

ORIGINAL ARTICLE

Life in an unusual intracellular niche: a bacterial symbiont infecting the nucleus of amoebae

Frederik Schulz¹, Ilias Lagkouvardos¹, Florian Wascher¹, Karin Aistleitner¹, Rok Kostanjšek² and Matthias Horn¹

¹Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna, Vienna, Austria and ²Department of Biology, University of Ljubljana, Ljubljana, Slovenia

Amoebae serve as hosts for various intracellular bacteria, including human pathogens. These microbes are able to overcome amoebal defense mechanisms and successfully establish a niche for replication, which is usually the cytoplasm. Here, we report on the discovery of a bacterial symbiont that is located inside the nucleus of its *Hartmannella* sp. host. This symbiont, tentatively named ‘*Candidatus Nucleicultrix amoebiphila*’, is only moderately related to known bacteria (~90% 16S and 23S rRNA sequence similarity) and member of a novel clade of protist symbionts affiliated with the *Rickettsiales* and *Rhodospirillales*. Screening of 16S rRNA amplicon data sets revealed a broad distribution of these bacteria in freshwater and soil habitats. ‘*Candidatus Nucleicultrix amoebiphila*’ traffics within 6 h post infection to the host nucleus. Maximum infection levels are reached after 96–120 h, at which time point the nucleus is pronouncedly enlarged and filled with bacteria. Transmission of the symbionts occurs vertically upon host cell division but may also occur horizontally through host cell lysis. Although we observed no impact on the fitness of the original *Hartmannella* sp. host, the bacteria are rather lytic for *Acanthamoeba castellanii*. Intranuclear symbiosis is an exceptional phenomenon, and amoebae represent an ideal model system to further investigate evolution and underlying molecular mechanisms of these unique microbial associations.

The ISME Journal (2014) 8, 1634–1644; doi:10.1038/ismej.2014.5; published online 6 February 2014

Subject Category: Microbe-microbe and microbe-host interactions

Keywords: symbiosis; nucleus; endonuclear; Rickettsia; *Hartmannella*; *Acanthamoeba*

Introduction

Free-living amoebae are important members of microbial communities in terrestrial and aquatic habitats, where they interact with other microorganisms in various ways. They primarily feed on bacteria, fungi and other protists, and thus represent main predators controlling the microbial populations (Rodríguez-Zaragoza, 1994). However, continuous grazing pressure has facilitated the evolution of mechanisms allowing some bacteria to evade degradation during phagocytosis and to thrive within amoebae (Barker and Brown, 1994; Casadevall, 2008). After uptake, these bacteria either transiently exploit their eukaryotic host eventually leading to its lysis, or they establish long-term relationships with these protists (Molmeret *et al.*, 2005).

Many bacteria that are able to infect and propagate within amoebae are well-known human pathogens, such as *Legionella pneumophila* (Swanson and

Hammer, 2000), *Francisella tularensis* (Abd and Johansson, 2003) and *Mycobacterium tuberculosis* (Mba Medie *et al.*, 2011). In addition, amoebae contain frequently obligate intracellular symbionts that appear to be cosmopolitan in distribution (Fritsche *et al.*, 1993; Horn and Wagner, 2004; Schmitz-Esser *et al.*, 2008). These include members of diverse evolutionary lineages, most notably the *Chlamydiae* (Greub *et al.*, 2003; Horn, 2008). The study of these amoeba-associated microbes provided novel insights in the evolution of pathogenicity and the molecular mechanisms underlying the interaction with eukaryotic host cells (Cazalet *et al.*, 2004; Horn *et al.*, 2004; Collingro *et al.*, 2005).

Intracellular bacteria generally replicate either within host-derived vacuoles or directly in the cytoplasm (Schmitz-Esser *et al.*, 2008; Friedrich *et al.*, 2012). For example, *L. pneumophila* blocks phagosome-lysosome fusion, modulates the endosome by manipulating host processes and finally uses it for replication (Isberg *et al.*, 2009). In contrast, *F. tularensis* is able to escape the phagosome and replicates inside the host cytoplasm (Santic *et al.*, 2009). There are only rare cases where other intracellular niches are exploited. *Midichloria mitochondrii* targets mitochondria (Sassera *et al.*, 2006), *Rickettsia bellii* and *Endonucleobacter*

Correspondence: M Horn, Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna, Althanstraße 14, 1090 Vienna, Austria.

E-mail: horn@microbial-ecology.net

Received 2 October 2013; revised 13 December 2014; accepted 1 January 2014; published online 6 February 2014

bathymodioli infect the nucleus of ticks or deep sea mussels (Ogata *et al.*, 2006; Zielinski *et al.*, 2009), and *Holospora* and *Caedibacter* species thrive in the nuclei of *Paramecium* hosts (Goertz, 1986; Fokin, 2004).

Here, we report the first isolation and in-depth characterization of bacteria thriving in the nucleus of free-living amoebae. These symbionts, for which we propose the name '*Candidatus Nucleicultrix amoebiphila*', are members of a novel clade within the *Alphaproteobacteria*. We describe the infection cycle of these bacteria, analyze their host range and their impact on host fitness, and we show that this novel symbiont clade represents a diverse group of bacteria with global distribution.

Materials and methods

Isolation and cultivation of amoebae

An activated sludge sample taken from a nitrifying bioreactor (maintained in our laboratory) was applied to a non-nutrient agar plate based on Page's amoebic saline (PAS; 0.12 g l⁻¹ NaCl, 0.004 g l⁻¹ MgSO₄ × 7 H₂O, 0.004 g l⁻¹ CaCl₂ × 2H₂O, 0.142 g l⁻¹ Na₂HPO₄, 0.136 g l⁻¹ KH₂PO₄) containing a layer of heat-inactivated *Escherichia coli* (Page, 1988). Amoebae propagating on these plates were repeatedly transferred to new non-nutrient agar plates seeded with heat-inactivated *E. coli* and incubated at room temperature. Finally, amoebae were transferred to culture flasks (Nunclon delta-surface, Thermo scientific, St Leon Rot, Germany) containing PAS and live *E. coli tolC*⁻ as well as ampicillin (200 ng ml⁻¹). Amoeba cultures were screened using fluorescence *in situ* hybridization (FISH) for the presence of bacterial symbionts.

To obtain clonal cultures with a high percentage of intracellular bacteria, single amoeba from cultures containing bacterial symbionts was picked with a micromanipulator (Eppendorf, Hamburg, Germany) and placed in PAS containing *E. coli tolC*⁻ and ampicillin (200 ng ml⁻¹) in 96-well culture dishes (Corning, Corning, NY, USA). Cultures were screened using FISH; uninfected cultures were pooled and used as symbiont-free host cells.

Hartmannella vermiformis A1Hsp, *Acanthamoeba castellanii* Neff, *Acanthamoeba polyphaga* DOME, *Acanthamoeba* sp. C1, *Acanthamoeba* sp. 5a2, *Acanthamoeba* sp. UWC8 and *Naegleria gruberi* NEG-M used for infection experiments were incubated at 24 °C in PAS supplemented with *E. coli tolC*⁻ and ampicillin (200 ng ml⁻¹).

Transmission electron microscopy

Amoebae were detached from culture flasks by shaking, and concentrated by centrifugation (3000 g, 8 min). The supernatant was discarded and samples were fixed for transmission electron microscopy (TEM) in 3.5% glutaraldehyde in 0.1 M

phosphate buffer (pH 7.2), washed in 0.1 M phosphate buffer (pH 7.2), post fixed in 1% OsO₄ for 1 h and dehydrated in an increasing ethanol series. Dehydrated samples were embedded in agar 100 resin and cut. Ultrathin sections were stained with uranyl acetate and Reynold's lead citrate before examination with a Philips CM 100 transmission electron microscope (FEI, Hillsboro, OR, USA) operating at 80 kV.

