Single-cell microbiomics unveils distribution and patterns of microbial symbioses in the natural environment.

Vittorio Boscaro[a], Vittoria Manassero[b], Patrick J. Keeling[a], Claudia Vannini[b1]

[a] University of British Columbia, Department of Botany, Vancouver, Canada

[b] University of Pisa, Department of Biology, Pisa, Italy

[1] Corresponding author: Claudia Vannini. Università di Pisa, Dipartimento di Biologia, Unità di Zoologia-Antropologia. Via A. Volta 4/6, 56126 Pisa, Italy. Fax: (0039) 0502211393. e-mail address: claudia.vannini@unipi.it

Abstract

Protist-bacteria associations are extremely common. Among them, those involving ciliates of the genus *Euplotes* are emerging as models for symbioses between prokaryotes and eukaryotes, and a great deal of information is available from cultured representatives of this system. Even so, as for most known microbial symbioses, data on natural populations is lacking and their ecology remain largely unexplored; how well cultures represent natural diversity is untested. Here we describe a survey on natural populations of *Euplotes* based on a single-cell microbiomic approach. The results reveal an unexpected variability in symbiotic communities, with different individual hosts of the same population harboring different sets of bacterial endosymbionts. In some cases, co-occurring *Euplotes* can even have different essential symbionts, *Polynucleobacter* and "*Candidatus* Protistobacter", suggesting that replacement events could be more frequent in nature than previously hypothesized. Accessory symbionts are even more variable: some showed a strong affinity for one host species, some for a sampling site, and two ("*Candidatus* Cyrtobacter" and "*Candidatus* Anadelfobacter") displayed an unusual pattern of competitive exclusion. These data represent the first insight into prevalence and patterns of microbial associations in protists in the natural environment.

Keywords: protist-bacteria symbiosis – protist microbiome - *Bandiella* - *Devosia* – *Francisella* – protist microbiota

Declarations

Funding: This work was supported by the University of Pisa (565-60% 2018, 565-60% 2019, 565-60% 2020, PRA_2018_63) and by the Gordon and Betty Moore Foundation (https://doi.org/10.37807/GBMF9201).

Conflict of interest/Competing interests: The authors declare no conflict of interest or competing interests.

Availability of data and material: Sanger nucleotide sequences and metabarcoding raw reads have been deposited in the European Nucleotide Archive (ENA) database under accession numbers OU070010-OU070076 and project number PRJEB44318, respectively.

Code availability: Not applicable

Authors' contributions: Vittorio Boscaro and Claudia Vannini conceived and designed the project, analyzed the data, wrote the first draft of the paper. Field and lab work were performed by Vittoria Manassero and Claudia Vannini. Patrick Keeling and Claudia Vannini supervised the work. All authors contributed to the final draft.

Acknowledgements

The authors wish to thank Simone Gabrielli and Davide Stano for support in sampling and for help with graphic artworks. The authors are grateful to the Migliarino San Rossore Massaciuccoli Regional Park for giving permission for sampling, and to Francesca Logli in particular for her assistance.

Introduction

Symbioses between bacteria and unicellular eukaryotes (protists) are an extraordinary common yet understudied phenomenon [1]. Among protists, ciliates are particularly prone to establishing symbiotic associations with prokaryotes, due to several distinctive characteristics such as bacterivorous feeding behavior, large size, and a variety of intracellular compartments, which offer different microhabitats for bacterial colonization [2]. Associations between bacteria and ciliates constitute a traditional field of research in ciliatology [3], and in the last decades the use of molecular and "-omics" approaches has renewed the interest in this topic. Many new bacterial symbionts of ciliates have been described (e.g. [4-7]) and characterized from the points of view of genomics and phylogeny [6-12], life cycle [13], and relationship with the host [14-17].

