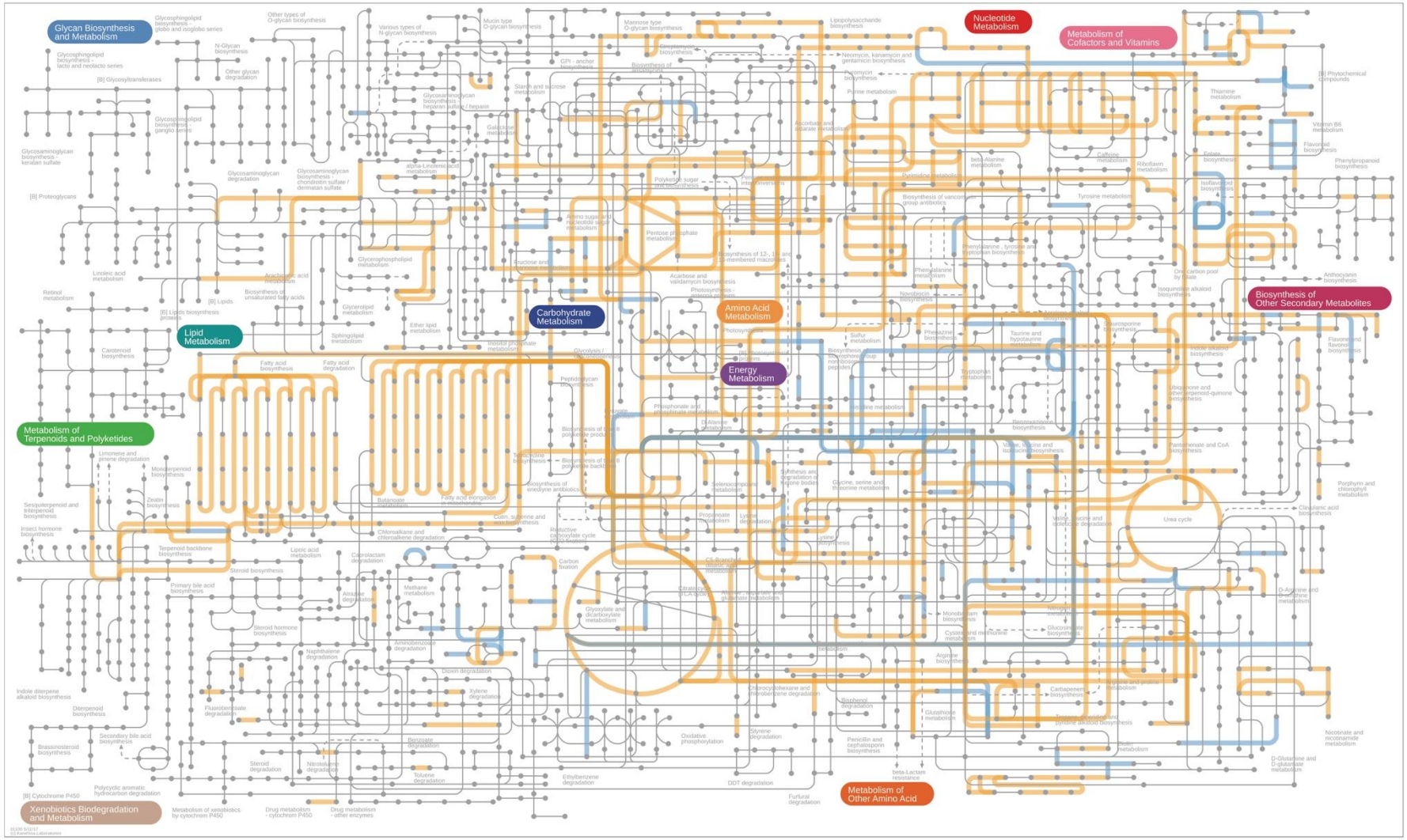
The phylogenetic analysis (Fig. 3A) places all four *Procabacter* species as a monophyletic group at the base of the former Neisseriales group, now reclassified and split into the families Neisseriaceae and Chromobacteriaceae under the order Burkholderiales (Parks et al. 2018). Other changes concerning this sub-branch include the reclassifications of the genera *Aquaspirillum*, *Chitinibacter, Leeia*, and *Chitinimonas* (formerly placed under the family Burkholderiaceae) as the families Aquaspirillaceae, Chitinibacteriaceae, Leeiaceae and Chitinimonadaceae (Supp. Fig. 2AB). The *Procabacter* group is a sister clade to the Chromobacteriaceae and Neisseriaceae and clearly differs from the outgroup. The AAI data further supports the phylogenetic results, as the *Procabacter* spp. shows higher identity with Chromobacteriaceae members (56 % - 58 %) than Neisseriaceae members (51% to 53%; Fig. 3B; Supp. Fig. 2C). The AAI value within the *Procabacter* group ranges between 72 % - 80 %. Both methods predicted the Chromobacteriaceae bacteria*Deefgea rivuli, Chinitibacter tainensis, Chinitilyticum* spp., *Andreprevotia* sp., and *Chitiniphilus* sp. as the closest relatives (Fig. 3 ABC). Taken together, our data present support for the genus *Procabacter* as a monophyletic group. In agreement with the newly proposed standardized bacterial taxonomy (Parks et al. 2018), we propose the placement of the Procabacteriaceae family within the order Burkholderiales.

3.2 Minorly reduced metabolic capability of *Procabacter*

Because the intracellular habitat acts as a readily available energy source, selection pressure on symbionts to retain their full metabolic capability is relaxed. The genes affected are usually involved in metabolic pathways or the biosynthesis of amino acids and metabolites, as those

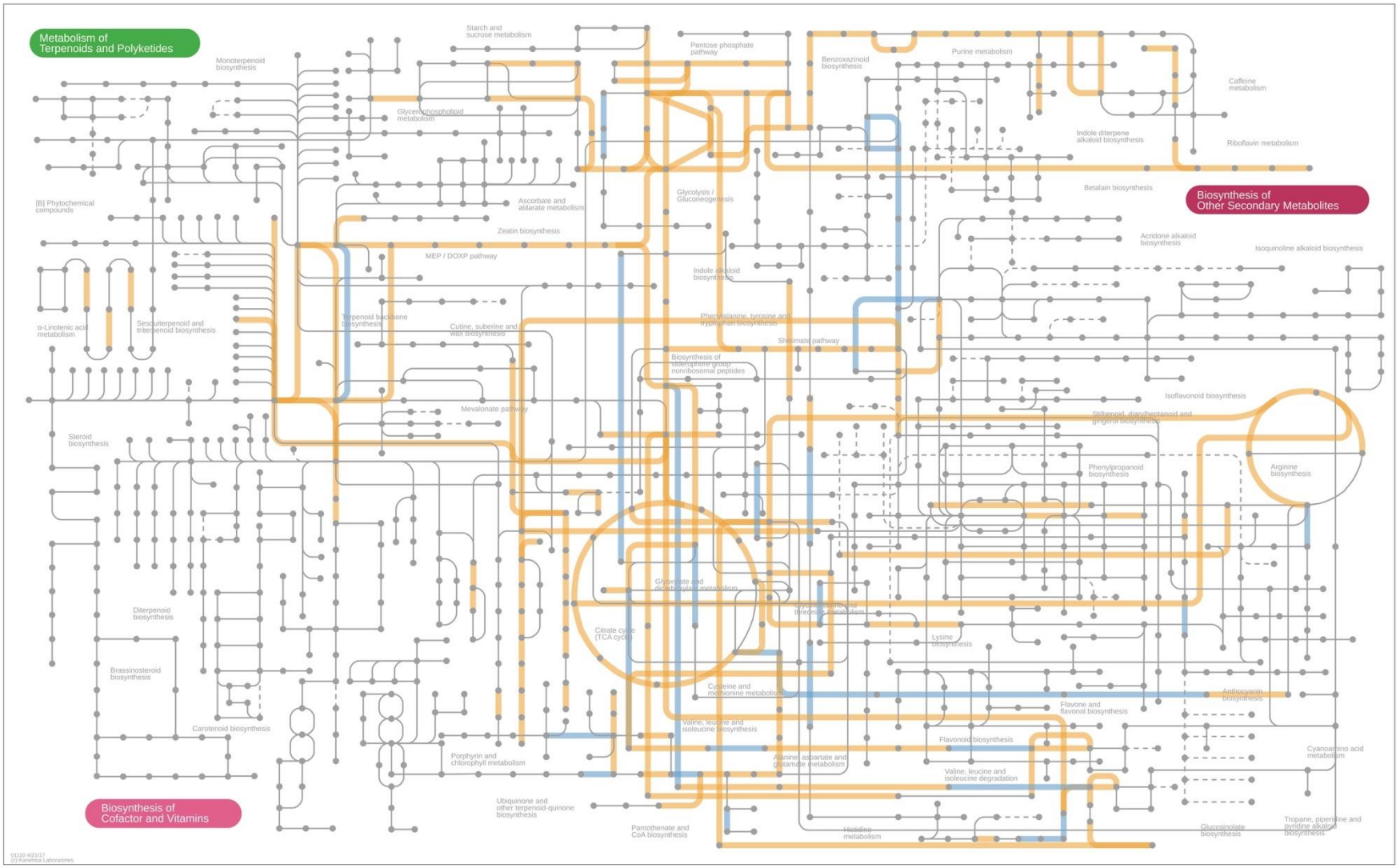
can be easily obtained from the host. As a result, genome size reduction is a prominent feature of endosymbionts, as the nonvital genes are reduced without compromising core functions (Moran 2002; McCutcheon and Moran 2011).

A


Metabolic pathways
Glycan Biosynthesis and Metabolism
Lipid Metabolism
Metabolism of Terpenoids and Polyketides
Xenobiotics Biodegradation and Metabolism
Carbohydrate Metabolism
Amino Acid Metabolism
Energy Metabolism
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Biosynthesis of Other Secondary Metabolites
Metabolism of Other Amino Acid

Secondary metabolites

B



Procabacter spp. core metabolic pathways

Procabacter spp. accessory metabolic pathways

Figure 4
A map visualizing (A) metabolic pathways and (B) secondary metabolite pathways found in *Procabacter* spp. genomes based on the proteome's KEGG Orthology annotation. The core fraction (orange) represents reactions found in all *Procabacter* genome. Blue marks the accessory fraction, reactions present in some genomes. The main energy-generating pathways are conserved in all *Procabacters*, with degradation in amino acid synthesis pathways, and truncated lipid metabolism. Different species vary in amino-acid and metabolite synthesis capabilities.

The *Procabacter* genomes encompass a surprisingly diverse metabolic capability, and the majority of genes in the predicted metabolic pathways are present. This includes glycolysis and gluconeogenesis, the whole citrate cycle, the pentose phosphate pathway, purine and pyrimidine metabolism, lipopolysaccharide biosynthesis, heme production, and ubiquinone biosynthesis (Fig. 4, Supp. Table 2 - Sheet 1). *Procabacter* spp. core metabolism pathways (present in all genomes) are quite conserved (Fig. 4AB) and the accessory pathways (pathways not present in all four genomes) are mainly involved in the synthesis of various amino acids (Fig. 4AB, Supp. Table 2 - Sheet 2). OEW1 retains almost the full array of all 20 amino acids, lacking only the capability to synthesize cysteine, histidine, and methionine. In contrast, EI5 and CRCP both are only able to synthesize about half of all amino acids. We also compared the percentage of proteins categorized as metabolism-related in COG (clusters of orthologous groups) categories (C, E, F, G, H, I, M, P, and Q) to other endosymbionts of amoeba and species of neighbouring family (Supp. Table 3). With around 30% of all proteins dedicated to metabolism, *Procabacter* spp. are similar to other known amoebal endosymbionts and are lower than the free-living members of Chromobacteriaceae and Neisseriaceae, which all range around 40%. However, among the 30%, *Procabacter* spp. have double the percentage of proteins classified to the category "H" (Coenzyme metabolism) than other amoebal endosymbiont. Based on the metabolic reconstruction, we see species-specific diversity in the ability to synthesise amino acids and suggest that amino acids are *Procabacter* spp. main metabolic bottleneck, relying on the uptake from the host.

4. *Procabacter* host cell interaction

A vital aspect of intracellular bacteria is the interaction with the host cell. The modi of interaction are very diverse, ranging from initial adhesion to habitation inside the intracellular space. Host interaction is generally facilitated by means of effector proteins,

usually either transported from the bacterial cytoplasm directly to the host or indirectly released into the environment.

4.1 Adhesion facilitating proteins

The adherence of symbionts to potential host cells is an essential first step in establishing an infection and colonization of cells. A range of extracellular structures or proteins such as pili/fimbriae and adhesins aids the bacteria in enhancing their host-attachment ability (Pizarro-Cerdá and Cossart 2006). One machinery involved in adhesion is the type 5 secretion system (T5SS). Different from the T3SS, which spans both inner- and outer membranes, the proteins secreted via the T5SS need to cross each membrane individually. All five subtypes of T5SS (T5a-eSS; (Leo, Grin, and Linke 2012) consist of two crucial parts, the translocation and the passenger domain. The translocator-domain protein forms a pore in the outer membrane and facilitates the secretion of the passenger-domain containing protein. Because T5SS facilitates its own translocation to the environment, it is also termed as "autotransporters". All *Procabacter* spp. genomes contain T5SS operons of the subtype A, B, and C (Supp. Table 4). It includes homologs of the ShlA/-B system (Cytolysin and its translocator of *Serratia marcescens*; part of TPS: Two-partner secretion pathway), IcsA/-B system of *Shigella flexneri*, AidA adhesins (T5aSS) and YadA-domain containing proteins (T5cSS). Although the different subtypes vary in size and functional domains, they all facilitate better adhesion to the host cells. Known for their enormous sizes of up to 5,600 residues, ShlA-like proteins, often described as filamentous haemagglutinin, have various functions important for pathogenesis, such as attachment to host cells, aggregation, and cell killing (Rojas et al. 2002; Kida et al. 2008). The UWC12 and CRCP ShlA-homologs are considerably larger (> 2,300 residues) than the EI5 and OEW1 homologs (~550 residues)(Supp. Table 4). *Procabacter's* T5cSS proteins contain YadA-domains and are smaller than 500 amino acids in length, ranging mostly around 300 aa. In comparison, the *Procabacter* T5aSS proteins have a longer average length (1079 aa) and contain AidA-repeats. AidA adhesins, described in *E.coli*, are directly anchored in the outer membrane and function in autoaggregation and biofilm formation (Wells et al. 2007).

We have also found another type of protein aiding adhesion. All four genomes have a σ – fimbriae operon, containing the Spore-Coat-Protein U domain (COG5430), the chaperone (with domain COG3121), the usher protein and the tip adhesin (Supp. Table 4).

The σ-type of non-flagellated filaments is part of the chaperone-usher assembly pathway and usually involved in biofilm formation and cell adhesion(Tomaras et al. 2003; Nuccio and Bäumler 2007). We speculate that the σ–fimbriae aid Procabacter cells in either their uptake by the amoeba host by either increasing the ability to attach themselves to the host cell or to generate biofilms on surfaces in their natural environments, which leads to amoeba grazing and subsequent contact with potential host cells (Parry 2004).

4.2 *Procabacter* spp. do not have a flagellum

One genomic difference between the members of Chromobacteriaceae and Neisseriaceae is the presence of flagella genes in the genome of the former ones (Adeolu and Gupta 2013). We did not detect any typical flagellum marker genes in the *Procabacter* genomes, a common trait with the members of the Neisseriaceae. Procabacter's lack of flagellum could be the result of a lifestyle specific adaptation to the intracellular niche, as the energetic cost of flagella expression and maintenance is high. In the cytoplasm, high motility is not needed and the flagellum expression becomes too energetically expensive, leading to the reduction of responsible genes as a result. This phenomenon can be observed in other endosymbiotic bacteria such as *Buchnera aphidicola*, *Blochmania* spp., and *Baumania* spp. (Maezawa et al. 2006; Gil et al. 2003; Toft and Fares 2008).