Fluorescence in situ hybridization

Amoeba cells were harvested using centrifugation (3000 g, 8 min), washed with PAS and applied on microscope slides for 30 min to allow attachment of cells before fixation with 4% formaldehyde (15 min at room temperature). A symbiont-specific probe (CBR125, 5'-ACCATTCGGCATGTTCCC-3') was designed using the probedesign/probematch tools of the ARB software package and deposited at probeBase (Ludwig *et al.*, 2004; Loy *et al.*, 2007). Optimal hybridization conditions for symbiont-specific probes were determined in a series of hybridization experiments with increasing formamide concentrations in the hybridization buffer. In addition, probe EUK516 (5'-ACCAGACTTGCCC TCC-3', Amann and Binder, 1990) targeting most eukaryotes, the probe mix EUB338 I-III (5'-GCTGCC TCCCGTAGGAGT-3', 5'-GCAGCCACCCGTAGGTGT-3', 5'-GCTGCCACCCGTAGGTGT-3; Amann and Binder, 1990; Daims *et al.*, 1999) targeting most bacteria and the negative control probe NONEUB (5'-ACTCCTAC GGGAGGCAGC-3') were used. All probes were purchased from Thermo Fisher Scientific (St Leon Rot, Germany). Hybridization was performed overnight at 46 °C at a formamide concentration of 20% using standard hybridization and washing buffers (Daims *et al.*, 2005). Cells were subsequently stained with 4', 6-diamidino-2-phenylindole (0.5 µg ml⁻¹ in PAS, 3 min), washed once with PAS and embedded in Citifluor (Agar-Scientific, Stansted, UK). Slides were examined using a confocal laser scanning microscope (LSM 510 Meta, Zeiss, Oberkochen, Germany) or an epifluorescence microscope (Axio-plan 2 imaging, Zeiss) equipped with a CCD camera (AxioCam HRC, Zeiss).

DNA extraction, polymerase chain reaction (PCR), cloning and sequencing

DNA was extracted from infected amoeba cultures using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Amoebal 18S rRNA genes were amplified by PCR using primers 18SF (5'-GTAGTCA TATGCTTGCTCTC-3') and 18SR (5'-CGRARACCTTG TTACGAC-3'; Schmitz-Esser *et al.*, 2008) at an annealing temperature of 50 °C. Bacterial 16S rRNA genes were amplified using primers 616V (5'-AGAG TTTGATYMTGGCTCAG-3') and 1492R (5'-GGYTAC CTTGTTACGACTT-3') at an annealing temperature of 52 °C (Juretschko *et al.*, 1998; Loy *et al.*, 2002).

Bacterial 23S rRNA genes were amplified using primers 129F (5'-CYGAATGGGRVAACC-3') and 2490R (5'-CGACATCGAGGTGCCAAAC-3') at an annealing temperature of 56 °C (Hunt and Klepac-Ceraj, 2006). PCRs typically contained 100 ng template DNA, 50 pmol of each primer, 1 unit of Taq DNA polymerase (Fermentas, St Leon Rot, Germany), 10 × Taq buffer with KCl and 2 μM MgCl₂ and 0.2 μM of each deoxynucleotide in a total volume of 50 μl. PCR products were purified using the PCR Purification Kit (Qiagen) and were either sequenced directly or cloned using the TOPO TA Cloning Kit (Invitrogen, Darmstadt, Germany) following the manufacturers' instructions. Nucleotide sequences were determined using the BigDye Terminator kit v3.1 (Applied Biosystems, Vienna, Austria) and an ABI 3130 XL genetic analyzer (Applied Biosystems). Newly obtained rRNA gene sequences were deposited at Genbank/EMBL/DDBJ under accession numbers KF697195, KF697196 and KF697197.

Infection experiments

Supernatant of infected amoebae was collected and passed through a 1.2-μm filter (Sartorius, Goettingen, Germany). The filtrate was centrifuged at 10 500 g for 10 min, and the resulting bacterial pellet was washed once with PAS, resuspended in PAS and directly used for infection experiments. Bacterial cell numbers were estimated by counting DAPI-stained bacteria in a Neubauer counting chamber.

All experiments were performed at room temperature, if not stated otherwise. Amoebae were seeded into 12-well plates (Thermoscientific, Hvidovre, Denmark; 3 × 10⁵ cells per well), and bacteria at an estimated multiplicity of infection of 500 were added. After 6 h, amoeba cells were washed twice with PAS to remove extracellular bacteria. At different time points post infection (p.i.), amoebae were detached and amoeba cell numbers were determined with a Neubauer counting chamber. Half of the cell suspension was then used for FISH, and the other half for viability staining.

The percentage of amoebae-containing bacteria in the nucleus was determined by analyzing FISH-stained samples. At least 100 amoeba cells were counted per sample. To assess the viability of amoeba cells during infection, amoebae were incubated with propidium iodide (PI; 1.5 μM; Molecular Probes, Eugene, OR, USA) in PAS for 20 min at room temperature in the dark. Samples were washed once with PAS and transferred to a black plate reader dish (Greiner Bio-One, Frickenhausen, Germany), and fluorescence intensities were determined with a Tecan Infinite M200 plate reader (Tecan, Groeding, Austria). Negative controls included PAS and suspensions of uninfected amoebae.

For each time point and condition, at least three replicate infection experiments were analyzed. To test for statistically significant differences ($P < 0.05$)

between different conditions and time points, a two-way analysis of variance including the Bonferroni post test was performed using GraphPad Prism (version 6; La Jolla, CA, USA).

Phylogenetic analysis

Phylogenetic analyses were conducted using MEGA version 5 (Tamura *et al.*, 2011) and ARB (Ludwig *et al.*, 2004). The alignment editor integrated in ARB was used to build alignments based on the current Silva ARB 16S and 23S rRNA databases (Quast *et al.*, 2013), which were updated with sequences from GenBank/EMBL/DDBJ obtained by sequence homology searches using BLASTn available at the NCBI web site (National Center for Biotechnology Information; Altschul *et al.*, 1997; Wheeler *et al.*, 2008). The alignment was trimmed to the length of the shortest sequence, manually curated and exported from ARB using a 50% conservation filter. The resulting alignments comprised 103 sequences and 1252 positions for the 16S rRNA and 75 sequences including 2235 positions for the 23S rRNA; the concatenated 16S–23S rRNA data set included 3487 positions. MEGA was used to find the optimal maximum likelihood model, and the GTR + I + G model of evolution was selected. Maximum likelihood trees were calculated in MEGA. To evaluate the robustness of the tree topology, 1000 bootstrap resamplings were performed. In addition, MrBayes 3.2 (Ronquist *et al.*, 2012) was used for tree reconstruction. The Bayesian analysis included four Markov chains, consisting of 5 000 000 generations and a sampling every 100 generations. Two simultaneous runs with different random start trees were performed, and the first 25% of samples were discarded. The resulting phylogenetic trees were visualized in iTOL (Letunic and Bork, 2007).

Screening of 16S rRNA gene amplicon data sets

The screening of amplicon data sets is based on a recently published approach (Lagkouvardos *et al.*, 2013). Briefly, all raw 454-based 16S rRNA gene amplicon sequence data from environmental samples in the databases SRA (Kodama *et al.*, 2012) and VAMPS (Huse *et al.*, 2010; <http://vamaps.mbl.edu/>) were extracted and organized by sample in independent data sets. The databases were searched using BLAST (Altschul *et al.*, 1997) and the full-length 16S rRNA gene sequence of '*Candidatus Nucleicultrix amoebiphila*' as query. The detected amplicon sequences were filtered with respect to size (>200 nucleotides) and alignment length (>80% of their size). Extracted sequences that were at least 95% similar to the 16S rRNA of '*Candidatus Nucleicultrix amoebiphila*' and that had a length of at least 300 nucleotides were assigned to either of the three data sets covering different regions of a reference alignment including the full-length sequence of '*Candidatus Nucleicultrix amoebiphila*'

and the most similar sequences from GenBank/EMBL/DDBJ (Wheeler *et al.*, 2008). To decrease the complexity of the data sets and to account for artifacts caused by sequencing errors, each set of sequences was clustered with UCLUST at a 97% threshold (Edgar, 2010). Singletons were omitted, and in the case where operational taxonomic units contained multiple sequences from the same sample, only one sequence representative for this sample was kept. Sequences were aligned using SSU-ALIGN for each of the three data sets (Nawrocki and Adviser-Eddy, 2009), manually curated and trimmed to the length of the shortest sequence in Jalview (Waterhouse *et al.*, 2009). Phylogenetic analysis was performed with FastTree2 using the GTR model and the gamma distribution with 20 categories option (Price *et al.*, 2010). The resulting phylogenetic tree was visualized in iTOL (Letunic and Bork, 2007).