Among ciliates, *Euplotes* represents one of the most extensively studied. Species in this genus are common inhabitants of most aquatic environments. Ancestrally marine, they have successfully invaded freshwater and soil habitats [18] and can be easily collected from the wild and maintained in laboratory. Therefore, *Euplotes* has been used as a model system for genetics, molecular biology, cell biology, ecology, and symbiosis [5, 19-21]. Indeed, a considerable amount of data on very different and multifaceted bacterial symbioses in *Euplotes* is available. Within the genus, a monophyletic clade ("clade B", Syberg-Olsen *et al.* 2016) depends on endosymbiotic bacteria for reproduction and survival [9, 22, 23]. These essential symbionts have been recruited many times during the evolutionary history of the hosts, depicting a complex and intriguing picture of loss and gain [8, 9]. Up to now, three bacteria have been described as essential symbionts of clade B *Euplotes*, namely *Polynucleobacter*, "*Candidatus* Protistobacter" (both *Betaproteobacteria*), and "*Candidatus* Devosia" (*Alphaproteobacteria*) [9, 16, 24]. In addition to the essential symbionts, a variety of accessory bacteria have been described within *Euplotes* species of clade B, mostly belonging to the orders *Rickettsiales* and *Holosporales* of *Alphaproteobacteria* or to *Gammaproteobacteria* [5]. Multiple accessory symbionts can coexist, with up to six stable different bacteria found in a single host strain [5]. Essential symbionts, as well as some whose relationship with the host still has to be clarified, are also harboured by *Euplotes* belonging to other phylogenetic clades [12, 14, 25].

Up to now all studies on the symbionts of *Euplotes* and other ciliates have been performed on laboratory cultures, whose representativeness of the real situation in the natural environment has not been assessed. Field research is completely lacking, and nothing is known about the ecology of these symbiotic systems, including even basic data like distribution, prevalence or co-occurrence patterns. The extreme instability of protist populations in the natural environment, coupled with the lack of suitable and reliable methods, have hindered such investigations.

Here we report the first survey of bacterial symbionts in natural populations of a ciliate, using a sensitive single-cell metabarcoding approach for simultaneous identification of both hosts and symbionts [26]. Detection and

documentation of all bacteria associated to single individuals of *Euplotes* in their natural habitat provides a reliable snapshot of the natural diversity at the level of individual cells in populations. Here, we analyze these data to (i) assess the prevalence of bacterial symbionts in natural populations of *Euplotes*; (ii) document the natural diversity of symbionts and specific patterns of association between host and symbiont species in natural populations; and (iii) identify patterns of symbiotic consortia inside the same host cell in nature.

Methods

Sampling and ciliate cell isolation. Samples were collected in autumn 2018 in two different areas along the Tuscany coast within the Migliarino San Rossore Massaciuccoli Regional Park. One site (SR2A) was located in the San Rossore estate, along a small ditch connected to the mouth of river Arno, and four other sites were located in coastal ponds next to Marina di Vecchiano (MdV; sites MdV1A, MdV1D, MdV3A, MdV3B), near the mouth of river Serchio (Fig. 1). Microhabitats in both areas were extremely variable, subjected to wide fluctuations in water level (up to complete drying during the warm season; water temperature values ranging during the year from 1° C up to more than 33° C) and salinity (due to frequent coastal storms; water salinity ranging during one year from 0‰ to 13‰). A total volume of about 45 mL (water and sediments) was collected from each site and immediately transported to the lab. After gentle mixing, a 30 mL aliquot from each sample was used for ciliate collection. *Euplotes* were detected by microscopical observation, individually washed (three steps in sterile, artificial brackish water followed by two additional, fast steps in sterile distilled water) and then stored in 70% v/v ethanol inside a 0.2 mL tube at -20°C. Different sterile glass micropipettes were employed for each ciliate cell during isolation and for each washing step. The remaining sample volume of 15 mL was fixed in 70% (v/ v) ethanol and divided into three aliquots used as controls, in order to characterize the background environmental microbial communities.