4.3 *Procabacter* spp. encodes for an arsenal of pathogenic factors

All four *Procabacter* spp. encode for two independent T1SS operons. Two out of the three components are likely regulated together, namely the ABC-transporter and the membrane-fusion protein (MFP) component, whereas the porin-encoding gene is located somewhere else in the genome. The genomes also encode for a type III secretion system (T3SS), a syringe-like protein machinery used for direct delivery of effectors into the target cell (Supp. Table 5; Sheet "T3SS"). T3SS are mainly distributed in members of the phyla Proteobacteria and Chlamydiae (Hu et al. 2017). The majority of the research was centered around the role of these translocation machineries in pathogenic organisms like *Salmonella*, *Yersinia*, *Chlamydia*, *Escherichia*, *Xanthomonas*, *Burkholderia,* and *Ralstonia* (Abby and Rocha 2017). Within the order Betaproteobacteriales, most species encoding one or multiple T3SSs belong to the family Burkholderiaceae (Hu et al. 2017). For the members of the phylogenetically closer

related family Chromobacteriaceae, only *Chromobacterium* spp., *Paludibacterium* spp., and *Pseudogulbenkiania* spp. encode T3SSs in their genome. *Chromobacterium violaceum* even has multiple T3SS operons with different functions, a trait shared with some Burkholdericeae members (e.g., *Burkholderia ambifaria* and *Burkholderia andropogonis*) (Miki et al. 2010; Holden et al. 2004; Hu et al. 2017). The rest of the Chromobacteriaceae and Neisseriaceae members do not encode T3SS but mainly rely on T1SS, Type IV secretion systems (T4SS), and Type V secretion systems (T5SS; Quillin and Seifert 2018). We used the tool EffectiveDB (Eichinger et al. 2016) to predict putative proteins secreted by the T3SS. The pipeline combines the analysis of Type III secretion signals, conserved chaperone binding sites, and searches for eukaryotic-like domains (ELDs) for the prediction. We could identify over 355 putative T3SS effectors (EffectiveT3 score > 0.99), more than 45 proteins with a chaperon around the expected region, and an arsenal of proteins with ELDs in the genomes of *Procabacter* spp. (Supp. Table 5). In comparison to Neisseriaceae and Chromobacteriaceae, the *Procabacter* spp. have a significantly higher percentage of predicted T3SS effector and ELD-containing proteins (Fig. 5, Supp. Table 5). ELDs occur in eukaryotic genomes and are more commonly found in bacteria with a host-associated lifestyle than a non-host-associated lifestyle. These ELD-containing proteins might facilitate host-interaction and are therefore speculated to be effectors. We made the same observation of a high quantity of putative effector proteins on the genomes of other amoeba symbionts (Schmitz-Esser et al. 2010; Schulz et al. 2015; Domman et al. 2014).

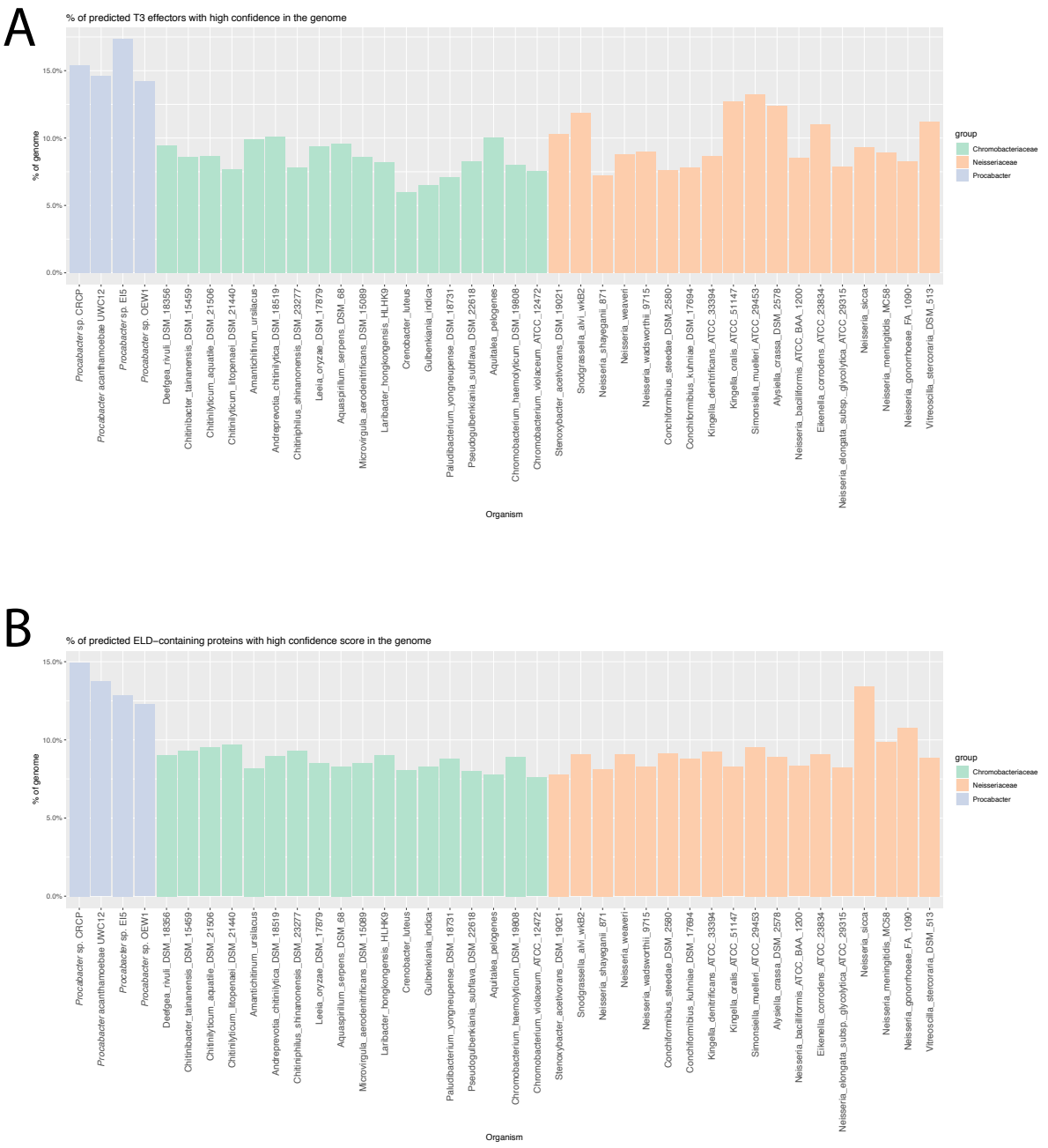


Figure 5

The percentage of (A) Type 3 secretion system effector proteins and (B) eukaryotic-like domains containing proteins in the genomes. *Procabacter* genomes are highlighted in blue, Chromobacteriaceae genomes in green, and Neisseriaceae genomes in orange. *Procabacter* genomes contain a higher percentage of potential effectors in comparison to members of the neighboring groups.

4.4 Presence of two different cytochrome types on the genome

All *Procabacter* genomes encode for two types of cytochrome oxidases, namely the cytochrome-c-oxidase (*Cox*; *cbb3*-type) and a cyanide insensitive quinol (CIO) subtype of cytochrome bd oxidase. Although both oxidases are the terminal enzymes of the respiratory electron transport chain, they are phylogenetically unrelated (Giuffrè et al. 2014). The heme-copper oxidase *Cox* uses copper in its catalytic center, whereas the cytochrome bd uses only hemes as cofactors. The canonical cytochrome bd oxidase usually has high oxygen affinities (Borisov et al. 2011), whereas the CIO has been attributed with a much lower affinity (Miura et al. 2013; Hirai et al. 2016; Jo, Price-Whelan, and Dietrich 2014). Interestingly, phylogenetic analysis places the *Procabacter* CIO subunits among homologs of Alphaproteobacteria or Gammaproteobacteria. The majority of closely related homologs either belong to Alphaproteobacteria or Gammaproteobacteria, with the subunits of the *Methylobacterium* spp. being the closest related, with only two homologs found in members of the Betaproteobacteria (Supp. Fig. 3).

This observation suggests either a total loss of this oxidase type from the majority of the Betaproteobacteria or the horizontal acquisition of the *Procabacter* CIO subunits, highly likely from the Alphaproteobacteria (Supp. Fig. 3). Cbb-3 contributes to the virulence of some pathogens, increasing the survival rate in low oxygen environments and resisting oxidative/nitrosative stress (Baughn and Malamy 2004; Giuffrè et al. 2014). The occurrence of more than one respiratory oxidase, each with different substrate affinity and stress tolerance has already been described (Kaminski et al. 1996; Arai 2011; Dalsing et al. 2015; Chang et al. 2010; Poole and Cook 2000). Having different terminal oxidases allows the bacteria to grow under various conditions or occupy a more diverse environmental niche regarding the oxygen concentration, and a putative adaptation to the intracellular lifestyle.

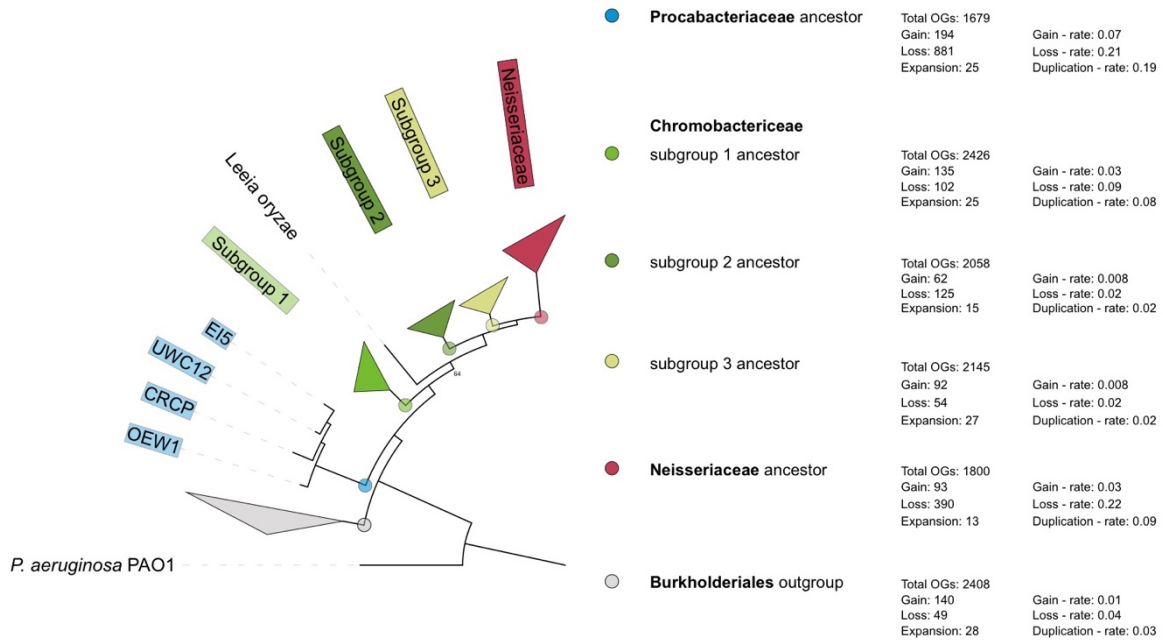


Figure 6
Protein-family gain and loss at different ancestral node. The *Procabacter* branch has lost a high number (881) of protein families at a high loss-rate (0.21) in comparison to other sister lineages, suggesting ongoing genome reduction. Gained protein-families contain groups of effector-like proteins and genomic mobile elements, also show elevated gain-rates in comparison.

5. *Procabacter's* ancestral genome suggests evolution towards endosymbiotic lifestyle

Procabacter's phylogenetic placement is at the root of the former order Neisseriales, which is now split into the families Neisseriaceae and Chromobacteriaceae under the order Burkholderiales. It is especially interesting, because the *Procabacter* group is the closest sister group consisting only of endosymbionts. Although some pathogenic members of the families Neisseriaceae and Chromobacteriaceae are able to invade the intracellular niche, neither of them are considered strictly intracellular. We used the software Count (Csurös 2010) to reconstruct the *Procabacter* ancestral genome state and infer gene loss and gain rates. The ancestor's genome has gained 194 (gain-rate: 0.07), expanded 25 and lost 881 gene families in comparison to the last common ancestor (loss-rate: 0.21) (Fig. 6). Compared to the subgroups of Chromobacteriaceae, the gene-family-loss rate of the *Procabacter*-ancestor is high (Fig. 6). The high loss-rate combined with the high quantity of lost gene families indicates a rather rapid than gradual process. In contrast, the Neisseriaceae ancestor has lost

only a moderate number of gene families (76 lost families; loss-rate: 0.22). Another noteworthy observation is the relatively large number of gained gene families in the *Procabacter* branch. 122 of the 194 gained gene families are *Procabacter* specific. We could assign functions to 76 gained gene families and 10 (of 25) expanded gene families either based on pre-annotated information (uniprot) or de novo annotation based on the internet tool Eggnog-Mapper (Huerta-Cepas et al. 2016). Most of the gained groups are classified as the eggnog category "S" (Function unknown), followed by "M" (Cell wall/ membrane/ envelope biogenesis), "L" (Replication, recombination and repair), and "K, T, U" (Transcription, Signal transduction, and secretion). However, the functional annotation revealed the presence of mobile genetic elements like transposases, prophages, and integrases, which are commonly found in genomes undergoing rearrangements. 10% (8 of 76; Supp. Table 6; marked in yellow) of the gained gene families are toxin-Antitoxin systems. Usually, these systems tend to accumulate within bacterial genomes, due to the more labile nature of the antitoxin in comparison to toxins, which selects against cells without both partners. Some of these systems are also responsible for managing stress regulations (Wang and Wood 2011). Interestingly, 21% (16 out of 76 OG; Supp. Table 6; marked in green) of these annotated gene families are annotated as proteins containing eukaryotic-like domains (e.g. ankyrin-repeats, leucine-rich-repeats, tetratricopeptide-repeats, pentatricopeptide-repeats, SEL1-domains or NACHT-domains), type III effector proteins or type III secretion system chaperones. Pathogenic bacteria often interact with the host cell through the means of effector proteins. Bacterial type III secretion effectors mimic eukaryotic proteins to manipulate the target eukaryotic cell with the benefit of cellular defense reduction, which further leads to increased bacterial virulence (Dean 2011). Endosymbiotic bacteria face similar host defense mechanisms and therefore often harbor a wide array of eukaryotic-like domain containing-effectors to localize the intracellular niche (Abby and Rocha 2017; Colin Dale and Moran 2006; Díez-Vives et al. 2017). In addition of the repeats-containing proteins, we also found other pathogenesis-related proteins as components of the type III secretion system, bacteriocins, adhesion proteins, lysozymes, two-component systems, quorum sensing regulators, serine proteases, autotransporters, translocons or mediator of bacterial evasion from the host cells among the gained proteins (Supp. Table 6). Metabolically wise, compound transporters and various transferases are also among the list. The predicted functions of the expanded families are similar to the gained gene families, e.g. proteins containing eukaryotic-like domains

(leucine-rich-repeats, tetratricopeptide-repeats), type 3 secretion system chaperone, adhesion protein, toxin, and some drug resistance proteins (Aminoglycoside phosphotransferase; Supp. Table 7).

Overall, these data suggest that the ancestor of Procabacteriaceae/ Chromobacteriaceae/ Neisseriaceae first divided into two major lineages. One is the free-living Proto-Chromobacteriaceae-Neisseriaceae lineage which further branches off into the individual Chromobacteriaceae and Neisseriaceae (Fig. 3). The other lineage, where the *Procabacter* spp. belong to, has abandoned the free-living lifestyle and evolved towards endosymbiosis. Early members of this lineage showed hallmarks of endosymbiotic genomes with loss of genes, not essential for the survival in the intracellular space, and gains and/or keeping genetic elements that are beneficial for host cell interaction.