Results

Isolation of a Hartmannella sp. with bacterial symbionts in the nucleus

During a routine screening of a nitrifying bioreactor for free-living amoebae, we obtained an amoeba isolate morphologically resembling *Hartmannella*

species. This was confirmed by 18S rRNA gene sequencing, demonstrating that this amoeba strain, termed *Hartmannella* sp. FS5, is closely related to known *Hartmannella* species (Supplementary Figure S1). Subsequent analysis of *Hartmannella* sp. FS5 using FISH demonstrated the presence of bacterial endosymbionts. Surprisingly, we observed that the bacterial FISH signal overlapped with DAPI (4', 6-diamidino-2-phenylindole)-stained amoeba nuclei, although no bacteria were detected in the cytoplasm (Supplementary Figure S2). This suggested that the bacterial symbionts in *Hartmannella* sp. FS5 are located in the host nucleus. TEM confirmed the intranuclear location (Figures 1a–c). Coccoid rods with a length of 0.5–1 µm, a diameter of 0.3–0.4 µm and a Gram-negative-type cell wall were observed directly in the nucleoplasm of infected amoebae.

A novel symbiont related to the Rickettsiales and the Rhodospirillales

Sequencing of the 16S and 23S rRNA genes revealed that the *Hartmannella* sp. FS5 symbionts are novel and only distantly related to known bacteria (~90% sequence similarity). Phylogenetic analysis showed that together with other symbionts of amoebae and

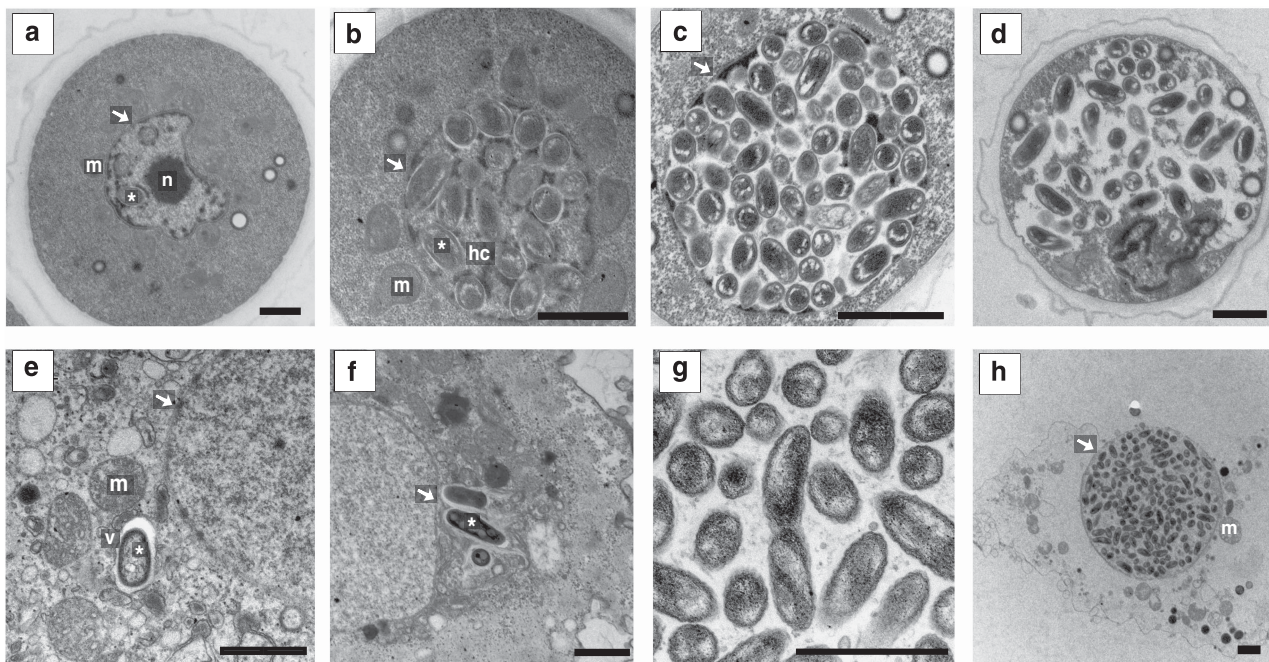


Figure 1 The intracellular niche of ‘*Candidatus Nucleicultrix amoebiphila*’. Transmission electron micrographs showing different infection stages of *Hartmannella* sp. (a–d) and *A. castellanii* (e–h). (a) Early infection stage in *Hartmannella* with three *Nucleicultrix* particles present in the nucleus, bacteria are surrounded by host electron-dense heterochromatin. (b) Intermediate infection stage; the nucleus is filled with bacteria. (c) Late infection stage with an enlarged nucleus. (d) Intact cyst containing symbionts; the nuclear membrane is ruptured. (e) Early infection stage in *A. castellanii*; after uptake into the cytoplasm, the bacteria remain enclosed by a vacuole and are surrounded by an electron-translucent space indicating the presence of extracellular polymeric substances. (f) Symbionts in close proximity to the host nuclear membrane; no clear phagosomal membrane can be recognized. (g) Ultrastructure of the symbionts inside an *Acanthamoeba* sp. nucleus; coccoid rods with a Gram-negative-type cell wall can be seen, one dividing cell is shown. (h) Lysis of an *A. castellanii* host cell; the nucleus is densely populated by bacteria, while the nuclear membrane is still intact. Asterisks indicate ‘*Candidatus Nucleicultrix amoebiphila*’; arrows indicate the nuclear membrane; m, mitochondria; n, nucleolus; hc, heterochromatin; v, vacuole. Bars, 1 µm.