Amplification and sequencing of SSU rRNA genes of hosts and associated bacteria. A simultaneous amplification of eukaryotic 18S rRNA gene and bacterial 16S rRNA gene was carried out directly on each individually stored ciliate cell, without performing DNA extractions, as described in Rossi et al [26]. Amplicons were purified with the Eurogold Cycle-Pure Kit (Euroclone) and diluted 1:100; aliquots were then processed differently for host and bacteria characterization. For ciliate host identification, two semi-nested amplifications were performed, products were further purified and Sanger sequenced using multiple appropriate internal primers by GATC Biotech (Cologne, Germany) [26].

In parallel, the characterization of host-associated bacteria was carried out with a metabarcoding approach, starting with a nested PCR using the KAPA HiFi HotStart Ready Mix with a prokaryotic primer set for the V3–V4 regions of the SSU rRNA gene: the forward primer S-D-Bact-0341-b-S-17 (5'-CCTACGGGNGGCWGCAG–3') and the reverse primer S-D-Bact-0785-a-A-21 (5'-GACTACHVGGGTATCTAATCC-3') [27]. Illumina overhang adapter

sequences added to the primers were 5′-TCGTCGGCAGCGTC AGATGTGTATAAGAGACAG-3′ and 5’-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3′, respectively (Illumina protocol, Part # 15044223, Rev. B). Amplification cycles (n = 25) were performed with an annealing temperature of 55 °C.

In order to characterize environmental prokaryotic communities, control aliquots were centrifuged at 10,000 *g* to pellet microbial organisms with the sediment; the supernatant was then removed and total genomic DNA was extracted from 0.25 g of each pellet using the PowerSoil DNA Isolation Kit (MoBio). Extracted DNA was then processed by the two amplification steps described above for bacterial community characterization.

Prokaryotic amplicons from single host cells as well as environmental controls were barcoded, pooled, and sequenced by IGA Technology (Udine, Italy) on the Illumina MiSeq platform (2 × 300 paired-end sequencing).

Sequence analysis. Gene sequences of eukaryotic 18S rRNA were inspected with NCBI Blast [28] for putative identification of the ciliate hosts, using a species identity cutoff of 99%.

Raw reads of bacterial V3–V4 regions were analyzed using the Quantitative Insights Into Microbial Ecology version 2 (QIIME2, https://qiime2.org) software package [29]. In order to remove the lower-quality ending base calls, forward and reverse reads were truncated at base 290 and 260, respectively. Quality filtering, primer trimming, pair-end read merging and clustering of reads in ASVs were performed with DADA2 [30], using default settings: sequences with any N character were discarded; sequences were truncated at any base with a quality score of 2 or lower; the maximum expected error allowed was 2; chimeras were detected with the *de novo* method. Taxonomic classification was performed using the SILVA release 132 [31]. Following Werner et al [32], the regions of interest were extracted from SSU rRNA reference sequences (99% similarity clustered database) and used to train a Naive Bayes classifier. ASVs identified as mitochondria or chloroplasts were removed before further data processing.

Data mining for bacterial symbiont detection. First, ASVs automatically assigned to *Polynucleobacter* or “*Ca*. Protistobacter”, plus all those classified within *Rickettsiales*, *Holosporales*, *Francisellaceae*, *Devosiaceae*, and *Verrucomicrobia*, which collectively cover all known symbionts of *Euplotes*, were extracted. ASVs assigned to the larger groups (with the exception of *Verrucomicrobia*, none of which was found) were added to previously curated alignments of reference full-length SSU rRNA gene sequences made with MAFFT [33]. Phylogenetic trees were inferred with IQ-TREE [34] using the –m TEST option to select the best-fitting substitution model, and manually inspected to cluster ASVs into species-like assemblages, which were then reclassified (Fig. 2 and Online Resource 1). Automated assignments made by the classifier were at this point disregarded or corrected. The relative abundances of named and unnamed taxa obtained this way were then assessed in host-derived libraries and environmental controls. Representative ASVs for putative bacterial taxa with provisional names mentioned here have been deposited in the European Nucleotide Archive (ENA) database under accession numbers OU452359-OU452364.