6. Conclusion

The monophyletic group of *Procabacter* consists of members only found as endosymbionts in *Acanthamoeba* hosts. Phylogenetic analysis revealed a close relationship to the families Neisseriaceae and Chromobacteriaceae, both groups comprised of free-living organisms and pathogens but not obligate endosymbionts. Genome analysis of *Procabacter* spp. has shown genomic hallmarks also found in genomes of other amoeba endosymbionts. Like other amoeba endosymbionts, *Procabacter* encodes secretion systems and probably rely on them for host cell communication. Until today, there is no indication about secretion system preferences among the amoebae endosymbionts, it varies between T2SS, T3SS, T4SS and T6SS (Böck et al. 2017; Schulz et al. 2015; König et al. 2017; Pagnier et al. 2015). Like the members of Chlamydiae, *Procabacter* encodes a type III secretion system. Despite the secretion system discrepancies between the various species, most of them do encode for an arsenal of effector-like protein, a trait also shared by *Procabacter* (Schmitz-Esser et al. 2010; Schulz et al. 2015; König et al. 2017). Where *Procabacter* differs from the rest, is the conservation of a great fraction of genomic repertoire for various metabolic pathways, unexpected from an obligate endosymbiont. Infection experiments suggest while *Procabacter* can survive outside of a host cell for a prolonged period, it ultimately still needs a suitable host for replication. This is likely the result of nutrient limitations, as we could observe dividing *Procabacter* cells in a host-free cultivation attempt with a nutrient-rich medium. We speculate that the group of *Procabacter*

spp. represents a clade of the relative recent endosymbionts. The reconstruction of the ancestral history shows the accumulation of effector proteins, likely used for host-interaction. Further, the gain of several mobile genetic elements in combination with the high gene lost-rate of this clade, suggests the undergoing reduction of Procabacter genomes.

Materials and Methods

Cultivation of Acanthamoeba hosts infected with and without *Procabacter* spp.

We used four different *Acanthamoeba* spp. infected with *Procabacter* spp. combinations for this study. For all four cultures, *Acanthamoeba* sp. containing *Procabacter* sp. endosymbionts were cultivated at 20°C in 8 mL Tryptone Soy Yeast Extract Broth (TSY; 30 g/L Trypticase Soy Broth, 10 g/L Yeast extract), a nutrient-rich growth medium. Cultures were maintained in cell culture flasks (Nunclon Delta-surface, Thermo Scientific, St Leon Rot, Germany) and the medium was exchanged every seven days.

Growth assessment of infected/non-infected *Acanthamoeba* sp.

We assessed the infection's impact on the host's growth by comparing the growth rate between infected and uninfected *Acanthamoeba* sp. UWC12. For the curing process, we incubated infected *Acanthamoeba* sp. UWC12 in TSY-Rifampicin medium (Final concentration: 100 µg/mL; filtered with 0.22 µm ø pore size) to remove UWC12 from the system. We chose two temperatures for our experiments; 20°C and 30°C. For each condition, two sets of two uninfected *Acanthamoeba* sp. UWC12 cultures with the same parent culture were incubated. For each setup, we infected one flask with UWC12 for a fresh infection and waited for seven days to let all cultures acclimate to their new conditions. We seeded 1000 amoebae from uninfected and infected cultures at both temperatures in duplicates into a well of a 12-well-plate. For each time-point and well, ten random pictures were taken with a camera at the inverted light microscope and the average number of attached amoeba cells was determined using the cell-counter tool of the software ImageJ (C. A. Schneider, Rasband, and Eliceiri 2012). The values were then averaged to get a mean cell density per well per time-point. We screened for contamination by FISH and estimated the correct infection rates (0 and 100 percent respectively) after the last time-point. To estimate the doubling time, first we calculated the intrinsic rate of increase $r_m = ln\frac{\frac{N_t}{N_0}}{t}$ (Dixon, 1987) and then the doubling

time was subsequently estimated with $PDT = \frac{ln\,(2)}{r_m}$ where N0 and Nt represent the amoeba cell numbers at the start of the experiment and time point t, respectively.

Purification of extracellular and intracellular *Procabacter acanthamoebae* UWC12 The intracellular fraction of *P. acanthamoebae* UWC12 was harvested by collecting amoebae containing *P. acanthamoebae* from a well-grown (mono-layer) culture in a culture flask. The cells were centrifuged (2600 g, 6 min, RT) in a 50 mL greiner tube and resuspended in eight mL of a medium of choice. We used a dounce homogenizer with the tight glass pestle (15 times) to break up the amoeba cells and release the intracellular bacteria. The homogenized suspension was filtered (ø: 5 µM) to remove cell debris and unlysed amoebae.
The extracellular fraction of *P. acanthamoebae* UWC12 was collected from a running culture's supernatant and filtered twice (ø: 5 µm) to remove larger cell debris. Afterward, the endosymbionts were pelleted by centrifugation (2600 g, 6 min, RT) resuspended in fresh medium.

Infection cycle of *Procabacter acanthamoebae* UWC12 in *Acanthamoeba* sp. UWC12

Axenic (endosymbiont-free) amoebae cultures of the original hosts were cured with TSY-Rifampicin (100 µg/mL) and exchanged every four days. In order to minimize cross-infection-spread from uncured amoebae, upon mono-layer formation, amoebae were detached and 100 µL suspension was passaged to a new flask to start a new culture. After each transfer, we used fluorescence in situ hybridization with a *Procabacter* specific probe to monitor the loss of infection rate. Once cured, axenic cultures were maintained separately from the infected ones to avoid contamination.

For an infection experiment, uninfected acanthamoebae were detached from the surface of a cultivation flask and pelleted by centrifugation (2600 g, 6 min, RT) in a 50 mL centrifugation tube. The supernatant was decanted and the pellet resuspended in 10 mL fresh medium. We used a hemocytometer (Neubauer counting chamber) to determine the number of amoeba in the suspension. Based on the settings, 3x10^4 cells were seeded in a well of a 12-well multidish or 5x10^6 cells into a 12.5 cm^2 culture flask.

Extracellular *Procabacter* cells MOI 200 were added to the uninfected *Acanthamoebae* and briefly centrifuged down (1350 rpm, 10 min, RT) to enhance symbiont-host-contact and facilitate a synchronous infection. The culture was kept at 20°C, washed after 2 h by removing the supernatant to remove all non-phagocytized *Procabacter* cells and fresh media were added. We monitored the infection process up to 12 days post infection with the time-points 2 h, 1 d, 3 d, 6 d, 9 d, and 12 d. At each time-point, a sample was taken and visualized by Fluorescence in situ hybridization (FISH; Daims et al., 2004). Host cells containing less than eight *P. acanthamoebae* were considered as lowly infected and amoebae containing more than eight endosymbionts as highly infected. Duplicates were performed for each time-point.

Fluorescence in situ hybridization (FISH) for cytoplasmatic *Procabacter* visualization

We performed FISH in combination with confocal laser scanning microscopy to visualize the infection process of UWC12. The host amoeba was centrifuged (6600 rpm, 6 min, RT), washed, and then resuspended in 1x page's amoeba saline solution (PAS; 0.12 g/L NaCl, 0.004 g/L MgSO4*H2O, 0.004 g/L CaCl*2H2O, 0.142 g/L Na_2HPO_4, 0.136 g/l KH_2PO_4). 20 µL of the suspension was applied to a well on a FISH-slide. Amoebae were allowed to attach for 20 min, and the well was washed once with PAS, followed by fixation with 4% formaldehyde (12 min at room temperature). The hybridization (Wagner et al. 2006) was performed at 20% formamide-concentration for 2 h at 46°C using the standard hybridization and washing buffers. We used the *Procabacter* genus-specific probe Proca-438 (5'- CGA TTT CCT CCC RGA CAA -3'; (Horn et al. 2002) in combination the probe mixture EUB338 I-III (5'-GCTGCCTCCCGTAGGAGT-3', 5'-GCAGCCACCCGTAGGTGT-3', and 5'-GCTGCCACCCGTAGGTGT-3'; Amann et al. 1990; Daims et al. 1999) targeting most bacteria , and the eukaryotic probe EUK516 (5'-ACCAGACTTGCCCTCC-3'; Amann et al. 1990) for better contrast of the host. Slides were embedded in Citifluor prior to examination with a confocal laser scanning microscope (LSM 510 Meta, Zeiss, Oberkochen, Germany) or an epifluorescence microscope (Axioplan 2 Imaging, Zeiss, Göttingen, Germany) equipped with a CCD camera (AxioCam HRc, Zeiss, Göttingen, Germany).

Infectivity maintenance in an extracellular environment

We analysed *Procabacter's* ability to maintain its infectivity without the presence of a host. Extracellular *Procabacter* cells were harvested and cultivated in a host-free environment in

TSY at 20°C in a 12 well-dish. For six continuous days, every 24 h we've added axenic *Acanthamoeba* sp. UWC12 (MOI 250) cells to two wells (duplicate) of incubated *Procabacter* and washed the culture with PAS after 2 h to stop reinfections with remaining symbionts in the supernatant. At 48 hpi of each infection, we've measured the percentage of infected *Acanthamoeba* sp. UWC12 by using FISH as described above.

Host-range assessment of *Procabacter acanthamoebae* UWC12

In addition to other *Acanthamoeba* spp., we tested the ability of UWC12 to infect *Drosophila* Schneider 2 (S2) insect cells, HeLa cells, and THP-1 cells. *Drosophila* Schneider 2 (S2) cells were maintained at 27°C (ambient atmospheric conditions) in five mL Schneider's medium (SM) in 25 cm^2 sterile tissue culture flasks with filter lids to facilitate gas exchange. The medium was supplemented with 10 % (v/v) fetal bovine serum (FBS, heat inactivated at 56°C for 30 minutes; Life Technologies, Carlsbad, United States of America) and sterilized by filtering (0.2 µm pore size). The medium was always preheated to 27°C before passaging cells. For both HeLa and THP-1 cells, we used pre-warmed (37°C) Roswell Park Memorial Institute 1640 medium (RPMI 1640; Life Technologies, Carlsbad, United States of America) amended with 25 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), 2.05 mM L-Glutamine and 10% (v/v) fetal bovine serum (FBS, heat inactivated at 56°C for 30 minutes; Life Technologies, Carlsbad, United States of America) as growth medium. The cultures were routinely checked for contaminants by FISH with the eukaryotic probe EUK516 in combination with the general bacteria probe. The cultures were split when the cells have formed a monolayer on the flask's bottom. We first removed the medium and used 1x DPBS (without Mg and Ca) to wash the cells. 0.5 mL / 10 cm^2 of 0.5% Trypsin-EDTA were carefully spread across the cells and incubated at 37°C for 10 min to detach the HeLa cells. 5 mL RPMI + 10% FBS was added to the detached cells and 0.5 mL of it were transferred to a new flask containing 5 mL of preheated RPMI medium + 10% FBS.

We performed the UWC12 infection experiments for S2, HeLa cells and THP-1 cells with MOI 150 in cell culture flasks. The infection was stimulated by centrifugation (10 minutes, 370 g, 27°C) to enhance the interaction between bacterial and insect cells. We performed a washing step with PAS to remove bacteria from the supernatant at 24 hpi to exclude extracellular

bacteria from starting an infection. For each time-point, the infection progress has been monitored by FISH and analyzed using fluorescence microscopy.

Axenic cultivation of *P. acanthamoebae* in Schneider's medium and biofilm quantification

Intracellular *P. acanthamoeba* were isolated from an infected amoeba culture as described above. The suspension was then diluted 1:20 (v/v) in Schneider's medium. 5 mL of the inoculate was transferred to a 25 cm2 culture flasks and incubated at 27°C in a microaerobic bags amended with an ascorbic acid containing paper sachet (CampyGen Compact, Oxoid, Thermo Fisher Scientific Inc., USA) to create a microaerobic environment.

For the quantification of biofilm, we compared cultivation under microaerobic and aerobic incubation conditions. Both were incubated within the microaerobic bags, but only the microaerobic group was supplemented with the ascorbic acid bag to form a microaerobic condition. We used *Pseudomonas aeruginosa* PAO1 as positive control and uninoculated Schneider's medium as a negative control. Incubations including controls were done in biological quadruplicates. The biofilm quantification assay with crystal violet to stain extracellular matrix produced in biofilms was described in Merritt et al., 2005. Washing and staining procedures were performed in 24 well multidish plates. After 16 incubation days, UWC12 containing wells were washed twice with H_2O_{bidest} to remove planktonic bacteria and incubated in 1 mL crystal violet solution (0.1 % w/v in H_2O_{bidest}) for 10 minutes to stain the biofilm. Afterwards we removed the crystal violet solution and washed the well three times with H_2O_{bidest} to remove residual crystal violet. The wells were air-dried and incubated in 1.2 mL of 30 % acetic acid for 15 min to dissolve the stained biofilm. 125 µL of the crystal violet/acetic acid solution from each well were transferred into a clear flat bottom 96 well plate (Greiner Bio-One, Kremsmünster, Austria) in technical duplicates. The optical density (OD) of 600 nm wavelength of each sample was then measured with the Infinite M200 microplate reader (Tecan, Männedorf, Switzerland).

DNA isolation and full genome sequencing

We isolated the genomic DNA of four *Procabacter* strains, namely UWC12, EI5, OEW1, and CRCP. Acanthamoeba cells infected with the respective endosymbiont *Procabacter* sp. were harvested from a flask (Nunc TripleFlask, 500 cm^2) and pelleted by centrifugation (4°C, 6600

rpm, 6 min). The pellet was resuspended in 10 mL Buffer A + EDTA (35 mM Tris-HCL, 250 mM Sucrose, 250 mM EDTA, 25 mM KCl, 10 mM $MgCl_2$, pH 7.5) and homogenized with a dounce tissue grinder (15 times, tight pestle) to release the endosymbiont by host cells breakup. The lysate was centrifuged a second time (4°C, 300g, 2 min) to increase the symbiont yield. The supernatant fraction containing the symbionts was stored at 4°C and the pellet was dounced a second time after resuspension in 10 mL Buffer A +EDTA. Centrifuge (4°C, 300g, 2 min) again to gather the symbiont fraction. All symbiont fractions were pooled, filtered (5 µm ø) and centrifuged once again (4°C, 11000 rpm, 5 min). The pellet was washed once with 5 mL ice-cold Buffer A (35 mM Tris-HCL, 250 mM Sucrose, 25 mM KCl, 10 mM $MgCl_2$, pH 7.5). After repeating the centrifugation and washing step, the pellet was resuspended in 1 mL Buffer A and treated with 10 µl DNAse I (Thermo Scientific) for 1 h on ice to remove extracellular residue DNA. The digestion was stopped by addition of 0.5M EDTA and followed by centrifugation (4°C, 11000 rpm, 5 min) and washing with Buffer A + EDTA. The purified *Procabacter* cells were resuspended in 250 µL TE-Buffer and then mixed with 675 µL DNA extraction buffer (100 mM Tris/HCl, 100 mM EDTA, 100 mM Na-Phosphate, 1.5 M NaCl, 1% CTAB, 200 µg/mL Proteinase K, pH 8). The mixture was incubated at 37°C for 30 min and inverted once every 10 min. Then 75 µL of 20% SDS was added, mixed properly, and incubated at 65°C for 1 h. The tube was inverted every 15 min. Following the incubation, an equal amount of Phenol/Chloroform/Isoamylalcohol (24:1 v/v) was added. After phase separation by centrifugation (room-temperature, 14000 rpm, 5 min), the aqueous phase was recovered. An equal amount of Chloroform/Isoamylalcohol was added to the phase, mixed and then centrifuged (room-temperature, 14000 rpm, 20 min). The aqueous phase was recovered once again and mixed with 1/10 volume of 3M sodium acetate and 1 µL of Glycogen (molecular grade, 20 mg/mL). Nucleic acids were precipitated by adding 0.7 volume of isopropanol and subsequently incubated at room temperature for 1h. The precipitated nucleic acids were pelleted by centrifugation (4°C, 16000 g , 30 min) and washed with ice-cold 70% ethanol once before being spun down again for 5 min at max speed (4°C). The pellet was air-dried for 5 min and resuspended in 30 µL TE-Buffer + RNase (1 µl RNAse by Qiagen 100 mg/mL + 1 mL TE-Buffer). Illumina TruSeq paired-end libraries were prepared and sequenced with the Illumina HiSeq 2000 instrument at the Vienna Biocenter Core Facilities (VBCF) in Vienna. For UWC12, we used additional 454-Sequencing (Library: 8 kb Paired-End, Roche 454 GS Flex) to create

larger contigs and MiSeq sequencing (250 bp Paired-End) for longer reads to reduce contig numbers.