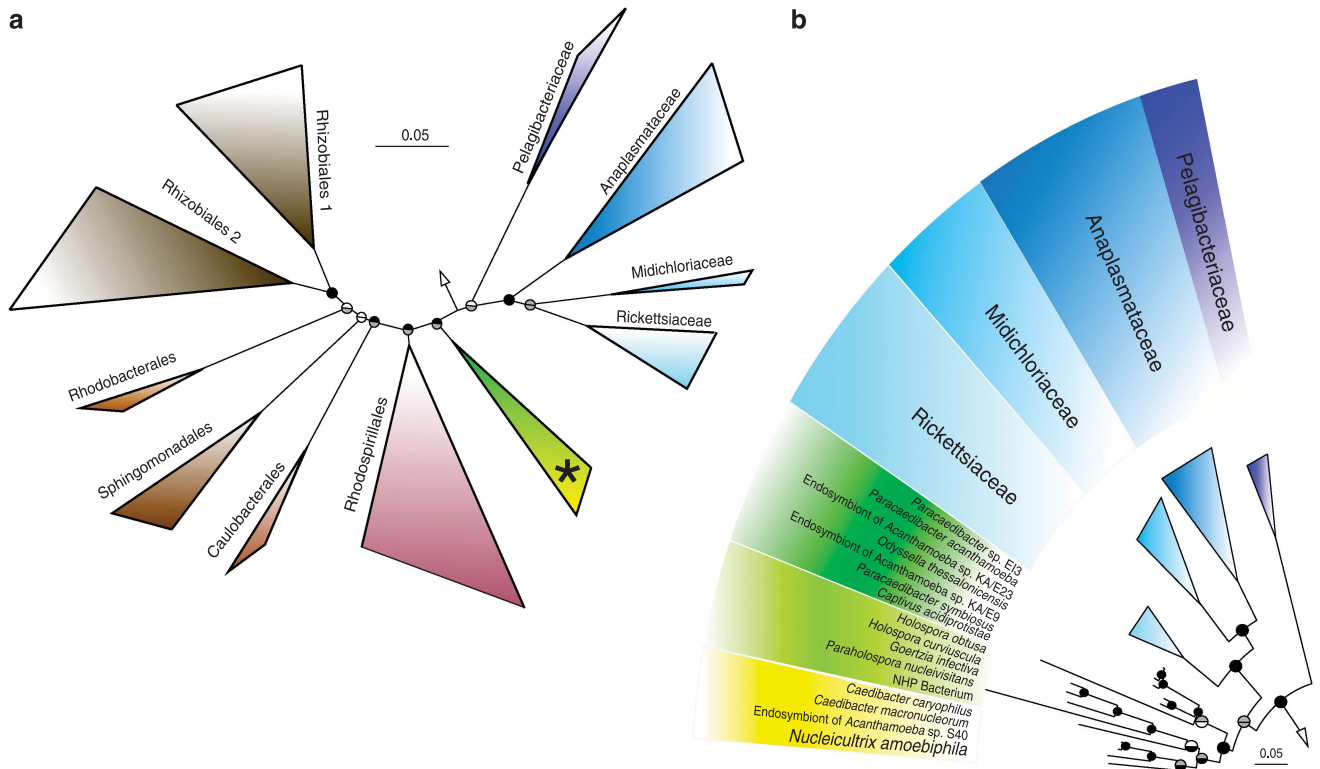


Figure 2 Phylogenetic relationship of ‘*Candidatus Nucleicultrix amoebiphila*’ with the *Alphaproteobacteria*. (a) Maximum likelihood tree based on a concatenated 16S–23S rRNA alignment. ‘*Candidatus Nucleicultrix amoebiphila*’ (indicated by an asterisk) together with other protist symbionts forms a deeply branching lineage within the *Alphaproteobacteria*. (b) 16S rRNA-based maximum likelihood tree showing members of the protist symbiont clade containing ‘*Candidatus Nucleicultrix amoebiphila*’. Circles displayed at the nodes represent bootstrap values (upper half) and Bayesian posterior probability values (lower half), respectively. Black indicates support values >80% or probability values >0.8; gray indicates support values of 50–80% or probability values of 0.5–0.8, white indicates no support. Fully resolved trees are shown in Supplementary Figures S3–S5; accession numbers of sequences used for phylogenetic analyses are listed in Supplementary Table S1.

ciliates they form a deep-branching lineage in the *Alphaproteobacteria* (Figure 2). However, the relationship of this symbiont clade with other *Alphaproteobacteria* could not be resolved unambiguously. Different data sets (16S rRNA, 23S rRNA or both genes concatenated) and different treeing methods resulted in conflicting, yet generally well-supported, tree topologies. The symbiont clade branched off within the *Rickettsiales* in 16S rRNA-based trees but grouped together with the *Rhodospirillales* in the 23S rRNA-based trees and in the trees inferred with the concatenated genes (Figure 2, Supplementary Figures S3–S5). The name ‘*Candidatus Nucleicultrix amoebiphila*’ is proposed for tentative classification of these novel intranuclear bacteria (a formal description is provided as Supplementary Text S1).

‘*Candidatus Nucleicultrix amoebiphila*’ also infects *Acanthamoeba* spp.

We next tested whether ‘*Candidatus Nucleicultrix amoebiphila*’ (hereafter referred to as *Nucleicultrix*) is able to infect amoebae other than its original *Hartmannella* sp. FS5 host. The symbionts were purified and added to various *Acanthamoeba*,

Hartmannella and *Naegleria* strains; amoeba cultures were screened with FISH 7 days p.i. *Nucleicultrix* infected *Hartmannella vermiformis* and four *Acanthamoeba* strains with different efficiencies, with *A. castellanii* Neff being the most susceptible *Acanthamoeba* strain. In these amoebae, *Nucleicultrix* was always located within the host nucleus (Supplementary Figure S7). It should be noted, however, that successful infection of *Hartmannella* spp. was dependent on the simultaneous presence of *E. coli* as food source (Supplementary Figure S8), and that no infection was observed in nutrient-rich media. Infection with *Nucleicultrix* was not detected for *Acanthamoeba* sp. UWC8 and *Naegleria gruberi*. Growth of purified *Nucleicultrix* was also not observed under host-free conditions using a variety of standard laboratory media.

Close association of *Nucleicultrix* with host cell heterochromatin

To obtain more detailed insights into the lifestyle of *Nucleicultrix*, we examined *Hartmannella* sp. and *A. castellanii* at different infection stages using TEM. In both amoebae, the symbiont was only rarely seen in the host cytoplasm, suggesting that

after uptake *Nucleicultrix* migrates to the nucleus and does not replicate in the cytoplasm. Bacteria in the cytoplasm were mostly enclosed in vacuoles (Figures 1e and f), which contained single or few bacterial cells often surrounded by an electron-translucent space, which might indicate the presence of extracellular polymeric substances. Within the host nucleus, the symbionts colocalized with electron-dense heterochromatin (Figures 1a and b). These structures more and more disappeared in strongly infected nuclei, and areas of white, unstructured electron-translucent space increased (Figure 1c). The infection leads to highly enlarged nuclei, densely packed with bacteria. Infected nuclei may burst (Figure 1d) or stay intact after host cell lysis (Figure 1h).

Infection cycle and impact on amoeba hosts

Monitoring of the infection of *Hartmannella* and *A. castellanii* using FISH and staining of endocytic compartments showed that *Nucleicultrix* is taken up by phagocytosis, resides within the phagosome and survives acidification by lysosome fusion. The symbionts escape from the phagolysosome and are detected as individual cells in the cytoplasm as early as 2 h p.i (Supplementary Figure S6). *Nucleicultrix* was first seen within the nuclear compartment 4–6 h (Figure 3, Supplementary Figure S7) p.i. Following invasion of the nucleus, the symbionts start replicating, and after 24 h, two to four bacteria were present per nucleus. Up to 32 bacteria could be found 48 h p.i., suggesting a doubling time of about 8 h at this stage of infection. Nuclei were densely

packed with *Nucleicultrix* after 72–96 h, and this compartment strongly increased in size at 120 h.

Host lysis was the predominant result of infection of *A. castellanii*, occurring from 72 h p.i. onwards and suggesting horizontal transmission of the bacterial symbionts (Supplementary Figure S7). In contrast, host lysis occurred later (at 96 h) and only rarely in *Hartmannella*. Instead, we observed dividing *Hartmannella* trophozoites with infected nuclei throughout the course of the experiment, with both daughter cells carrying the symbiont (Figure 4). This suggests that *Nucleicultrix* in *Hartmannella* sp. is transmitted both horizontally and vertically (Figure 4).