Results

Ciliate host identification. A total of 62 ciliate cells collected from ephemeral brackish water environments (Fig. 1A) were successfully processed both for ciliate host identification and for characterization of associated prokaryotes (up to 20 ciliate cells per morphospecies in the same site). For each ciliate cell, an almost complete 18S rRNA gene sequence was obtained (1,311-1,886 bp, with one 769 bp-long outlier, Online Resource 2). The two most frequently retrieved *Euplotes* species were *Euplotes platystoma* (39 cells) and *Euplotes woodruffi* (19 cells), both belonging to clade B and known to host essential bacterial symbionts. Data analyses and discussion are therefore focused on these two species.

Distribution of bacterial symbionts in* Euplotes *hosts. Ciliate-associated libraries averaged $1.35\ 10^5 \pm 3.57\ 10^4$ (SD) raw read pairs, $6.75\ 10^4 \pm 2.07\ 10^4$ (SD) merged sequences after quality control, and 102 ± 59.3 (SD) unique amplicon sequence variants (ASVs) (Online Resource 3). Instead of taking into account the entire prokaryotic community associated with the hosts (Online Resource 4) and trying to identify all potential symbionts, we focused only on bacterial groups already known to include symbionts of *Euplotes*, and manually refined promising species-like taxa (Fig. 2, Online Resource 1). Within the targeted bacterial groups, we identified 11 endosymbiotic species, each with a much higher relative abundance in host-derived libraries (1% or more of the sequences in at least one *Euplotes* library) than in environmental controls (at least 100 times more); they were in fact mostly absent from the latter (Online Resource 5), suggesting that signals from potential free-living forms and even symbionts are negligible compared to the overall background community. These 11 species, 5 of which were previously known as *Euplotes* symbionts, are here subdivided in three categories for convenience: (i) known essential symbionts, (ii) common accessory symbionts found in multiple sites, and (iii) uncommon accessory symbionts found in only one site or less than five host cells.

Essential symbionts. The two betaproteobacterial essential symbionts of clade B *Euplotes* species, namely *Polynucleobacter* and "*Ca.* Protistobacter", were the most common and abundant symbionts observed in our survey: every *E. platystoma* and *E. woodruffi* cell harbored one or the other. *Polynucleobacter* was the most prevalent, being detected in 52 out of 58 cells, while the remaining 6 cells, all from one population of *E. woodruffi*, contained "*Ca.* Protistobacter" (Fig. 1B). Their relative abundances were generally high, with *Polynucleobacter* averaging 56.9% of the sequences in host cells and "*Ca.* Protistobacter" averaging 14.9%. "*Ca.* Protistobacter" was never detected in controls, while *Polynucleobacter* was present at very low abundances (average: 0.043%), possibly reflecting the presence of free-living strains in the environment [35]. Interestingly, both betaproteobacteria were found in the *E. woodruffi* population from site SR2A (Fig. 1B), although each individual cell harbored only one of the two symbionts (with the possible exception of four cells which displayed an additional low signal from the other essential symbiont: <0.5% of the total sequences, or <50 times the abundance of the predominant one).

Common accessory symbionts. All other potential symbionts belonged to two alphaproteobacterial orders of intracellular bacteria, *Rickettsiales* and *Holosporales* (Fig. 1B). Three "*Candidatus* Midichloriaceae" (*Rickettsiales*) species previously described were the most common in our survey. "*Candidatus* Anadelfobacter veles" was only found in *E. platystoma* (in 11 out of 40 cells, average abundance: 1.46%), the host species it was originally described from (at the time identified as *Euplotes harpa* [36]. "*Candidatus* Cyrtobacter comes", also originally characterized in *E. platystoma* [36], was also more prevalent in that species (detected in 17 out of 39 cells, average abundance: 6.12%; versus in 4 out of 19 cells of *E. woodruffi*, with a lower 0.45% average abundance). "*Ca.* Anadelfobecter veles" and "*Ca.* Cyrtobacter comes" were never found in the same host cell (Fig. 1B,C). "*Ca.* Bandiella numerosa", originally described in *E. woodruffi* [5], was detected in every cell of that species surveyed here (average abundance: 13.53%), and never in *E. platystoma*. None of these midichloriaceae were ever detected in environmental controls.