Whole genome assembly and annotation

We applied the tool bbduk (BBMap-35.82 suite, https://sourceforge.net/projects/bbmap) to trim the reads and remove the adapters. The read quality was checked with FastQC (Version 0.11.3; http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). We performed an iterative assembled with SPAdes 3.6.2 (Bankevich et al. 2012) and assessed the assembly quality with QUAST 3.0 (Gurevich et al. 2013). Only contigs larger than 800 bp were included into the downstream analysis. We estimated the completeness and contamination level of all genomes with CheckM 1.0.6 (Parks et al. 2015) and AMPHORA2 (Parks et al. 2015; Wu and Scott 2012). We used ConsPred (V1.24; Weinmaier et al. 2016) for gene prediction and genome annotation. We screened the *Procabacter* genomes for known protein secretion apparatus, mainly Type I to Type VI secretion systems (T1SS - T6SS), with the software MacSyFinder (Abby and Rocha 2017) by using specific hidden Markov model (HMM) protein profiles. At the time of the analysis, no models were available for the newly described T6SS^{IV}; (Böck et al. 2017), flagella, T9SS, Type IV pili, and the Tad pili.

Dataset creation and comparative genomics

For comparative genomics, a dataset encompassing 71 genomes from the class Betaproteobacteria, the *Procabacter* genomes, and outgroup was created (Supp. Table 8). Mainly refseq genomes of the Neisseriales, Burkholderiales, Rhodocyclales, and Nitrosomonadales were chosen. We included Pseudomonas *aeruginosa* PAO1 genome as an outgroup for better rooting of the tree. The software checkM was used to estimate the quality of the genomes and the presence of potential contaminations by assessing the amount of almost universal single-copy genes specific for each phylogenetic group (Parks et al. 2015). Protein sequences were aligned using BLAST v2.6.0 (Altschul et al. 1990) with an expectation value cutoff set to 1 and were subsequently clustered into orthogroups using OrthoFinder v2.0.0 (Altschul et al. 1990; Emms and Kelly 2015). Bidirectional best blast hit average amino acid identities were calculated as described in (Hausmann et al. 2018).

Phylogenomic tree and *Procabacter* spp. ancestral genome reconstruction

A phylogenomic species tree was reconstructed using the concatenated alignment of marker genes as provided in the CheckM workflow (Parks et al. 2015). We used iqtree 1.6.2 using model test to find the best fitting model for the marker gene supermatrix (Minh, Nguyen, and von Haeseler 2013). The species tree was reconstructed using iqtree with the LG substitution matrix, 4 discrete gamma categories (+G4), modeling of invariable sites (+I) model, and the C60 empirical mixture model to account for site specific amino-acid replacement patterns (Quang, Gascuel, and Lartillot 2008). We then used Count (Csurös 2010), a gene tree unaware method, to infer the gains, losses, and duplications within genomes along the species tree. We chose the Wagner Parsimony model (Gene Gain Penalty = 2) to infer gene family gain and loss. The protein family annotation was performed as follows: information extracted from the uniprot database for the refseq genome were applied and the annotation manually curated. For protein families which include only *Procabacter* proteins, we inferred the function by using the online tool Eggnog Mapper (Csurös 2010; Huerta-Cepas et al. 2015, 2016) with the standard settings.

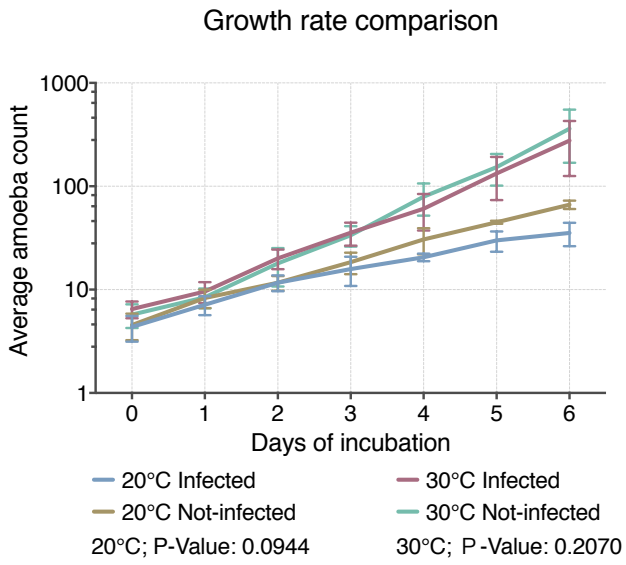
Phylogenetic analysis of cytBD subunits

For phylogenetic placement of cytBD subunits 1 and 2 we aligned *Procabacter* spp. copies to the assigned eggnog orthologous group proteins with mafft v7.312 (Katoh and Standley 2013) using the L-INS-i algorithm and de-replicated the alignment at a cutoff of 95 % amino acid identity with trimAl v1.4.1 (Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009). We used the unaligned proteins of the dereplicated dataset and denovo aligned with mafft as described above and trimmed the alignment with noisy v1.5.12 (Dress et al. 2008). We inferred single gene phylogenies in iqtree with the best fit model and 1000 parametric bootstraps (parameter "-m TEST -bb 1000").

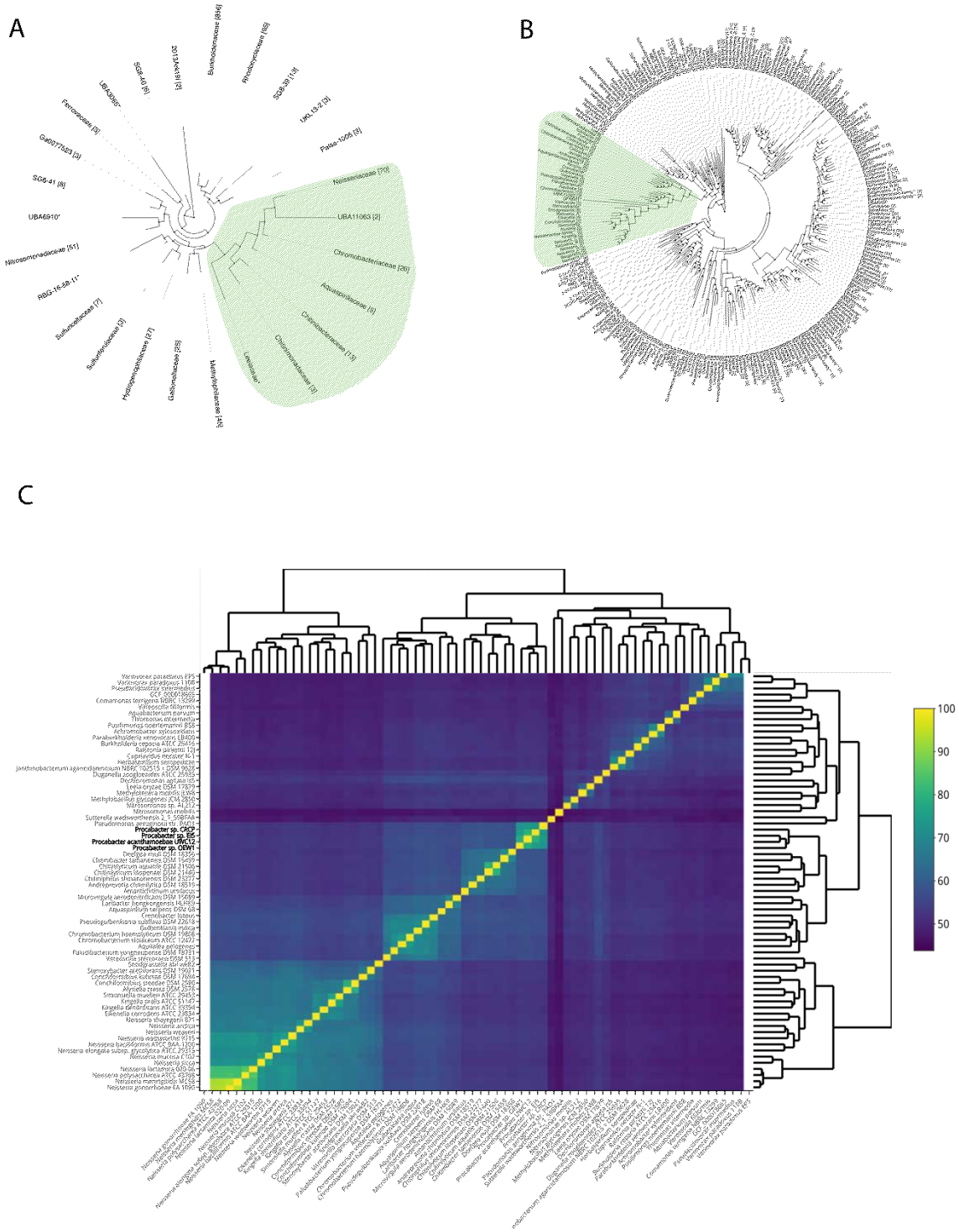
Author Contributions

H.F.T., S.K., and M.H. conceived the study. H.F.T. and S.K. performed the genomic DNA isolation. S.K. performed the infection cycle, biofilm-related experiments, and the genome assembly. H.F.T. and S.K. did the phylogenetic analysis. H.F.T. did the genome annotation and ancestor genome reconstruction. H.F.T. prepared the figures and tables. H.F.T. wrote the manuscript. All authors reviewed and edited the manuscript.

Supplementary Materials

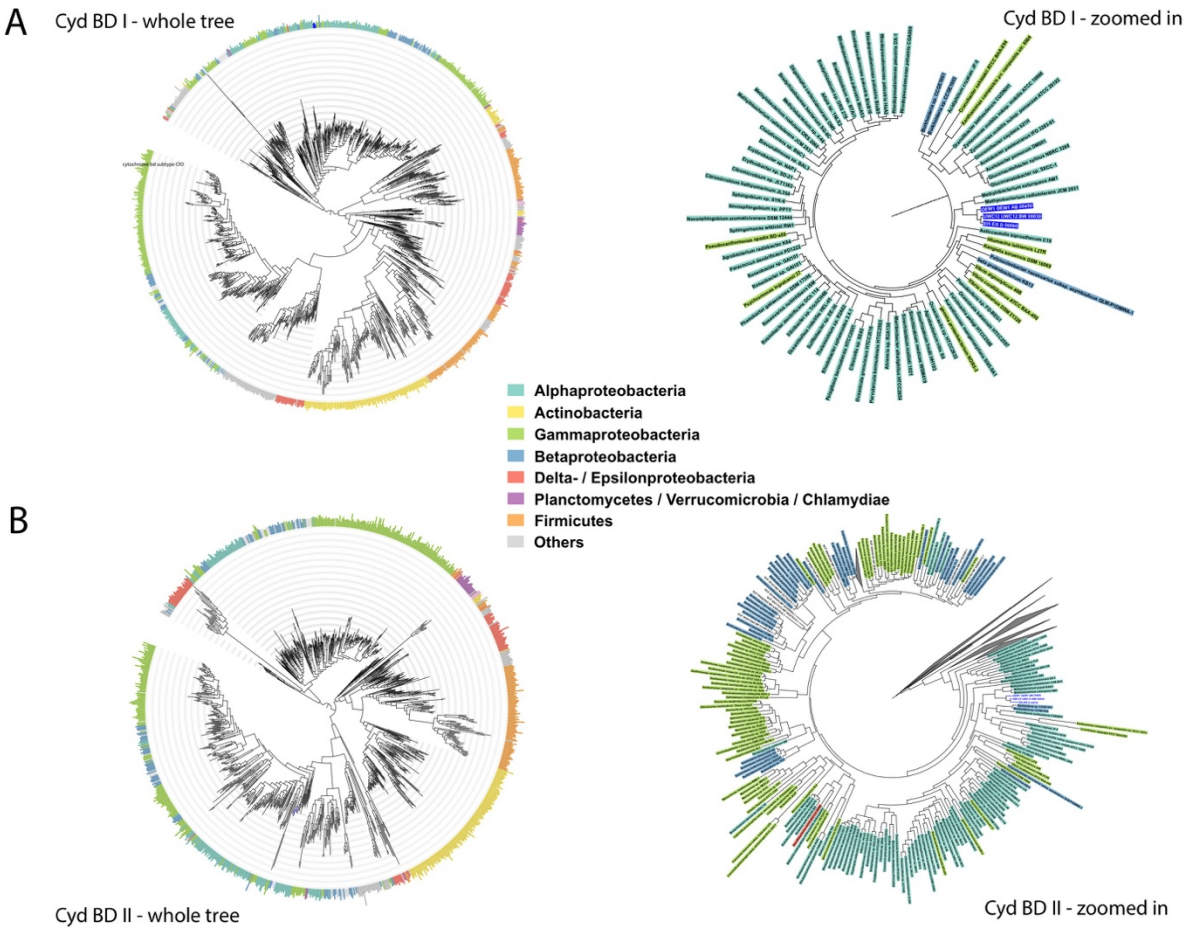


Supp. Figure 1
Growth rate comparison between *Acanthamoeba castellanii* infected with *Procabacter acanthaemobae* UWC12 and non-infected control at 20°C and 30°C. At both temperatures, the infected amoeba cultures grew slower in comparison to the control; inhibition in fitness is more pronounced at 20°C.



Supp. Figure 2

Current phylogeny of the order Burkholderiales, acting as a comparison for the phylogeny analysis based on marker genes performed in this study. (A) Reclassified former Neisseriales group as novel families are highlighted and (B) a detailed genus level view. The proposed family Procabacteriaceae falls at the root of the highlighted branch.



Supp. Figure 3
Phylogeny of *Procabacter's* cydBD I and II subunits with homologs in the same eggnog orthologous group. The colors represent different taxa. Both subunits are placed among homologs of mainly Alphaproteobacteria and Gammaproteobacteria. The data suggests horizontally transferred cytochrome BD, expanding Procabacter's survival in environments of varying oxygen conditions.

References

Abby, S.S. and Rocha, E.P.C. (2017) Identification of Protein Secretion Systems in Bacterial Genomes Using MacSyFinder. In *Methods in Molecular Biology*. pp. 1–21.

Adeolu, M. and Gupta, R.S. (2013) Phylogenomics and molecular signatures for the order Neisseriales: proposal for division of the order Neisseriales into the emended family Neisseriaceae and Chromobacteriaceae fam. nov. *Antonie Van Leeuwenhoek* **104**: 1–24.

Alsam, S., Jeong, S.R., Sissons, J., Dudley, R., Kim, K.S., and Khan, N.A. (2006) Escherichia coli interactions with Acanthamoeba: a symbiosis with environmental and clinical implications. *J Med Microbiol* **55**: 689–694.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., and Stahl, D.A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919–1925.

Arai, H. (2011) Regulation and Function of Versatile Aerobic and Anaerobic Respiratory Metabolism in Pseudomonas aeruginosa. *Front Microbiol* **2**: 103.

Auwerx, J. (1991) The human leukemia cell line, THP-1: a multifacetted model for the study of monocyte-macrophage differentiation. *Experientia* **47**: 22–31.

Avery, S.V., Harwood, J.L., and Lloyd, D. (1995) Quantification and Characterization of Phagocytosis in the Soil Amoeba Acanthamoeba castellanii by Flow Cytometry. *Appl Environ Microbiol* **61**: 1124–1132.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Baughn, A.D. and Malamy, M.H. (2004) The strict anaerobe Bacteroides fragilis grows in and benefits from nanomolar concentrations of oxygen. *Nature* **427**: 441–444.

Benamar, S., Bou Khalil, J.Y., Blanc-Tailleur, C., Bilen, M., Barrassi, L., and La Scola, B. (2017) Developmental Cycle and Genome Analysis of Protochlamydia massiliensis sp. nov. a New Species in the Parachlamydiacae Family. *Frontiers in Cellular and Infection Microbiology* **7.**:1-7.

Böck, D., Medeiros, J.M., Tsao, H.-F., Penz, T., Weiss, G.L., Aistleitner, K., et al. (2017) In situ architecture, function, and evolution of a contractile injection system. *Science* **357**: 713–717.

Borisov, V.B., Gennis, R.B., Hemp, J., and Verkhovsky, M.I. (2011) The cytochrome bd respiratory oxygen reductases. *Biochim Biophys Acta* **1807**: 1398–1413.

Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., et al. (2009) Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. *Proc Natl Acad Sci U S A* **106**: 21848–21853.

Braun, R.D., Lanzen, J.L., Snyder, S.A., and Dewhirst, M.W. (2001) Comparison of tumor and normal tissue oxygen tension measurements using OxyLite or microelectrodes in rodents. *Am J Physiol Heart Circ Physiol* **280**: H2533–44.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Chang, H.-Y., Ahn, Y., Pace, L.A., Lin, M.T., Lin, Y.-H., and Gennis, R.B. (2010) The diheme cytochrome c(4) from Vibrio cholerae is a natural electron donor to the respiratory cbb(3) oxygen reductase. *Biochemistry* **49**: 7494–7503.