Host cell lysis in infected *Acanthamoeba* cultures was further analyzed and quantified using a PI assay, which detects dead cells with disintegrated membranes (Figure 5a). A significant increase in PI fluorescence intensity was observed at 72 h p.i. in *A. castellanii* cultures ($P < 0.05$). This assay also demonstrated the general lack of host cell lysis in the original *Hartmannella* sp. host (Figure 5b), in which no significant increase in PI fluorescence intensities was detected. This is particularly noteworthy as the proportion of infected cells was generally higher in *Hartmannella* than in *A. castellanii* cultures (Figures 5a and b). At 24 h p.i., *Nucleicultrix* was present in 30% of all *Hartmannella* sp. cells but in only 10% of *A. castellanii* trophozoites when the same multiplicity of infection was used. Although the percentage of infected cells remained at a similar level during the course of the experiment in case of *A. castellanii*, nearly 80% of all *Hartmannella* sp. trophozoites contained

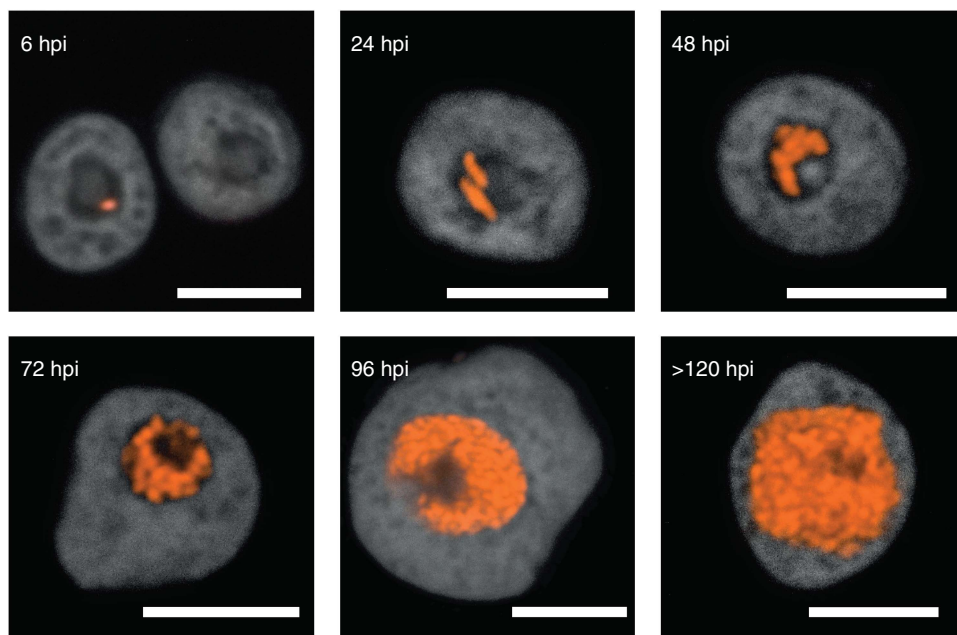


Figure 3 Infection cycle of ‘*Candidatus Nucleicultrix amoebiphila*’ in *Hartmannella* sp. FS5. Infection was monitored over a course of 144 h and visualized with FISH using probes CBR125 (orange) and EUK516-Cy5 (gray). Single bacterium starts colonizing the nuclear compartment within the first 6 h post infection; the nucleus is completely filled with bacteria after 96 h. Fully infected and highly enlarged nuclei can be observed after 120 h. The course of infection of *Acanthamoeba* sp. cells is similar (Supplementary Figure S7). Bar, 5 μ m.

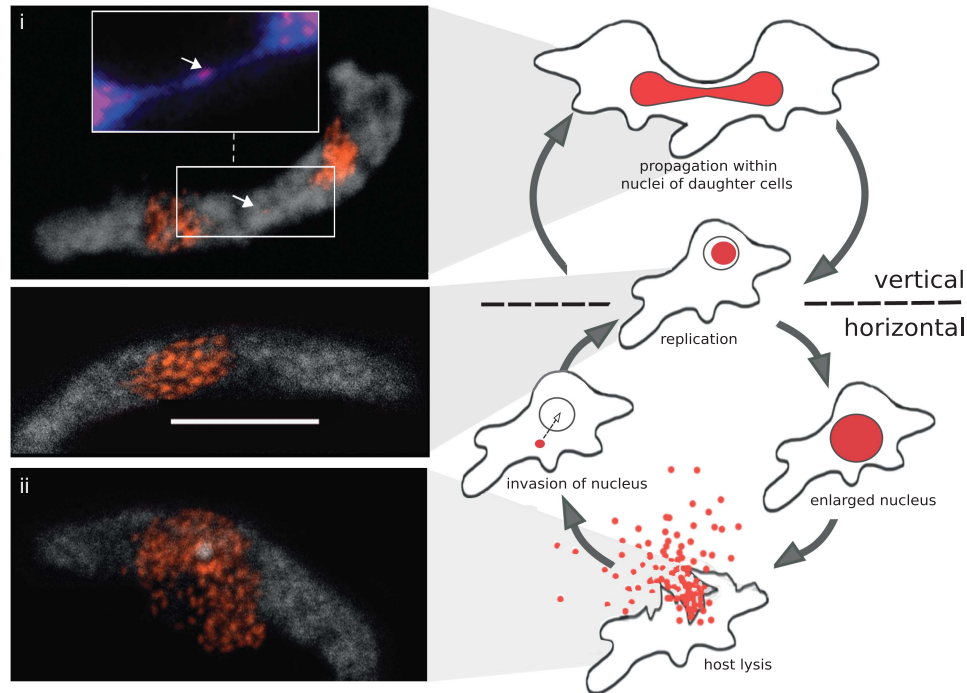


Figure 4 Modes of transmission of ‘*Candidatus Nucleicultrix amoebiphila*’. Key events during transmission of *Nucleicultrix* in *Hartmannella* sp. visualized by FISH are shown in the left panel. Bacteria and amoebae were labeled with probes CBR125 (orange) and EUK516 (gray), respectively. A dividing *Hartmannella* trophozoite is shown (i) in which the distribution of *Nucleicultrix* among both daughter cells can be seen. Note the single bacterial cell (arrows) within the amoeba nucleus undergoing segregation. Host cell lysis (ii) occurs only rarely in *Hartmannella* sp.; bar, 10 μm . A model of the life cycle and transmission of *Nucleicultrix* in amoebae is shown in the right panel. Although vertical transmission is predominant in *Hartmannella* sp., host lysis and horizontal transmission is the major route of *Nucleicultrix* in acanthamoebae.

Nucleicultrix at 144 h p.i. (Figures 5a and b). However, the presence of *Nucleicultrix* did not have a negative effect on the net growth of both *Acanthamoeba* and *Hartmannella* cultures (Figures 5c and d). The presence of *Nucleicultrix* also did not affect *Hartmannella* growth in continuous cultures, even when different incubation temperatures were used (Supplementary Figure S9).

We were able to maintain *Nucleicultrix* in continuous *Hartmannella* sp. cultures for more than 2 years so far, with 30–80% of all amoebae being infected. Moreover, infected *Acanthamoeba* cultures could be propagated for several months, yet with lower infection rates of 5–20%. *Nucleicultrix* is able to persist in *Hartmannella* sp. cysts over longer periods of starvation (Figure 1d), whereas we never observed infected *Acanthamoeba* cysts.

Environmental distribution of *Nucleicultrix*

To investigate the environmental distribution and diversity of *Nucleicultrix*, we searched 16S rRNA amplicon data sets deposited at SRA and the VAMPS (Huse *et al.*, 2010; Kodama *et al.*, 2012; <http://vamaps.mbl.edu>; Lagkouvardos *et al.*, 2013). We retrieved a total of 1356 sequences with a similarity >95% to the 16S rRNA of *Nucleicultrix*, with the majority of sequences being nearly identical to that of the symbiont ($\geq 99\%$; Supplementary Figure S10; Supplementary Table S2).

Grouping these sequences into operational taxonomic units, dereplication and subsequent phylogenetic analysis showed that *Nucleicultrix*-related sequences form two clearly distinct clades (Figure 6). The vast majority of sequences originated from freshwater samples (Supplementary Figure S10, Supplementary Table S2); however, both clades also contained sequences from a variety of environments including soil, anthropogenic habitats, humans and animals (Figure 6). Nearly identical sequences were found in different environmental samples originating from geographically distinct regions.