Collingro, A., Tischler, P., Weinmaier, T., Penz, T., Heinz, E., Brunham, R.C., et al. (2011) Unity in variety--the pan-genome of the Chlamydiae. *Mol Biol Evol* **28**: 3253–3270.

Csurös, M. (2010) Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* **26**: 1910–1912.

Daims, H., Brühl, A., Amann, R., Schleifer, K.H., and Wagner, M. (1999) The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst Appl Microbiol* **22**: 434–444.

Daims, H., Stoecker K., Wagner M. (2004) Fluorescence in situ hybridization for the detection of prokaryotes. *Molecular Microbial Ecology* 208-228

Dale, C. and Maudlin, I. (1999) Sodalis gen. nov. and Sodalis glossinidius sp. nov., a microaerophilic secondary endosymbiont of the tsetse fly Glossina morsitans morsitans. *Int J Syst Bacteriol* **49**: 267–275.

Dale, C. and Moran, N.A. (2006) Molecular interactions between bacterial symbionts and their hosts. *Cell* **126**: 453–465.

Dalsing, B.L., Truchon, A.N., Gonzalez-Orta, E.T., Milling, A.S., and Allen, C. (2015) Ralstonia solanacearum uses inorganic nitrogen metabolism for virulence, ATP production, and detoxification in the oxygen-limited host xylem environment. *MBio* **6**: e02471.

Dean, P. (2011) Functional domains and motifs of bacterial type III effector proteins and their roles in infection. *FEMS Microbiol Rev* **35**: 1100–1125.

Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L., and Héchard, Y. (2015) Shedding light on microbial dark matter: a TM6 bacterium as natural endosymbiont of a free-living amoeba. *Environ Microbiol Rep* **7**: 970–978.

Díez-Vives, C., Moitinho-Silva, L., Nielsen, S., Reynolds, D., and Thomas, T. (2017) Expression of eukaryotic-like protein in the microbiome of sponges. *Mol Ecol* **26**: 1432–1451.

Domman, D., Collingro, A., Lagkouvardos, I., Gehre, L., Weinmaier, T., Rattei, T., et al. (2014) Massive expansion of Ubiquitination-related gene families within the Chlamydiae. *Mol Biol Evol* **31**: 2890–2904.

Dress, A.W.M., Flamm, C., Fritzsch, G., Grünewald, S., Kruspe, M., Prohaska, S.J., and Stadler, P.F. (2008) Noisy: identification of problematic columns in multiple sequence alignments. *Algorithms Mol Biol* **3**: 7.

Eichinger, V., Nussbaumer, T., Platzer, A., Jehl, M.-A., Arnold, R., and Rattei, T. (2016) EffectiveDB--updates and novel features for a better annotation of bacterial secreted proteins and Type III, IV, VI secretion systems. *Nucleic Acids Res* **44**: D669–74.

Elwell, C., Mirrashidi, K., and Engel, J. (2016) Chlamydia cell biology and pathogenesis. *Nature Reviews Microbiology* **14**: 385–400.

Emms, D.M. and Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**: 157.

Gil, R., Silva, F.J., Zientz, E., Delmotte, F., Gonzalez-Candelas, F., Latorre, A., et al. (2003) The genome sequence of Blochmannia floridanus: Comparative analysis of reduced genomes. *Proceedings of the National Academy of Sciences* **100**: 9388–9393.

Giuffrè, A., Borisov, V.B., Arese, M., Sarti, P., and Forte, E. (2014) Cytochrome bd oxidase and bacterial tolerance to oxidative and nitrosative stress. *Biochim Biophys Acta* **1837**: 1178–1187.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013) QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**: 1072–1075.

Hahn, M.W. and Höfle, M.G. (2001) Grazing of protozoa and its effect on populations of aquatic bacteria. *FEMS Microbiol Ecol* **35**: 113–121.

Hausmann, B., Pelikan, C., Herbold, C.W., Köstlbacher, S., Albertsen, M., Eichorst, S.A., et al. (2018) Peatland Acidobacteria with a dissimilatory sulfur metabolism. *ISME J* **12**: 1729–1742.

Heinz, E., Kolarov, I., Kästner, C., Toenshoff, E.R., Wagner, M., and Horn, M. (2007) An Acanthamoeba sp. containing two phylogenetically different bacterial endosymbionts. *Environ Microbiol* **9**: 1604–1609.

Henning, K., Lothar, Z., Baerbel, H., Helmut, H., and Rolf, M. (2007) Hartmanella vermiformis (Hartmannellidae) harboured a hidden chlamydia-like endosymbiont. *Endocytobiosis Cell Res*.

Hirai, T., Osamura, T., Ishii, M., and Arai, H. (2016) Expression of multiple cbb3cytochromecoxidase isoforms by combinations of multiple isosubunits inPseudomonas aeruginosa. *Proceedings of the National Academy of Sciences* **113**: 12815–12819.

Holden, M.T.G., Titball, R.W., Peacock, S.J., Cerdeño-Tárraga, A.M., Atkins, T., Crossman, L.C., et al. (2004) Genomic plasticity of the causative agent of melioidosis, Burkholderia pseudomallei. *Proc Natl Acad Sci U S A* **101**: 14240–14245.

Horn, M. (2008) Chlamydiae as symbionts in eukaryotes. *Annu Rev Microbiol* **62**: 113–131.

Horn, M., Fritsche, T.R., Linner, T., Gautom, R.K., Harzenetter, M.D., and Wagner, M. (2002) Obligate bacterial endosymbionts of Acanthamoeba spp. related to the beta-Proteobacteria: proposal of "Candidatus Procabacter acanthamoebae" gen. nov., sp. nov. *Int J Syst Evol Microbiol* **52**: 599–605.

Huerta-Cepas, J., Forslund, K., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2016) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution* **8**: 2115-2118

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2015) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44**: D286–D293.

Huws, S.A., McBain, A.J., and Gilbert, P. (2005) Protozoan grazing and its impact upon population dynamics in biofilm communities. *J Appl Microbiol* **98**: 238–244.

Hu, Y., Huang, H., Cheng, X., Shu, X., White, A.P., Stavrinides, J., et al. (2017) A global survey of bacterial type III secretion systems and their effectors. *Environ Microbiol* **19**: 3879–3895.

Ishida, K., Sekizuka, T., Hayashida, K., Matsuo, J., Takeuchi, F., Kuroda, M., et al. (2014) Amoebal endosymbiont Neochlamydia genome sequence illuminates the bacterial role in the defense of the host amoebae against Legionella pneumophila. *PLoS One* **9**: e95166.

Jo, J., Price-Whelan, A., and Dietrich, L.E.P. (2014) An aerobic exercise: defining the roles of Pseudomonas aeruginosa terminal oxidases. *J Bacteriol* **196**: 4203–4205.

Kaminski, P.A., Kitts, C.L., Zimmerman, Z., and Ludwig, R.A. (1996) Azorhizobium caulinodans uses both cytochrome bd (quinol) and cytochrome cbb3 (cytochrome c) terminal oxidases for symbiotic N2 fixation. *J Bacteriol* **178**: 5989–5994.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Kida, Y., Higashimoto, Y., Inoue, H., Shimizu, T., and Kuwano, K. (2008) A novel secreted protease from Pseudomonas aeruginosa activates NF-kappaB through protease-activated receptors. *Cell Microbiol* **10**: 1491–1504.

König, L., Siegl, A., Penz, T., Haider, S., Wentrup, C., Polzin, J., et al. (2017) Biphasic Metabolism and Host Interaction of a Chlamydial Symbiont. *mSystems* **3.**:

König, L., Wentrup, C., Schulz, F., Wascher, F., Escola, S., Swanson, M.S., et al. (2019) Symbiont-Mediated Defense against Legionella pneumophila in Amoebae. *MBio* **10.**:

Leo, J.C., Grin, I., and Linke, D. (2012) Type V secretion: mechanism(s) of autotransport through the bacterial outer membrane. *Philos Trans R Soc Lond B Biol Sci* **367**: 1088–1101.

Maezawa, K., Shigenobu, S., Taniguchi, H., Kubo, T., Aizawa, S.-I., and Morioka, M. (2006) Hundreds of flagellar basal bodies cover the cell surface of the endosymbiotic bacterium Buchnera aphidicola sp. strain APS. *J Bacteriol* **188**: 6539–6543.

McCutcheon, J.P. (2010) The bacterial essence of tiny symbiont genomes. *Curr Opin Microbiol* **13**: 73–78.

McCutcheon, J.P. and Moran, N.A. (2011) Extreme genome reduction in symbiotic bacteria. *Nat Rev Microbiol* **10**: 13–26.

Merritt JH., Kadouri DE., O'Toole GA. (2005) Growing and analyzing static biofilms. *Curr Protoc Microbiol* Chapter 1: Unit 1B.1

Miki, T., Iguchi, M., Akiba, K., Hosono, M., Sobue, T., Danbara, H., and Okada, N. (2010) Chromobacterium pathogenicity island 1 type III secretion system is a major virulence determinant for Chromobacterium violaceum-induced cell death in hepatocytes. *Mol Microbiol* **77**: 855–872.

Minh, B.Q., Nguyen, M.A.T., and von Haeseler, A. (2013) Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol* **30**: 1188–1195.

Miura, H., Mogi, T., Ano, Y., Migita, C.T., Matsutani, M., Yakushi, T., et al. (2013) Cyanide-insensitive quinol oxidase (CIO) from Gluconobacter oxydans is a unique terminal oxidase subfamily of cytochrome bd. *J Biochem* **153**: 535–545.

Moran, N.A. (2002) Microbial minimalism: genome reduction in bacterial pathogens. *Cell* **108**: 583–586.

Moran, N.A. (2007) Symbiosis as an adaptive process and source of phenotypic complexity. *Proc Natl Acad Sci U S A* **104 Suppl 1**: 8627–8633.

Nilsson, A.I., Koskiniemi, S., Eriksson, S., Kugelberg, E., Hinton, J.C.D., and Andersson, D.I. (2005) Bacterial genome size reduction by experimental evolution. *Proc Natl Acad Sci U S A* **102**: 12112–12116.

Nuccio, S.-P. and Bäumler, A.J. (2007) Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek. *Microbiol Mol Biol Rev* **71**: 551–575.

Ogata, H., La Scola, B., Audic, S., Renesto, P., Blanc, G., Robert, C., et al. (2006) Genome sequence of Rickettsia bellii illuminates the role of amoebae in gene exchanges between intracellular pathogens. *PLoS Genet* **2**: e76.

Omsland, A., Hackstadt, T., and Heinzen, R.A. (2013) Bringing culture to the uncultured: Coxiella burnetii and lessons for obligate intracellular bacterial pathogens. *PLoS Pathog* **9**: e1003540.

Pagnier, I., Yutin, N., Croce, O., Makarova, K.S., Wolf, Y.I., Benamar, S., et al. (2015) Babela massiliensis, a representative of a widespread bacterial phylum with unusual adaptations to parasitism in amoebae. *Biol Direct* **10**: 13.

Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**: 996–1004.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Parry, J.D. (2004) Protozoan Grazing of Freshwater Biofilms. In *Advances in Applied Microbiology* **54**: 167–196.

Pizarro-Cerdá, J. and Cossart, P. (2006) Bacterial adhesion and entry into host cells. *Cell* **124**: 715–727.

Poole, R.K. and Cook, G.M. (2000) Redundancy of aerobic respiratory chains in bacteria? Routes, reasons and regulation. In *Advances in Microbial Physiology*. pp. 165–224.

Quang, L.S., Gascuel, O., and Lartillot, N. (2008) Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24**: 2317–2323.

Quillin, S.J. and Seifert, H.S. (2018) Neisseria gonorrhoeae host adaptation and pathogenesis. *Nat Rev Microbiol* **16**: 226–240.

Rodríguez-Zaragoza, S. (1994) Ecology of Free-Living Amoebae. *Crit Rev Microbiol* **20**: 225–241.

Rojas, C.M., Ham, J.H., Deng, W.-L., Doyle, J.J., and Collmer, A. (2002) HecA, a member of a class of adhesins produced by diverse pathogenic bacteria, contributes to the attachment, aggregation, epidermal cell killing, and virulence phenotypes of Erwinia chrysanthemi EC16 on Nicotiana clevelandii seedlings. *Proc Natl Acad Sci U S A* **99**: 13142–13147.

Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., and Horn, M. (2010) The genome of the amoeba symbiont "Candidatus Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J Bacteriol* **192**: 1045–1057.

Schmitz-Esser, S., Toenshoff, E.R., Haider, S., Heinz, E., Hoenninger, V.M., Wagner, M., and Horn, M. (2008) Diversity of bacterial endosymbionts of environmental acanthamoeba isolates. *Appl Environ Microbiol* **74**: 5822–5831.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**: 671–675.

Schneider, I. (1972) Cell lines derived from late embryonic stages of Drosophila melanogaster. *J Embryol Exp Morphol* **27**: 353–365.

Schulz, F., Lagkouvardos, I., Wascher, F., Aistleitner, K., Kostanjšek, R., and Horn, M. (2014) Life in an unusual intracellular niche: a bacterial symbiont infecting the nucleus of amoebae. *ISME J* **8**: 1634–1644.

Schulz, F., Martijn, J., Wascher, F., Lagkouvardos, I., Kostanjšek, R., Ettema, T.J.G., and Horn, M. (2015) A Rickettsiales symbiont of amoebae with ancient features. *Environ Microbiol* **18**: 2326–2342.

Sherr, E.B. and Sherr, B.F. (2002) Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek* **81**: 293–308.

Sixt, B.S., Hiess, B., König, L., and Horn, M. (2012) Lack of effective anti-apoptotic activities restricts growth of Parachlamydiaceae in insect cells. *PLoS One* **7**: e29565.

Thomas, V. and Greub, G. (2010) Amoeba/amoebal symbiont genetic transfers: lessons from giant virus neighbours. *Intervirology* **53**: 254–267.

Toft, C. and Fares, M.A. (2008) The evolution of the flagellar assembly pathway in endosymbiotic bacterial genomes. *Mol Biol Evol* **25**: 2069–2076.

Tomaras, A.P., Dorsey, C.W., Edelmann, R.E., and Actis, L.A. (2003) Attachment to and biofilm formation on abiotic surfaces by Acinetobacter baumannii: involvement of a novel chaperone-usher pili assembly system. *Microbiology* **149**: 3473–3484.

Tsao, H.-F., Scheikl, U., Volland, J.-M., Köhsler, M., Bright, M., Walochnik, J., and Horn, M. (2017) "Candidatus Cochliophilus cryoturris" (Coxiellaceae), a symbiont of the testate amoeba Cochliopodium minus. *Sci Rep* **7**: 3394.

Wagner, M., Nielsen, P.H., Loy, A., Nielsen, J.L., and Daims, H. (2006) Linking microbial community structure with function: fluorescence in situ hybridization-microautoradiography and isotope arrays. *Curr Opin Biotechnol* **17**: 83–91.

Wang, X. and Wood, T.K. (2011) Toxin-antitoxin systems influence biofilm and persister cell formation and the general stress response. *Appl Environ Microbiol* **77**: 5577–5583.

Weinmaier, T., Platzer, A., Frank, J., Hellinger, H.-J., Tischler, P., and Rattei, T. (2016) ConsPred: a rule-based (re-)annotation framework for prokaryotic genomes. *Bioinformatics* **32**: 3327–3329.

Wells, T.J., Tree, J.J., Ulett, G.C., and Schembri, M.A. (2007) Autotransporter proteins: novel targets at the bacterial cell surface. *FEMS Microbiol Lett* **274**: 163–172.

Wernegreen, J.J. (2002) Genome evolution in bacterial endosymbionts of insects. *Nat Rev Genet* **3**: 850–861.

Wu, M. and Scott, A.J. (2012) Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**: 1033–1034.