Discussion

This study represents the first identification and characterization of bacterial symbionts colonizing the nucleus of amoebae. So far, bacteria sharing this unusual lifestyle have been mainly found in protozoan hosts: in ciliates (Goertz, 1986; Fokin, 2004) and also in *Euglenida* (Shin *et al.*, 2003), *Trichonymphida* (Dolan *et al.*, 2004), *Cristamonadida* (d’Ambrosio *et al.*, 1999), *Spirotrichonymphida* (Hollande and Carruette-Valentin, 1971) and dinoflagellates (Leedale, 1969; Alverca *et al.*, 2002). Most of these bacteria were described only at the phenotypic level, using light- and electron microscopy to confirm the intranuclear location. The phenomenon of intranuclear symbiosis was investigated in more detail only for the *Paramecium*

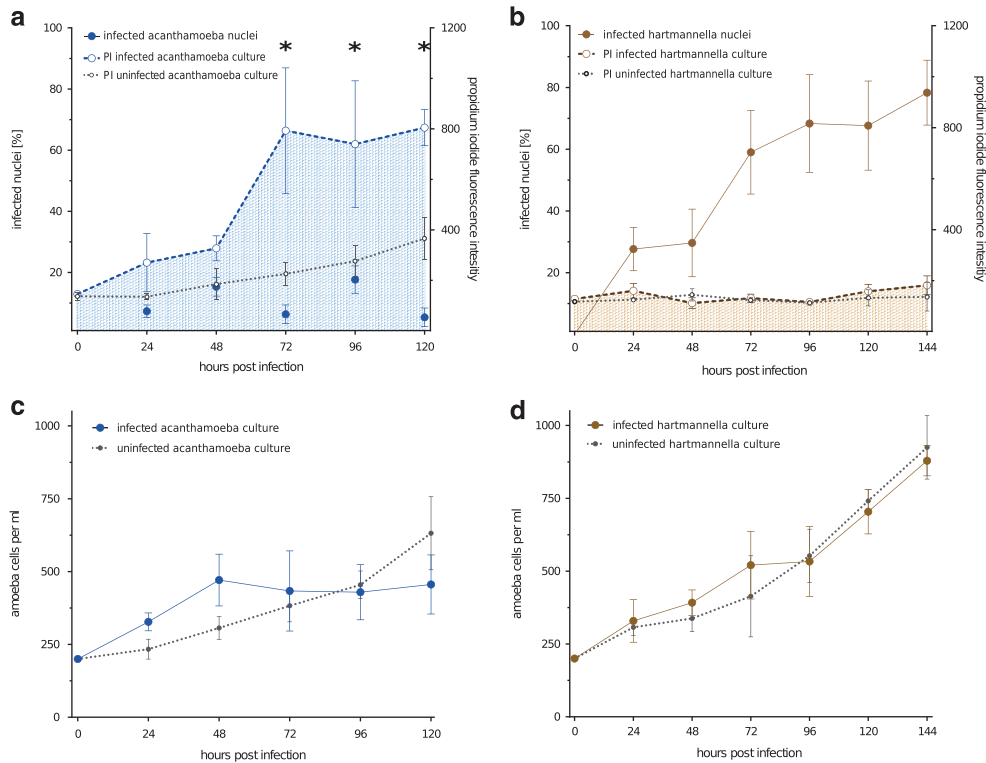


Figure 5 Fate of amoeba hosts during infection with ‘*Candidatus Nucleicultrix amoebiphila*’. (a, b) Viability of amoeba cells during infection with *Nucleicultrix*. PI fluorescence intensities indicating dead amoeba cells were determined during the course of infection. *A. castellanii* but hardly *Hartmannella* sp. cells are lysed in the presence of *Nucleicultrix*. The percentage of infected amoebae during the course of infection is shown by filled circles. (c, d) Amoeba cell numbers during the course of infection. Asterisks indicate significant differences to the control ($P < 0.05$), error bars show s.d. based on three replicate infection experiments.

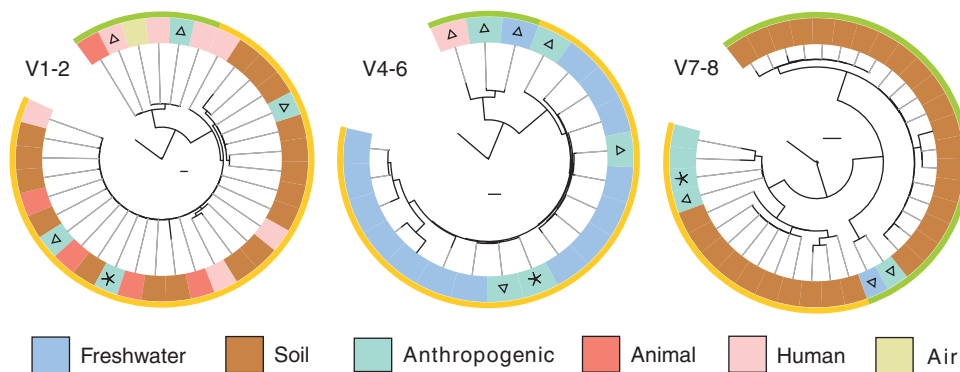


Figure 6 Diversity and environmental distribution of ‘*Candidatus Nucleicultrix amoebiphila*’. Phylogenetic trees based on amplicon sequences (>300 nucleotides length, >95% sequence similarity to the *Nucleicultrix* 16S rRNA) are shown. As 16S rRNA sequences from amplicon-based studies generally cover only short regions of the full-length gene, separate trees comprising sequences spanning variable regions V1–V2, V4–V6 and V7–V8, respectively, were calculated. Each tree depicts a subset of representative sequences retrieved from public databases. With all three data sets, *Nucleicultrix*-related sequences group into two distinct clades (marked yellow and green in the outer circle). Colors in the inner circle represent the environmental origin of the sequences; triangles indicate full-length sequences; the position of *Nucleicultrix* is labeled with an asterisk. Bars indicate an estimated evolutionary distance of 0.01.

symbionts *Caedibacter* spp. (Schrallhammer and Schweikert, 2009), *Holospira* spp. (Fujishima and Fujita, 1985; Fokin and Görtz, 2009), ‘*Candidatus Paraholospira nucleivisitans*’ (Eschbach *et al.*, 2009) and for the recently described ‘*Candidatus Endonucleobacter bathymodioli*’ in deep sea mussels (Zielinski *et al.*, 2009).

The affiliation of most intranuclear bacteria is unknown; however, *E. bathymodioli* has been identified

as a member of the *Gammaproteobacteria*, and the *Paramecium* symbionts *Holospira* and *Caedibacter* were described as *Alphaproteobacteria*, forming a monophyletic group in the order *Rickettsiales* (Zielinski *et al.*, 2009; Boscaro *et al.*, 2013). *Nucleicultrix* is only distantly related to *Holospira* and *Caedibacter* species, but it groups with the latter, as well as with the cytoplasmic *Acanthamoeba* symbiont S40 (Figure 2b) (Matsuo

et al., 2009). We noted that—in contrast to previous analysis and 16S rRNA-based phylogenetic trees—the 23S rRNA-based tree and a tree inferred from concatenated 16S and 23S rRNA genes placed this symbiont clade outside of the *Rickettsiales* and together with the *Rhodospirillales*, which comprise mainly free-living bacteria (Figure 2a). It has previously been shown that the relationships among *Alphaproteobacteria* are difficult to resolve based on rRNA gene analysis alone (Williams *et al.*, 2007). Although the exact affiliation of *Nucleicultrix* with other *Alphaproteobacteria* remains elusive at this point, future phylogenomic analysis might help to clarify this. Despite this uncertainty, the occurrence of the intranuclear lifestyle in only distantly related evolutionary lineages (*Alpha*- and *Gammaproteobacteria*, respectively) suggests that the ability to use the nucleus of eukaryotic cells as a niche has evolved multiple times independently in different bacterial groups.

Querying sequence databases containing environmental 16S rRNA gene sequences derived from amplicon-based studies provided a first glimpse at diversity and environmental distribution of *Nucleicultrix*-related bacteria. In addition to only 17 near-full-length sequences in the main public databases Genbank/EMBL/DDBJ, this analysis recovered more than 1300 partial 16S rRNA gene sequences with similarity to the 16S rRNA of *Nucleicultrix*. On the basis of this data set, *Nucleicultrix*-related bacteria belong to either of the two phylogenetically distinct clades and occur in a variety of habitats (Figure 6, Supplementary Figure S10). The presence of nearly identical sequences in samples from geographically distinct locations and in environments where protists are typically found (such as freshwater and soil) suggests that *Nucleicultrix*-related bacteria occur world-wide and are likely dispersed by protozoa.

Known bacterial symbionts targeting the nucleus of their eukaryotic host cells display a number of remarkable morphologic features different from those of *Nucleicultrix*. Both the mussel symbiont *E. bathymodioli* and the *Paramecium* symbionts *Holospira* spp. show a developmental cycle with morphologically distinct stages (Zielinski *et al.*, 2009; Fujishima and Kodama, 2012). The shape of these symbionts alters between short coccoid rods and elongated cells with a length of up to 15–20 μm . These elongated forms represent the reproductive stage of *E. bathymodioli* or the infectious form of *Holospira* species. The latter is further characterized by a so-called invasion tip, a subcellular structure involved in the invasion of the host nucleus (Kawai and Fujishima, 2000). In contrast, we found no evidence for different cell shapes or developmental forms for *Nucleicultrix* during the infection of its amoeba hosts. *Nucleicultrix* is also different from the *Paramecium* symbionts *Caedibacter caryophila* and *C. macronucleorum*, which contain characteristic ‘refractile bodies’ known as

R-bodies that mediate the killer trait of their hosts (Schmidt and Goertz, 1987; Schrällhammer and Schweikert, 2009). Taken together, this suggests that the strategies for host interaction are fundamentally different between known endonuclear bacteria.

After uptake into the host cell by phagocytosis, *Nucleicultrix* escapes the phagolysosome and is temporarily located in the cytoplasm (Supplementary Figure S6). The exact mode of entry into the nucleus remains unknown. However, during cell division, *Hartmannella* and *Acanthamoeba* undergo open mitosis (Levandowsky and Hutner, 1979; Khan, 2009), in which the nuclear envelope disintegrates. A conceivable scenario is that *Nucleicultrix* interacts with chromatin during host cell division and is enclosed in the nucleus after reassembly of the nuclear envelope. Once inside the nucleus, *Nucleicultrix* colocalizes with host chromatin and initiates replication. During the course of infection, the electron-dense heterochromatin gradually disappears and a white, unstructured electron-lucent space around the symbionts occurs in TEM images (Figure 1). This is reminiscent of nuclei of paramecia infected with *Holospira* spp. and suggests that endonuclear bacteria might feed directly on host chromatin (Goertz, 1986). However, the overall neutral effect of *Nucleicultrix* on host fitness seems to rule out an extensive degradation of host DNA. In its original *Hartmannella* host, *Nucleicultrix* does not negatively affect growth (Supplementary Figure S9); the symbionts are transmitted vertically during binary fission of their host cells and amoeba cultures can be maintained for prolonged periods at high infection rates. This situation is different in *Acanthamoeba* host cells where *Nucleicultrix* is more lytic and the percentage of infected amoeba is generally lower (Figure 5).

Although rarely used as intracellular niche, the nucleus is a potentially attractive compartment for intracellular bacteria, as it is rich in proteins, nucleic acids and nucleoside triphosphates required for DNA replication and transcription. It also represents a shelter protecting bacteria from cellular defense mechanism in the cytoplasm. A number of bacterial pathogens that reside in the cytoplasm have evolved sophisticated strategies to target bacterial effector proteins (nucleomodulins) to the nucleus (Bierne *et al.*, 2012). Residing within this compartment should facilitate manipulations of the host cell by targeting basic processes in the nucleus such as gene expression or DNA replication. The identification of bacteria living in the nucleus of amoebae, which can easily be kept in culture, represents a great opportunity to further investigate molecular mechanisms underlying intranuclear symbiosis.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Christiane Dorninger, Hanna Koch and Jan Dolinšek for providing samples. Frederik Schulz is a recipient of the DOC fellowship of the Austrian Academy of Sciences at the Division of Microbial Ecology, University of Vienna, Austria. Matthias Horn acknowledges the support from the European Research Council (ERC StG ‘EvoChlamy’).

References

- Abd H, Johansson T. (2003). Survival and growth of *Francisella tularensis* in *Acanthamoeba castellanii*. *Appl Environ Microbiol* **69**: 600–606.
- Altschul SF, Madden TL, Schäffer AA, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.
- Alverca E, Biegala I, Kennaway G. (2002). In situ identification and localization of bacteria associated with *Gyrodinium instriatum* (Gymnodiniales, Dinophyceae) by electron and confocal microscopy. *Eur J Phycol* **37**: 37–41.
- Amann R, Binder B. (1990). Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl Environ Microbiol* **56**: 1919–1925.
- Barker J, Brown M. (1994). Trojan horses of the microbial world: protozoa and the survival of bacterial pathogens in the environment. *Microbiology* **140**: 1253–1259.
- Bierne H, Cossart P, Bact I. (2012). When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14**: 622–633.
- Boscaro V, Fokin S, Schrällhammer M. (2013). Revised systematics of *Holospora*-like bacteria and characterization of “*Candidatus Gortzia infectiva*” a novel macronuclear symbiont of *Paramecium jenningsi*. *Microb Ecol* **65**: 255–267.
- Cazalet C, Rusniok C, Brüggemann H, Zidane N, Magnier A, Ma L *et al.* (2004). Evidence in the *Legionella pneumophila* genome for exploitation of host cell functions and high genome plasticity. *Nat Genet* **36**: 1165–1173.
- Casadevall A. (2008). Evolution of intracellular pathogens. *Annu Rev Microbiol* **62**: 19–33.
- Collingro A, Toenshoff ER, Taylor MW, Fritsche TR, Wagner M, Horn M. (2005). “*Candidatus* Protochlamydia amoebophila” an endosymbiont of *Acanthamoeba* spp. *Int J Syst Evol Microbiol* **55**: 1863–1866.
- d’Ambrosio U, Dolan M, Wier AM, Margulis L. (1999). *Devescovinid trichomonad* with axostyle-based rotary motor (‘*Rubberneckia*’): taxonomic assignment as *Caduceia versatilis* sp. nov. *Eur J Protistol* **35**: 327–337.
- Daims H, Brühl A, Amann R, Schleifer KH, 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. (2005). Fluorescence *in situ* hybridization for the detection of prokaryotes. *Mol Microb Ecol* **213**: 239.
- Dolan MF, Wier AM, Melnitsky H, Whiteside JH, Margulis L. (2004). Cysts and symbionts of *Staurojoenina assimilis* Kirby from *Neotermes*. *Eur J Protistol* **40**: 257–264.
- Edgar RC. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.
- Eschbach E, Pfannkuchen M, Schweikert M, Drutschmann D, Brümmer F, Fokin S *et al.* (2009). ‘*Candidatus Paraholospora nucleivisitans*’ an intracellular bacterium in *Paramecium sexaurelia* shuttles between the cytoplasm and the nucleus of its host. *Syst Appl Microbiol* **32**: 490–500.
- Fokin SI. (2004). Bacterial endocytobionts of *Ciliophora* and their interactions with the host cell. *Int Rev Cytol* **236**: 181–249.
- Fokin SI, Görtz HD. (2009). Diversity of *Holospora* bacteria in *Paramecium* and their characterization. In: Fujishima M (ed) *Endosymbionts in Paramecium*. Springer-Verlag: Heidelberg, Germany, pp 161–199.
- Friedrich N, Hagedorn M, Soldati-Favre D, Soldati T. (2012). Prison break: pathogens’ strategies to egress from host cells. *Microbiol Mol Biol Rev* **76**: 707–720.
- Fritsche TR, Gautom RK, Seyedirashti S, Bergeron DL, Lindquist TD. (1993). Occurrence of bacterial endosymbionts in *Acanthamoeba* spp isolated from corneal and environmental specimens and contact lenses. *J Clin Microbiol* **31**: 1122–1126.
- Fujishima M, Fujita M. (1985). Infection and maintenance of *Holospora obtusa* a macronucleus-specific bacterium of the ciliate *Paramecium caudatum*. *J Cell Sci* **76**: 179–187.
- Fujishima M, Kodama Y. (2012). Endosymbionts in paramecium. *Eur J Protistol* **48**: 124–137.
- Gortz H. (1986). Endonucleobiosis in ciliates. *Int Rev Cytol* **102**: 169–213.
- Greub G, Mege J, Raoult D. (2003). *Parachlamydia acanthamoeba* enters and multiplies within human macrophages and induces their apoptosis. *Infect Immun* **59**: 5979–5985.
- Hollande A, Carruette-Valentin J. (1971). Les attractophores l’induction du fuseau et la division cellulaire chez les *Hypermastigines*. *Protistologica* **7**: 5–100.
- Horn M. (2008). *Chlamydiae* as symbionts in eukaryotes. *Annu Rev Microbiol* **62**: 113–131.
- Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B *et al.* (2004). Illuminating the evolutionary history of *Chlamydiae*. *Science* **304**: 728–730.
- Horn M, Wagner M. (2004). Bacterial endosymbionts of free-living amoebae. *J Eukaryot Microbiol* **51**: 509–514.
- Hunt D, Klepac-Ceraj V. (2006). Evaluation of 23S rRNA PCR primers for use in phylogenetic studies of bacterial diversity. *Appl Environ Microbiol* **72**: 2221–2225.
- Huse SM, Welch DM, Morrison HG, Sogin ML. (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12**: 1889–1898.
- Isberg RR, O’Connor TJ, Heidtman M. (2009). The *Legionella pneumophila* replication vacuole: making a cosy niche inside host cells. *Nat Rev Microbiol* **7**: 13–24.
- Juretschko S, Timmermann G, Schleifer KH, Schmid M, Pommerening-Röser A, Koops HP *et al.* (1998). 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.
- Kawai M, Fujishima M. (2000). Invasion of the macronucleus of *Paramecium caudatum* by the bacterium *Holospora obtusa*: fates of the bacteria