Chapter VI

In situ architecture, function, and evolution of a novel contractile injection system

TYPE 6 SECRETION

In situ architecture, function, and evolution of a contractile injection system

Désirée Böck,[1]* João M. Medeiros,[1]* Han-Fei Tsao,[2] Thomas Penz,[2]† Gregor L. Weiss,[1] Karin Aistleitner,[2]‡ Matthias Horn,[2]§ Martin Pilhofer[1]§

Contractile injection systems mediate bacterial cell-cell interactions by a bacteriophage tail–like structure. In contrast to extracellular systems, the type 6 secretion system (T6SS) is defined by intracellular localization and attachment to the cytoplasmic membrane. Here we used cryo-focused ion beam milling, electron cryotomography, and functional assays to study a T6SS in *Amoebophilus asiaticus*. The in situ architecture revealed three modules, including a contractile sheath-tube, a baseplate, and an anchor. All modules showed conformational changes upon firing. Lateral baseplate interactions coordinated T6SSs in hexagonal arrays. The system mediated interactions with host membranes and may participate in phagosome escape. Evolutionary sequence analyses predicted that T6SSs are more widespread than previously thought. Our insights form the basis for understanding T6SS key concepts and exploring T6SS diversity.

Contractile injection systems (CISs) deliver effectors to mediate bacterial cell-cell interactions. Their structural components are homologous to the contractile tails of phages (*1*). CISs consist of an inner tube surrounded by a contractile sheath, a spike capping the inner tube, and a baseplate complex at the base of the sheath. Sheath contraction propels the inner tube into the target. Two modes of action divide CISs into "extracellular CISs" (eCISs) and "type 6 secretion" (T6S) systems (T6SSs). eCISs resemble headless phages; they are released into the medium and bind to the target cell surface. Examples are antibacterial R-type bacteriocins (*2*), insecticidal antifeeding prophages (Afps) (*3*), and metamorphosis-inducing structures (MACs) (*4*). By contrast, the T6SS is defined by its cytoplasmic localization and anchoring to the inner membrane (*5*–*9*).

Amoebophilus asiaticus (hereafter referred to as *Amoebophilus* or amoebophili) is an obligate intracellular bacterial symbiont of amoebae (*10*). The *Amoebophilus* genome does not encode known secretion systems (*11*); however, it contains a gene cluster with similarities to that of Afps (*12*). We reasoned that the Afp-like gene cluster might encode a system that would give insight into T6SS structure, function, and evolution.

To investigate whether *Amoebophilus* produced any CISs, we imaged bacterial cells that were purified from infected amoeba cultures by electron cryotomography (ECT). Fifty percent of the imaged cells (n = 92) contained phage tail–like assemblies. Like T6SSs and unlike eCISs, the structures were always located in the *Amoebophilus* cytoplasm and attached to the cytoplasmic membrane. The structures were always found in bundles of 2 to 34 parallel individual systems (8 on average; Fig. 1, A to C; fig. S1, A to D; and movie S1). Cells contained either one or two (85%) bundles (Fig. 1D and fig. S1E). Inside a bundle, the structures were arranged in ordered hexagonal arrays (Fig. 1, E and F). Extended and contracted conformational states could be distinguished by differences in length (242 ± 7 nm, n = 254, and 122 ± 6 nm, n = 153, respectively), diameter (14 ± 2 nm and 19 ± 2 nm, respectively), and surface properties of the sheath (helical ridges on the contracted structures) (Fig. 1, B and C, and fig. S1, F to I). The narrow distribution of sheath lengths indicates a mechanism for length control. In addition, not all structures inside an array appeared to fire together (Fig. 1, B to D and F, and fig. S1E).

The arrays of contractile structures were encoded by the *Amoebophilus* Afp-like gene cluster. Twelve of its components were detected in a sheath preparation (table S1). Sheath structures were labeled by specific antibodies (fig. S1, J and K). Furthermore, tomograms of purified sheath revealed contracted sheaths whose structure (fig. S1, L and M) and dimensions (length 115 ± 7 nm, diameter 19 ± 2 nm, n = 51) were consistent with the structures observed in situ (fig. S1, F to I).

To observe macromolecular details, we averaged subvolumes of extended T6S-like machines (Fig. 2, A to K, and movies S2 and S3). The structure revealed three major modules: a sheath-tube assembly, a baseplate, and an anchoring complex (Fig. 2, A to I). This segmentation was supported by the comparison of the average with the structure of the minimal composition of a contractile injection system derived from the T4 phage tail (*13*) (Fig. 2B). All three modules showed sixfold rotationally symmetric features (fig. S2). Densities for the inner tube (7 nm diameter) and sheath (14 nm diameter) could be clearly discerned (Fig. 2, A and I). The baseplate was overall hexagonal (Fig. 2, G and H) and established connections with baseplates of neighboring structures, thereby coordinating multiple systems in hexagonal arrays (Fig. 2, J and K). The central baseplate region featured additional densities that reinforced in a threefold symmetrized average (fig. S2, I to M). The anchoring complex consisted of a six-footed platform that attached the baseplate to the inner membrane (Fig. 2, A and D to F). Densities whose dimensions were consistent with a spike complex (*14*) were seen capping the inner tube and protruded through the centers of baseplate and anchor (Fig. 2, A and B and E to G). The averages lacked densities that would indicate the presence of an elaborate trans-envelope complex (Fig. 2A and fig. S2, A and E to G) such as TssJLM (*9*).

The T4 phage baseplate triggers sheath contraction by a large-scale conformational change (*13*). We therefore calculated an average of contracted structures (Fig. 2, L to R, and movies S4 and S5). Again, sheath, baseplate, and anchor modules were identified (Fig. 2, L to R). All three modules revealed pronounced conformational changes as compared to the extended state (movie S6). Similar to the *Vibrio cholerae* T6SS (*5*), the sheath diameter increased upon contraction, along with the appearance of helical surface ridges and the disappearance of the inner tube (Fig. 2, I and R). The baseplate showed a widening and a loss of densities in the center (Fig. 2, G and H and P and Q). Likewise, the anchoring platform showed lateral expansion (distance between opposing feet at the inner membrane increased from 16 to 19 nm) and a loss of the spike density in the center (Fig. 2, E and F and N and O). On a larger scale, the spacing between contractile structures increased from 22 nm between extended structures to 30 nm between contracted structures.

We then tested whether the *Amoebophilus* system secreted tube protein into extracellular space. Immunoblotting detected tube protein (Hcp), but no sheath, in the supernatant of a culture (fig. S3), indicating Hcp translocation. The system thus fulfills the functional hallmark of canonical T6SSs. Together with the structural data, this suggests that the *Amoebophilus* Afp-like gene cluster encodes a T6SS rather than an eCIS.

Next, we investigated the function of T6SS arrays. Amoebophili were internalized in the first 2 hours postinfection (hpi) and exited the amoebae ~144 hpi (fig. S4). To observe intracellular amoebophili by ECT, we used cryo-focused ion beam milling to generate lamellae that were suitable for ECT (fig. S5) (*15*, *16*). At 0.25 hpi, coccoid amoebophili were inside phagosomes (80%, n = 20; Fig. 3A and movie S7). At later stages, most amoebophili had escaped into the cytosol (94%, n = 94), differentiated into rods, and replicated (Fig. 3, B to D; fig. S6A; and movies S8 to S10). The sheath mRNA level was 230-fold higher in extracellular amoebophili compared to the replicative stage (table S2). Likewise, cryotomograms of amoebophili from different infection stages showed that T6SSs were most abundant in extracellular amoebophili (58%, n = 19) and at early infection stages (1 hpi, 96%, n = 25; 2 hpi, 69%, n = 13), whereas replicative amoebophili

[1]Institute of Molecular Biology and Biophysics, Eidgenössische Technische Hochschule Zürich, CH-8093 Zürich, Switzerland. [2]Division of Microbial Ecology, University of Vienna, A-1090 Vienna, Austria.
*These authors contributed equally to this work. †Present address: CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria. ‡Present address: National Institute of Allergy and Infectious Diseases, Hamilton, MT 59840, USA. §Corresponding author. Email: horn@microbial-ecology.net (M.H.); pilhofer@biol.ethz.ch (M.P.)

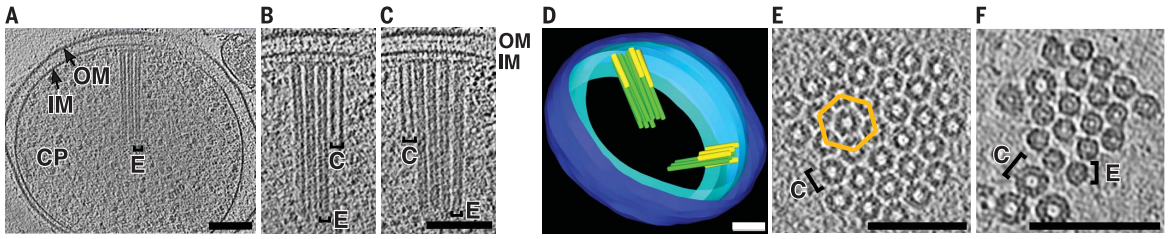


Fig. 1. The *Amoebophilus* Afp-like gene cluster encodes ordered arrays of T6S-like structures. (**A** to **C**) Cryotomograms of purified *Amoebophilus* cells revealed bundles of cytoplasmic, membrane-bound, contractile structures in extended ("E") and contracted ("C") conformations. CP, cytoplasm; IM, inner membrane; OM, outer membrane. Shown are three different 14-nm slices through the same tomogram. (**D**) Bundles comprised 2 to 34 individual machines (green, extended; yellow, contracted) and were organized in one or two bundles per cell. Shown is a model of the tomogram shown in (A) to (C). Blue, outer membrane; cyan, inner membrane. (**E** and **F**) Structures were arranged in hexagonal arrays (lattice indicated in orange). Shown are 15-nm (E) and 8-nm (F) cross-sectional slices. Bars: 100 nm.

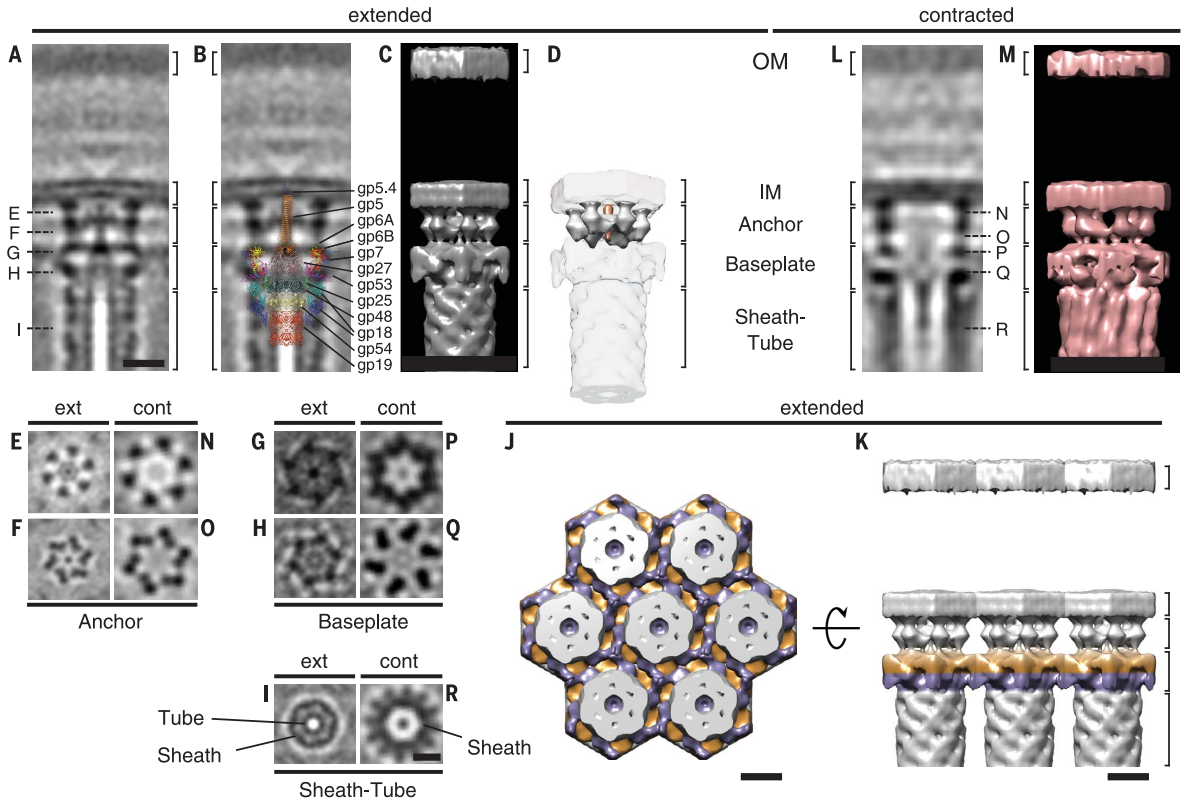


Fig. 2. The T6SS in situ architecture reveals three major modules, conformational changes upon firing, and the structural basis of array formation. Subtomogram averages of extended (**A** to **K**) and contracted (**L** to **R**) T6SSs revealed three major cytoplasmic modules: Sheath-Tube, Baseplate, and Anchor (indicated by brackets). Shown are 0.81-nm [(A) and (B), (E) to (I)] and 1.38-nm [(L), (N) to (R)] longitudinal [(A), (B), and (L)] and perpendicular [(E) to (I), (N) to (R)] slices through sixfold rotationally symmetrized averages, and their three-dimensional models [(C), (D), (J), (K), and (M)]. The positions of perpendicular sections are indicated in (A) and (L). OM, outer membrane; IM, inner membrane. Bars: 10 nm [(A) to (D), (L), and (M) to scale; (E) to (I), (N) to (R) to scale]. The segmentation in three modules was supported by the overlay (B) with the structure of the minimal composition of a contractile injection system derived from the T4 phage tail [from (*13*)]. It allowed the putative assignment of densities corresponding to tube (gp19/gp48/gp54), sheath (gp18), spike complex (gp5/gp5.4/gp27), and baseplate wedge components (gp6/gp7/gp25/gp53). Densities that were not accounted for were assigned to the anchor module [segmented in (D); white, anchor; orange, spike]. Upon firing, pronounced conformational changes were detected in all modules [movie S6 shows a morph between models shown in (C) and (M)]. The sheath increased in diameter and formed surface ridges [(A), (C), (I), (L), (M), and (R)]. Baseplate and anchor showed widening and loss of densities in the center [(A), (C), (E) to (H), (L), (M), (N) to (Q)]. Ordered arrays were established by lateral interactions of the hexagonal baseplates [(J) and (K)]. Shown are top (J) and side (K) views of a model that was assembled from masked averages. For viewing purposes, two different baseplate levels are colored in purple and orange.