- and timings of invasion steps. *Eur J Protistol* **36**: 46–52.
- Khan NA. (2009). *Acanthamoeba: biology and pathogenesis*. Causter Academic Press: Norfolk, UK.
- Kodama Y, Shumway M, Leinonen R. (2012). The Sequence read archive: explosive growth of sequencing data. *Nucleic Acids Res* **40**: D54–D56.
- Lagkouvardos I, Weinmaier T, Lauro FM, Cavicchioli R, Rattei T, Horn M. (2013). Integrating metagenomic and amplicon databases to resolve the phylogenetic and ecological diversity of the *Chlamydiae*. *ISME J* **8**: 115–125.
- Leedale G F. (1969). Observations on endonuclear bacteria in euglenoid flagellates. *Österreichische Bot Zeitschrift* **116**: 279–294.
- Letunic I, Bork P. (2007). Interactive tree of life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**: 127–128.
- Levandowsky M, Hutner SH. (1979). *Biochemistry and physiology of protozoa 1*. Academic Press: New York, USA.
- Loy A, Lehner A, Lee N. (2002). 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.
- Loy A, Maixner F, Wagner M, Horn M. (2007). probeBase—an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucleic Acids Res* **35**: 800–804.
- Ludwig W, Strunk O, Westram R, Richter L, Meier H, Buchner A *et al*. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Matsuo J, Kawaguchi K, Nakamura S, Hayashi Y, Yoshida M, Takahashi K *et al*. (2009). Survival and transfer ability of phylogenetically diverse bacterial endosymbionts in environmental *Acanthamoeba* isolates. *Environ Microbiol Rep* **2**: 524–533.
- Mba Medie F, Ben Salah I, Henrissat B, Raoult D, Drancourt M. (2011). *Mycobacterium tuberculosis* complex mycobacteria as amoeba-resistant organisms. *PLoS One* **6**: e20499.
- Molmeret M, Horn M, Wagner M. (2005). Amoebae as training grounds for intracellular bacterial pathogens. *Appl Environ Microbiol* **71**: 20–28.
- Nawrocki E, Eddy S. (2009). Structural RNA homology search and alignment using covariance models. *Electronic Theses and Dissertations*. Paper 256. Washington University.
- 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.
- Page FC. (1988). *A new key to freshwater and soil Gynnamoebae: with instructions for culture*. Freshwater Biological Association: Cumbria, UK, pp 95–96.
- Price MN, Dehal PS, Arkin AP. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* **5**: e9490.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P *et al*. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.
- Rodríguez-Zaragoza S. (1994). Ecology of free-living amoebae. *Crit Rev Microbiol* **20**: 225–241.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S *et al*. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539–542.
- Santic M, Akimana C, Asare R, Kouokam JC, Atay S, Kwaik YA. (2009). Intracellular fate of *Francisella tularensis* within arthropod-derived cell. *Environ Microbiol* **11**: 1473–1481.
- Sassera D, Beninati T, Bandi C, Bouman EP, Sacchi L, Fabbri M *et al*. (2006). “*Candidatus* Midichloria mitochondrii” an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. *Int J Syst Evol Microbiol* **56**: 2535–2540.
- Schmidt HJ, Goertz HD. (1987). *Caedibacter caryophila* sp. nov., a killer symbiont inhabiting the macronucleus of *Paramecium caudatum*. *Int J Syst Evol Microbiol* **37**: 6–9.
- Schmitz-Esser S, Toenshoff E, Haider S, Heinz E, Hoenninger VM, Wagner M *et al*. (2008). Diversity of bacterial endosymbionts of environmental *Acanthamoeba* isolates. *Appl Environ Microbiol* **74**: 5822–5831.
- Schrallhammer M, Schweikert M. (2009). The killer effect of *Paramecium* and its causative agents. In: Fujishima M (ed) *Endosymbionts in Paramecium*. Springer-Verlag: Heidelberg, Germany, pp 227–245.
- Shin W, Boo S, Fritz L. (2003). Endonuclear bacteria in *Euglena hemichromata* (*Euglenophyceae*): a proposed pathway to endonucleobiosis. *Phycologia* **42**: 198–203.
- Swanson M S, Hammer BK. (2000). Legionella pneumophila pathogenesis: a fateful journey from amoebae to macrophages. *Annu Rev Microbiol* **54**: 567–613.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood evolutionary distance, maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.
- Waterhouse AM, Procter J B, Martin DM, Clamp M, Barton GJ. (2009). Jalview version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–1191.
- Wheeler DL, Barrett T, Benson D, Bryant SH, Canese K, Chetvernin V *et al*. (2008). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* **36**: D13–D21.
- Williams KP, Sobral BW, Dickerman AW. (2007). A robust species tree for the alphaproteobacteria. *J Bacteriol* **189**: 4578–4586.
- Zielinski FU, Perenthaler A, Duperron S, Raggi L, Giere O, Borowski C *et al*. (2009). Widespread occurrence of an intranuclear bacterial parasite in vent and seep bathymodiolin mussels. *Environ Microbiol* **11**: 1150–1167.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

Supplementary Information accompanies this paper on The ISME Journal website (<http://www.nature.com/ismej>)