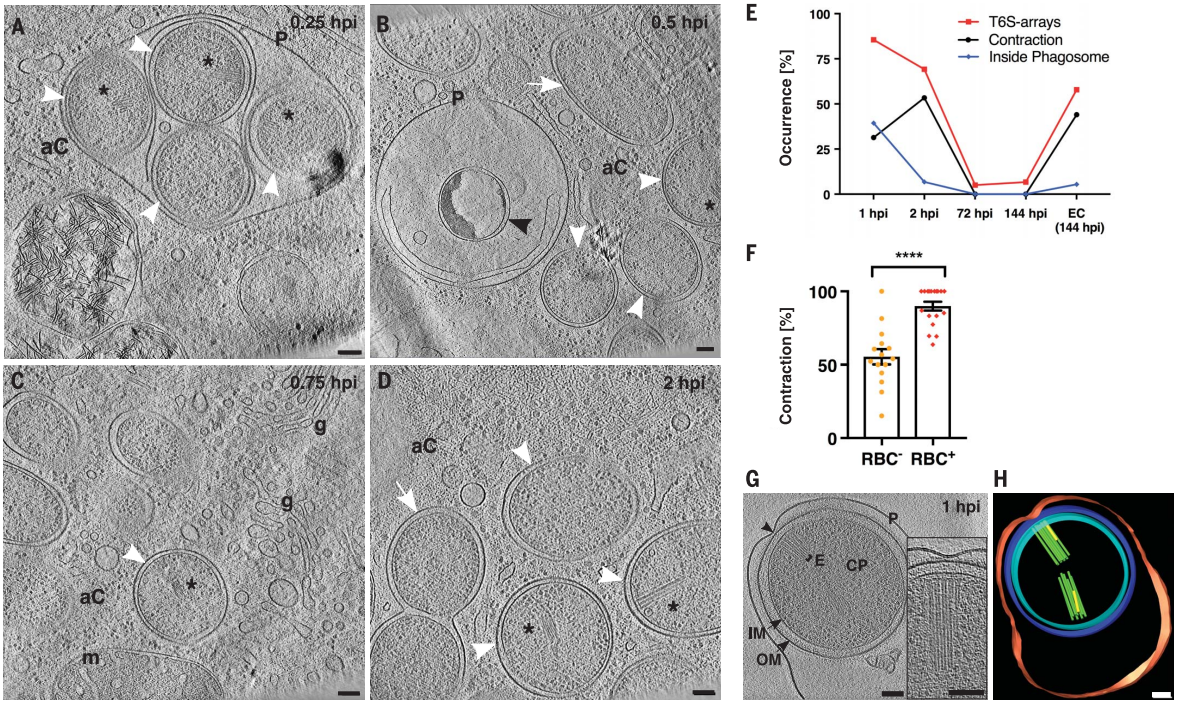


Fig. 3. T6S arrays are required during early infection stages and mediate interactions with host membranes. (**A** to **D**) Bacteria inside their host were imaged by cryo-focused ion beam milling and ECT (fig. S5). At 0.25 hpi, most coccoid amoebophili (white arrowheads) were found inside phagosomes (P). At later time points (0.5 to 2 hpi), amoebophili had escaped into the amoeba cytosol (aC). Amoebophili differentiated into rods (white arrows) and divided. A small fraction did not escape, showing signs of degradation (black arrowhead). Shown are 15-nm slices through cryotomograms. Asterisk, T6S array; g, golgi apparatus; m, mitochondrion. Bars: 100 nm. (**E**) Abundance of T6S arrays was determined by ECT of amoebophili purified from synchronized cultures, and found to be highest in extracellular "EC (144 hpi)" and early intracellular infection stages (1 hpi, 2 hpi). The increase in the contraction rate between 1 and 2 hpi correlated with the escape from the phagosome. Shown are the percentages of cells with T6S arrays (red), percentages of T6S structures that were contracted (black), and percentages of cells found inside phagosomes (blue). T6S arrays, number of quantified amoebophili: n^{1hpi} = 25, n^{2hpi} = 13, n^{72hpi} = 20, n^{144hpi} = 15, $n^{EC(144\ hpi)}$ = 19; Contraction, number of quantified T6S structures: n^{1hpi} = 168, n^{2hpi} = 88, n^{72hpi} = 4, n^{144hpi} = 4, $n^{EC(144\ hpi)}$ = 59; Inside Phagosome, number of quantified amoebophili: n^{1hpi} = 121, n^{2hpi} = 118, n^{72hpi} = 218, n^{144hpi} = 337, $n^{EC(144\ hpi)}$ = 55). (**F**) Amoebophili showed hemolytic activity (fig. S7). Extracellular amoebophili that interacted with RBCs showed an increased T6S contraction rate (****$P < 0.0001$; n^{RBC^-} = 506; n^{RBC^+} = 480; mean ± SEM). (**G** and **H**) Amoebophili residing in phagosomes revealed contact sites (black arrowhead) between the *Amoebophilus* outer membrane and the phagosome. Any such contact site correlated with a T6S array (n = 14). Shown are a 15-nm tomographic slice (G) and the corresponding model (H). P/red, phagosome; OM/blue, outer membrane; IM/cyan, inner membrane; CP, cytoplasm; E/green, extended T6SS; yellow, contracted T6SS; Bars: 100 nm.

did usually not harbor structures (5%, n = 20) (Fig. 3E and fig. S6, B to G). The process of exiting the phagosome during early infection (up to 2 hpi) correlated with increased fractions of contracted structures (Fig. 3E). Experiments comparing the potential of amoebophili from different infection stages to establish new infections showed that host infection rates positively correlated with T6SS expression (fig. S6H). Likewise, we tested amoebophili from different infection stages for hemolytic activity and found that red blood cell (RBC) lysis also positively correlated with T6SS expression (fig. S7). ECT imaging of amoebophili that had the possibility to interact with RBCs showed a 30% increased fraction of contracted structures, compared to a control sample (Fig. 3F; $P < 0.0001$). The analysis of tomograms of purified amoebophili that were found inside phagosomes (39% at 1 hpi) revealed that any contact site between the phagosome membrane and the outer membrane of the bacterium correlated with the presence of T6SSs (with at least one contracted structure) (n = 14; Fig. 3, G and H; fig. S8; and movie S11). Together, our data suggest that T6S arrays mediate interactions with host membranes and may participate in phagosome escape. It remains open whether phagosome rupturing is mediated by mechanical forces or membrane-targeting effectors.

Next, we sought to understand the evolutionary history of the *Amoebophilus* Afp-like gene cluster. We compared three key components (sheath, tail tube, baseplate component gp25) to other CISs (table S3). Similarities were highest with a gene cluster of unknown function in the related symbiont *Cardinium hertigii* (*17*). Moderate similarities were found for Afps and MACs, both mediating interactions with animal larvae (*3*, *4*). Lowest (or no similarities at all) were detected for T6SSs [subtypes *i*, *ii*, *iii* (*18*)] and contractile phages. Likewise, phylogenetic analyses revealed that *Amoebophilus* (and *Cardinium*) sequences stably clustered in a monophyletic group with Afps and MACs, rather than with T6SS$^{i\text{-}iii}$ (Fig. 4A and fig. S9). With the exception of gp7, the analysis of gene content detected *Amoebophilus* homologs of all components that are conserved across CISs and phages (*13*). These include putative sheath (gp18/TssBC), inner tube (gp19/Hcp, gp48, gp54), spike (gp5/VgrG, gp5.4/PAAR, gp27), and three baseplate wedge components (gp6/TssF, gp25/TssE, gp53) (Fig. 4B and table S4). The lack of gp7 might be explained by the presence of two gp6 homologs, and the suggestion that gp6 and gp7 diverged from a common ancestor (*13*, *19*).

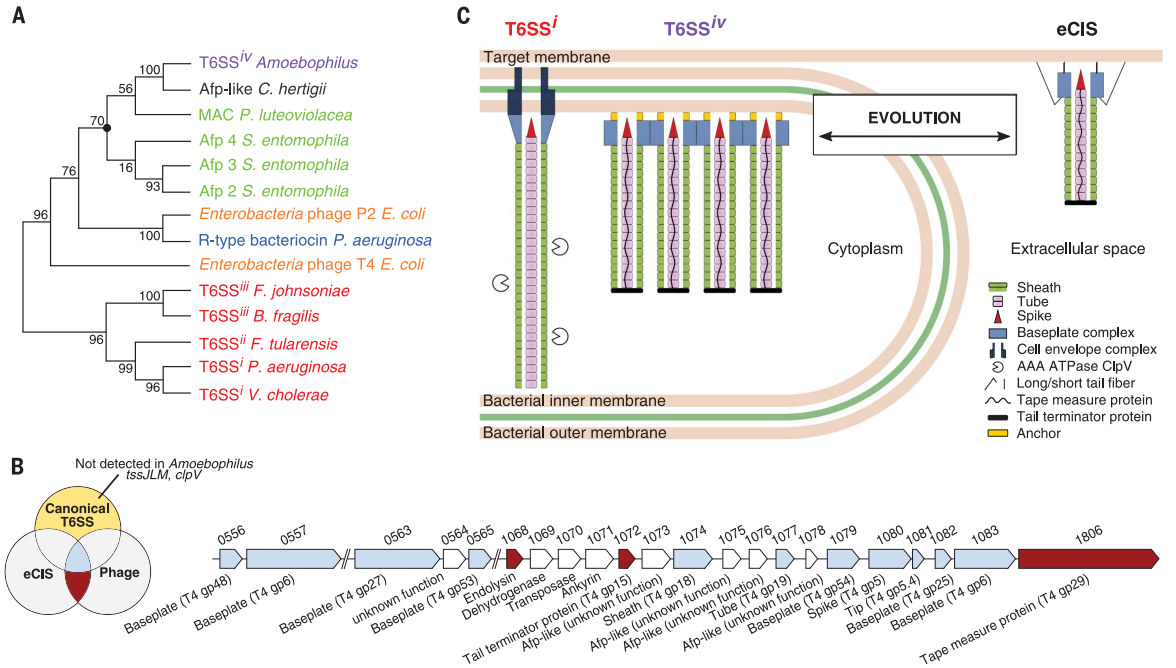


Fig. 4. *Amoebophilus* T6SS^{iv} is closely related to eCIS. (**A**) Phylogenetic analyses of sheath protein sequences showed that T6SS$^{i\text{-}iii}$ formed a monophyletic group with high support (bootstrap supports are indicated at nodes). *Amoebophilus* T6S sheath, however, stably clustered in a group with the sheath from a gene cluster in *Cardinium* (structure and function unknown), and with the eCIS sheaths of Afp and MAC. The marked node was stable in all calculated trees as determined with different treeing methods and different components (fig. S9). (**B**) The *Amoebophilus* Afp-like gene cluster encodes all components that are conserved (blue) among all contractile injection systems (canonical T6SS/eCIS/phages), although it lacks any homologs of components that are specific for canonical T6SSs (yellow). Instead, the cluster harbors genes that are typical of eCISs and phages (red). Shown are two *Amoebophilus* genomic regions, locus tags, and gene annotations. See also table S4. (**C**) Schematic showing the major components of canonical T6SS, *Amoebophilus* T6SS^{iv}, and eCIS (homologs in same colors). T6SS^{iv} evolved either from eCISs, or alternatively, T6SS^{iv} represents a primordial system that gave rise to eCISs/phages/T6SS$^{i\text{-}iii}$. Both scenarios predict that T6SSs are more abundant than previously estimated (table S5).

Components that are exclusively conserved in canonical T6SSs (and absent in eCISs/phages) were not found, including TssJLM [trans-envelope complex (*9*)] or ClpV [sheath recycling (*20*)]. By contrast, the *Amoebophilus* cluster encodes proteins that were thought to be specific for eCISs and contractile phages. The length of the *Amoebophilus* T6SSs, for instance, is likely controlled by Aasi_1072/1806, which are homologs of tail terminator and tape measure proteins in Afps and phages (*21*, *22*). Indeed, sheath length can be predicted from TmP sequence (*22*) and correlates well with the length of *Amoebophilus* sheaths (fig. S10). Another example is an Rz-like endopeptidase (Aasi_1068), which usually co-occurs with a holin to mediate the release of eCISs and phages from the bacterial cytoplasm (*3*, *23*).

In conclusion, our structural and functional data showed that the *Amoebophilus* Afp-like gene cluster encodes a T6SS, whereas sequence analyses indicated a close relationship to eCISs. We therefore introduce the term "T6SS subtype 4" (T6SS^{iv}). In contrast to the distant relationships of T6SS$^{i\text{-}iii}$ to eCISs and phages that obstruct the reconstruction of an evolutionary path (*1*, *24*), we can hypothesize that T6SS^{iv} evolved from an Afp/MAC-like eCIS (independently of T6SS$^{i\text{-}iii}$) by the loss of tail fibers, loss of holin, and the establishment of interactions with the cytoplasmic membrane. Alternatively, T6SS^{iv} represents a primordial system from which eCISs, phages, and T6SS$^{i\text{-}iii}$ have evolved (Fig. 4C). Both scenarios predict that T6SSs are more abundant than previously thought. Indeed, T6SS^{iv}-like gene clusters were detected in six diverse bacterial phyla (table S5). The finding that diverse T6SS subtypes do not share a conserved gene set that would distinguish them from eCISs or phages emphasizes the necessity of an integrative approach to discover and characterize new systems.

REFERENCES AND NOTES

1. P. G. Leiman, M. M. Shneider, *Adv. Exp. Med. Biol.* **726**, 93–114 (2012).
2. Y. Uratani, T. Hoshino, *J. Bacteriol.* **157**, 632–636 (1984).
3. M. R. H. Hurst, T. R. Glare, T. A. Jackson, *J. Bacteriol.* **186**, 5116–5128 (2004).
4. N. J. Shikuma *et al.*, *Science* **343**, 529–533 (2014).
5. M. Basler, M. Pilhofer, G. P. Henderson, G. J. Jensen, J. J. Mekalanos, *Nature* **483**, 182–186 (2012).
6. A. Hachani, T. E. Wood, A. Filloux, *Curr. Opin. Microbiol.* **29**, 81–93 (2016).
7. Y.-W. Chang, L. A. Rettberg, D. R. Ortega, G. J. Jensen, *EMBO Rep.* **18**, 1090–1099 (2017).
8. M. Basler, *Philos. Trans. R. Soc. London B Biol. Sci.* **370**, 20150021 (2015).
9. E. Durand *et al.*, *Nature* **523**, 555–560 (2015).
10. M. Horn *et al.*, *Environ. Microbiol.* **3**, 440–449 (2001).
11. S. Schmitz-Esser *et al.*, *J. Bacteriol.* **192**, 1045–1057 (2010).
12. T. Penz, M. Horn, S. Schmitz-Esser, *Virulence* **1**, 541–545 (2010).
13. N. M. I. Taylor *et al.*, *Nature* **533**, 346–352 (2016).
14. M. M. Shneider *et al.*, *Nature* **500**, 350–353 (2013).
15. M. Marko, C. Hsieh, R. Schalek, J. Frank, C. Mannella, *Nat. Methods* **4**, 215–217 (2007).
16. A. Rigort *et al.*, *J. Struct. Biol.* **172**, 169–179 (2010).
17. T. Penz *et al.*, *PLOS Genet.* **8**, e1003012 (2012).
18. A. B. Russell *et al.*, *Cell Host Microbe* **16**, 227–236 (2014).
19. C. R. Büttner, Y. Wu, K. L. Maxwell, A. R. Davidson, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10174–10179 (2016).
20. G. Bönemann, A. Pietrosiuk, A. Diemand, H. Zentgraf, A. Mogk, *EMBO J.* **28**, 315–325 (2009).
21. D. Rybakova *et al.*, *Mol. Microbiol.* **89**, 702–714 (2013).
22. D. Rybakova, P. Schramm, A. K. Mitra, M. R. H. Hurst, *Mol. Microbiol.* **96**, 815–826 (2015).
23. I. N. Wang, D. L. Smith, R. Young, *Annu. Rev. Microbiol.* **54**, 799–825 (2000).
24. P. G. Leiman *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4154–4159 (2009).

ACKNOWLEDGMENTS

We thank F. Bosia, A. Harreither, P. Gasser, S. Rutz, P. Szwedziak, P. Tittmann, R. Wepf, and C. Zaubitzer for technical support. ScopeM is acknowledged for instrument access at ETH Zürich. We thank O. Medalia for microscope access at the University of Zürich. T. Ishikawa, R. Wepf,

and Pilhofer Lab members are acknowledged for discussions. R. Kostanjsek is acknowledged for generating preliminary data. W.-D. Hardt, H. Hilbi, R. Kooger, V. Panse, N. Shikuma, M. Swulius, and E. Tocheva are acknowledged for comments on the manuscript. We thank N. Taylor for providing us with the T4 model shown in Fig. 2B. Preliminary observations were made in the lab of G. J. Jensen. M.P. is supported by the European Research Council, the Swiss National Science Foundation, and the Helmut Horten Foundation. M.H. is supported by the European Research Council (ERC StG, no. 281633) and Austrian Science Fund (I 1628-B22). Subtomogram averages and tomograms were deposited in the Electron Microscopy Data Bank (accession numbers EMD-3791 and EMD-3793 to EMD-3801).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/357/6352/713/suppl/DC1

19 May 2017; accepted 17 July 2017
10.1126/science.aan7904

Chapter VII

Synthesis

The goal of this thesis was to decipher different aspects of amoeba-bacteria interaction, starting with a microbial diversity study to provide an overview and exemplify the relevance of this topic. I investigated not only the microbial but also the protist community structures and their seasonal variations in cooling towers, as they are known sources of legionellosis outbreaks. I then focused on two endosymbionts with different phylogeny to illustrate the various amoeba symbiosis. The first endosymbiont, the *Cochliophilus cryoturris*, was isolated from one of the cooling towers that we have studied in the previous chapter, further highlighting the importance of artificial water systems as a source for undiscovered amoeba-bacteria interactions (Tsao et al. 2019). The second was an endosymbiont of *Acanthamoeba* called *Procabacter acanthamoebae*. With only 90% 16S rRNA sequence similarity to known members within the Betaproteobacteria, the genus *Procabacter* was first proposed in 2002 (Horn et al. 2002). Their members reside within the cytosol of their amoeba host, as shown with fluorescence-in-situ hybridization and transmission electron microscopy. Here I've examined *Procabacter's* infection process, host range, host-free cultivation, the genomic features, and the evolution of this lineage extensively. The results provided insights into the diverse lifestyle of intracellular organisms, and mechanisms to occupy the niche. Last, I highlighted one key component of intracellular survival on a molecular basis. Here I focused on a novel secretion apparatus of the endosymbiont *Amoebophilus asiaticus*, which fulfills essential functions in host-bacteria communication.

FLA-bacteria interactions are manifold and not only have tremendous implications for the ecology of the surrounding environments but, more importantly, also for the field of public health. Human pathogens such as *Legionella pneumophila, Mycobacterium* spp., and *Pseudomonas aeruginosa* can all reside in man-made water systems, as the low temperature variations and constant influx of nutrients provide a suitable habitat for replication. Furthermore, the low water flow rates and the small diameters of the pipes benefit biofilm formation and further aid the propagation of human pathogens in these systems (Newsome et al. 1985; Thomas and McDonnell 2007; Mena and Gerba 2009; Abdel-Nour et al. 2013). In chapter III, we described in detail the microbial and protist community structure of three different cooling towers, not as a temporary snapshot, but as a year-long observation of their dynamics. I've found out that the cooling towers taxa resemble known freshwater microbial communities, which is consistent with the system's primary water source being drinking

water. Once thought of as a relatively stable system due to the semi-open design and constant water temperature, it was surprising to observe evident seasonal dynamics with large community shifts at all locations to different degrees. Interestingly, the microbial and protist changes do not necessarily agree with each other, probably due to their different reaction towards environmental stimuli. Another fact that needed to be highlighted was the high abundance of taxa traditionally associated with biofilms, especially *Pseudomonas*, *Methyloversatilis,* and *Blastomonas*. These taxa's occurrence and persistence hints at the significance of biofilm formation at all three locations. Coincidentally, some properties that make cooling towers well-suited for biofilm formation, also favor protists. Cooling tower basins offer a large surface area, rather constant and elevated temperature, and a slow water flow speed. The temperature supports not only higher microbial but also protists' growth. The low flow speed benefits the attachment of microbes on the surface, which in turn favors protist grazing and biofilm formation (Türetgen and Cotuk 2007). As biofilm matrices are aggregations of extracellular polysaccharides, proteins, lipids, DNA, and especially microbial cells, they further attract other protists (Limoli et al. 2015). This is problematic in two ways. First, biofilms contribute to resistance against disinfection procedures due to their multilayered structure, where treatments only affect the surface layer. Second, biofilms are known as reservoirs and potential sources for many bacterial pathogens. While *L. pneumophila* replicates primarily within environmental protists, they colonize and persist in the natural environment by biofilm formation and colonization within multispecies microbial communities (Abdel-Nour et al. 2013). As the attracted protists graze upon biofilms, they are also more likely to come in contact with *Legionella*, further aiding the growth of the pathogen. But not only pathogens are associated with biofilms, but also other amoeba-associated bacteria. The host-free cultivation data of *Procabacter acanthamoebae* showed biofilm formation under aerobic conditions (Chapter V). The formation of biofilm could be a stress response to restore microaerobic concentration as *P. acanthamoebae* has adapted to the intracellular niche. On the other hand, a host-free environment with elevated oxygen concentration could signal *Procabacter* to form biofilm to increase its probability of coming in contact with a potential host. Chapter III further emphasizes the role of biofilms in an aquatic niche in terms of pathogen propagation and their further implications for protists. With the improvement of next-generation sequencing speed and depth, the monitoring of variations of biofilm-associated taxa could be a suitable method for assessing the efficiency of different

disinfection strategies for removal of biofilms and reducing the number of protists. Next, specific microbial community patterns can serve as an early indicator for potential outbreaks, trigger disinfection measures, and save valuable time to take proper precautions.

The co-occurrence analysis described in chapter III captured only a fraction of potential interactions out there, as in a natural setting, bacterial symbionts are probably more promiscuous and thrive in a wide range of various protists. This is further highlighted in chapter IV, the description of *Cochliophilus cryoturris*. The endosymbiont in its natural host, *Cochliopodium minus,* was isolated from the water samples in one of the cooling towers through an agar-plate based isolation technique (Tsao et al. 2017). Not only is it the first endosymbiont found in a testate amoeba, but it is also the representative of a novel monophyletic group within the family Coxiellaceae. The family of Coxiellaceae includes several members with intracellular lifestyles, such as *Coxiella burnetii,* a human pathogen, and members of the genera *Rickettsiella* and *Diplorickettsia*, parasites and symbionts of arthropods (Cordaux et al. 2007). *Aquicella* species were first isolated from water and borehole samples and can infect *Vermamoeba vermiformis,* but are not dependent on them (Cordaux et al. 2007; Santos et al. 2003). *C. cryoturris* shares similar traits with its relatives *Coxiella burnetii* and *Legionella* spp., as they all reside in vacuole-like structures in their host's cytosol. Coincidentally many other obligate endosymbiont of amoeba such as *Vermiphilus pyriformis*, *Parachlamydia*, *Protochlamydia, Simkania negevensis*, *Waddlia chondrophila, Paracaedibacter, Jidaibacter acanthamoebae*, *Occultobacter vannellae,* and *Babela massiliensis* also remain in sub-compartments such as vacuoles or inclusions (Delafont et al. 2015; Pagnier et al. 2015; Schulz et al. 2016; Schulz et al. 2015; Pilhofer et al. 2014). Yet, there are exceptions, as *Amoebophilus asiaticus* and *Procabacter acanthamoebae* reside most of their time freely in the cytosol (Chapter V, Chapter VI, Horn et al. 2002).

Why do some bacteria reside in vacuoles whereas others are directly in the cytosolic space? The answer might have to do with how amoebae hunt for their food. Amoebae take up food bacteria by a process called phagocytosis, where the bacteria are internalized in vacuole-like structures called phagosomes (Wetzel and Korn 1969). When the phagosome fuses with lysosomes, a maturation process where a phagolysosome is formed, it will lead to acidification and degradation of bacteria. However, pathogens as well as facultative and obligate

intracellular bacteria have developed similar strategies to survive or escape the phagocytosis process (Cosson and Soldati 2008). The prevention of digestion can either happen by inhibition of the maturation process or by escape from the phagosome. This vital function is often facilitated by a type of bacterial molecular apparatus called a secretion system (Green and Mecsas, 2016). The system can deliver bacterial molecules across the host membrane, which are then recognized and thus disrupt the regular phagocytic function. In general, secretion systems can have multiple ways of promoting bacterial virulence, ranging from attachment enhancement to host cells, scavenging nutrients, to killing of target cells, and disruption of cellular functions. Today, we differentiate between eight distinct forms and termed them "secretion system of type I - VII and IX" (Lasica et al. 2017; Green and Mecsas 2016). Describing each type in detail would largely exceed the scope of this chapter. However, known endosymbionts of amoeba either use type I secretion systems (*Procabacter* spp.), type II secretion systems (*Babela massiliensis, Procabacter* spp., *Cytomitobacter, Nesciobacter*), type III secretion systems (Chlamydiae endosymbiont of amoeba, *Procabacter* spp.), type IV secretion systems (*Jidaibacter acanthamoebae, Cytomitobacter, Nesciobacter)*, type V secretion system (*Procabacter* spp.), and variations of type VI secretion systems (*Amoebophilus asiaticus,* endosymbionts of diplonemid; George et al. 2020), as discussed in the chapter VI.

The novel type VI secretion system (T6SS) of the subtype IV described in chapter VI, was initially characterized as an antifeeding-prophage (AFP)-like component in *Amoebophilus asiaticus* (Penz et al. 2010). Structural and functional wise, there are many similarities between T6SS and extracellular contractile injection system (eCIS) of phages. The main difference is the anchoring of T6SS to the inner membrane, whereas eCIS is released extracellularly to bind to the target's cell (Chen et al. 2019). The *Amoebophilus*' T6SS consists of an anchor protein, a base plate, and a sheath-protein with an inner tube. Screening of extracellular space resulted only in the presence of tube but no sheath proteins. *Amoebophilus* switches between two life phases, the replicative phase, and the infective phase. Functional experiments comparing the ability of *Amoebophilus* from different life phases to establish new infections showed that host infection rates positively correlated with T6SS expression. Further experiment has also shown a positive correlation between red blood cell lysis and T6SS expression. Interestingly, phylogenetic analysis consistently places *Amoebophilus*' T6SS stably with other antifeeding prophages and phage-derived structures into the same

monophyletic group. As structural and functional data showed that the *Amoebophilus*' Afp-like gene cluster encodes a T6SS, the sequence analyses indicated a close relationship to eCISs. Therefore the term T6SS of subtype 4 (T6SS*iv*) has been proposed. T6SS*iv* either (i) evolved from Afp-like apparatus by loss of specific subunits and establishment with the inner membrane, or (ii) represents a primordial system from which eCISs and the canonical T6SS have evolved. The result challenged what we previously thought as a T6SS and indicated a higher abundance, as we detected similar gene clusters in six diverse bacterial phyla.

An amoeba endosymbiont, which encodes a diverse set of the aforementioned secretion system is the *Procabacter acanthamoebae.* Our experimental effort provides detailed insights into the life cycle of *Procabacter acanthamoebae* (Chapter V). In parallel, we analyzed the genomic data sets, further extending our knowledge of this unique symbiosis (Chapter V). Phylogenetic analysis in combination with amino acid identity (AAI) data has shown that *Procabacter* spp. are closely related to members of the families Neisseriaceae and Chromobacteriaceae (Chapter V). While both families mainly consist of free-living bacteria, some species are considered pathogenic facultative intracellular organisms. On the other hand, the proposed family *Procabacteriaceae* consists only of obligate endosymbionts associated with *Acanthamoeba*. Genomes of intracellular organisms are mostly undergoing reduction, starting with the abolishment of metabolic genes, as the nutrients are easily obtained from the host. The *Procabacter* genomes encompass a surprisingly diverse metabolic capability, and the majority of genes in the predicted metabolic pathways are present. Metabolic reconstruction also revealed species-specific diversity to synthesize amino acids, suggesting *Procabacter* spp. relies on amino acid uptake from the host. From an evolutionary point of view, there is evidence that Procabacter genomes are still undergoing rearrangements. Ancestral genome analysis unveiled not only does the *Procabacter* group, in comparison to the related Neisseriaceae and Chromobacteriaceae group, exhibit a high number of lost gene families at an extraordinary rate. At the same time, it also acquired a relatively large number of new gene families. Among the newly acquired families are many mobile genetic elements such as transposases, prophages, and integrases. Mobile genetic elements usually trigger horizontal gene transfers, which are the driving force of genome evolution. By introducing novel functional genes and disruption of current ones, it will eventually lead to fixation of benefitting features and loss of unneeded functions. Further

analysis of the gained gene families revealed some interesting results, as many of them are effector proteins containing eukaryotic-like domains, probably delivered by the type III secretion system for interaction with eukaryotic host cells. Combined with the observation that *Procabacter* encodes (i) multiple secretion systems (type I, II, III, and V), (ii) factors enhancing host cell adhesion, (iii) many other predicted effector proteins, and (iv) two distinct cytochrome oxidases to cope with dynamic oxygen concentrations, we conclude that *Procabacter* very likely undergo evolution towards an endosymbiotic lifestyle.

References

Abdel-Nour, M., Duncan, C., Low, D.E., and Guyard, C. (2013) Biofilms: The stronghold of Legionella pneumophila. *Int J Mol Sci* **14**: 21660–21675.

Chen, L., Song, N., Liu, B., Zhang, N., Alikhan, N.F., Zhou, Z., et al. (2019) Genome-wide Identification and Characterization of a Superfamily of Bacterial Extracellular Contractile Injection Systems. *Cell Rep* **29**: 511-521.e2.

Cordaux, R., Paces-Fessy, M., Raimond, M., Michel-Salzat, A., Zimmer, M., and Bouchon, D. (2007) Molecular characterization and evolution of arthropod-pathogenic Rickettsiella bacteria. *Appl Environ Microbiol* **73**: 5045–5047.

Cosson, P. and Soldati, T. (2008) Eat, kill or die: when amoeba meets bacteria. *Curr Opin Microbiol* **11**: 271–276.

Delafont, V., Samba-Louaka, A., Bouchon, D., Moulin, L., and Héchard, Y. (2015) Shedding light on microbial dark matter: A TM6 bacterium as natural endosymbiont of a free-living amoeba. *Environ Microbiol Rep* **7**: 970–978.

George, E.E., Husnik, F., Tashyreva, D., Prokopchuk, G., Horák, A., Kwong, W.K., et al. (2020) Highly reduced genomes of protist Endosymbionts show evolutionary convergence. *Curr Biol* **30**: 925-933.e3.

Green, E.R. and Mecsas, J. (2016) Bacterial Secretion Systems: An Overview. *Microbiol Spectr* **4**:.

Horn, M., Fritsche, T.R., Linner, T., Gautom, R.K., Harzenetter, M.D., and Wagner, M. (2002) Obligate bacterial endosymbionts of Acanthamoeba spp. related to the β-Proteobacteria: Proposal of "Candidatus Procabacter acanthamoebae" gen. nov., sp. nov. *Int J Syst Evol Microbiol* **52**: 599–605.

Lasica, A.M., Ksiazek, M., Madej, M., and Potempa, J. (2017) The Type IX Secretion System (T9SS): Highlights and Recent Insights into Its Structure and Function. *Front Cell Infect Microbiol* **7**:.

Limoli, D.H., Jones, C.J., and Wozniak, D.J. (2015) Bacterial Extracellular Polysaccharides in Biofilm Formation and Function. *Microbiol Spectr* **3**:.

Mena, K.D. and Gerba, C.P. (2009) Risk assessment of pseudomonas aeruginosa in water. *Rev Environ Contam Toxicol* **201**: 71–115.

Newsome, A.L., Baker, R.L., Miller, R.D., and Arnold, R.R. (1985) Interactions between Naegleria fowleri and Legionella pneumophila. *Infect Immun* **50**: 449–452.

Pagnier, I., Yutin, N., Croce, O., Makarova, K.S., Wolf, Y.I., Benamar, S., et al. (2015) Babela massiliensis, a representative of a widespread bacterial phylum with unusual adaptations to parasitism in amoebae. *Biol Direct* **10**: 13.

Penz, T., Horn, M., and Schmitz-Esser, S. (2010) The genome of the amoeba symbiont "candidatus amoebophilus asiaticus" encodes an afp-like prophage possibly used for protein secretion. *Virulence* **1**: 541–545.

Pilhofer, M., Aistleitner, K., Ladinsky, M.S., König, L., Horn, M., and Jensen, G.J. (2014) Architecture and host interface of environmental chlamydiae revealed by electron cryotomography. *Environ Microbiol* **16**: 417–429.

Schulz, F., Martijn, J., Wascher, F., Lagkouvardos, I., Kostanjšek, R., Ettema, T.J.G., and Horn, M. (2016) A Rickettsiales symbiont of amoebae with ancient features. *Environ Microbiol* **18**: 2326–42.

Schulz, F., Tyml, T., Pizzetti, I., Dyková, I., Fazi, S., Kostka, M., and Horn, M. (2015) Marine amoebae with cytoplasmic and perinuclear symbionts deeply branching in the Gammaproteobacteria. *Sci Rep* **5**: 13381.

Thomas, V. and McDonnell, G. (2007) Relationship between mycobacteria and amoebae: Ecological and epidemiological concerns. *Lett Appl Microbiol* **45**: 349–357.

Tsao, H.F., Scheikl, U., Volland, J.M., Köhsler, M., Bright, M., Walochnik, J., and Horn, M. (2017) "Candidatus Cochliophilus cryoturris" (Coxiellaceae), a symbiont of the testate amoeba Cochliopodium minus. *Sci Rep* **7**: 3394.

Türetgen, I. and Cotuk, A. (2007) Monitoring of biofilm-associated Legionella pneumophila on different substrata in model cooling tower system. *Environ Monit Assess* **125**: 271–279.

Wetzel, M.G. and Korn, E.D. (1969) Phagocytosis of latex beads by Acahamoeba castellanii (Neff). 3. Isolation of the phagocytic vesicles and their membranes. *J Cell Biol* **43**: 90–104.