Uncommon accessory symbionts. Six more species-like taxa of *Rickettsiales* and *Holosporales* were found more sporadically in *Euplotes* (and never in controls). They could not be ascribed to known bacterial species and were given provisional names: "JLMS", "LLMS", and "UMS1" belong to "*Ca.* Midichloriaceae", "URS3" to *Rickettsiaceae* (*Rickettsiales*), "HLHS" and "UHS2" to *Holosporaceae* (*Holosporales*). These taxa were generally present only in a few host cells, sometimes at low abundance (Fig. 1B), but LLMS and UHS2 were prevalent in *Euplotes* cells of both species in site SR2A (detected in all *E. platystoma* cells; 7 out of 12 *E. woodruffi* cells had both, 2 only had UHS2), and were not found elsewhere (Fig. 1B,C).

Distribution of bacteria from "opportunistic" genera. Other *bona fide* non-alphaproteobacterial symbionts of *Euplotes* such as *Nebulobacter* [37] and *Pinguicoccus* [12] were not found at all in the surveyed populations. However, two bacterial genera sometimes associated to *Euplotes* deserve mention despite (or rather, due to) not fitting our criteria for symbiont detection: *Francisella* and *Devosia*.

By far the most abundant *Francisella* species in our survey was *Francisella philomiragia*, detected here in low abundance from a few cells (7 out of 58 *Euplotes* divided between both species, average abundance: 1.91%) and in many environmental controls (average abundance: 0.062%) (Fig. 1B).

Two *Devosia* species were described as symbionts of marine and freshwater *Euplotes* [14, 16], and form a phylogenetic clade putatively considered *Euplotes*-specific (Online Resource 1). We found sequences belonging to this clade associated both to *Euplotes* cells (9, belonging to both species, average abundance: 0.235%) and environmental controls (average abundance: 0.041%) (Fig. 1B). *Devosia* sequences not belonging to this clade presented a similar profile, although they were more abundant in controls (14 cells in both host species; average abundance: 0.230% in hosts, 0.620% in controls). Whenever *Devosia* were detected associated to *Euplotes*, their abundance was considerably lower than that of the essential betaproteobacterium (from approximately 2 to 700 times so).

Discussion

Suitability of microbiomic methods on unicellular eukaryotes. Interpreting microbiomics data, especially those based on metabarcoding and relying on low DNA input, is often challenging. The molecular techniques employed here were previously tested on ciliates [26, 38, 39], but on much lower numbers of cells and focusing on whole microbial communities' composition instead of target symbiotic bacterial species. The results were compatible with the existing knowledge, but details were hard to pin down due to the huge diversity within observed microbial communities and the high potential for procedural artifacts. The analysis of this survey's data was designed to avoid two main pitfalls: first, by using a well-known host model with partially predictable outcomes (the presence of essential symbionts) we added a strong layer of control on top of routine environmental library collection; second, by focusing on symbiotic bacterial groups we could largely reduce the problem of differentiating between "symbionts", "food", and "loosely host-associated bacteria", admittedly sacrificing the possibility to detect new symbiotic lineages in *Euplotes*.

The unfailing detection of predicted essential symbionts, the recovering of several previously known species of *Euplotes* symbionts, and the absence of signal from most of them in environmental controls, are all consistent with the conclusions of this and previous attempts to describe bacterial communities within unicellular eukaryotes being accurate presentations of the communities associated with individual cells.

***Diversity of bacterial symbiont communities in natural populations of* Euplotes.** The most striking observation from our data is that members of the same natural host population can harbor different communities of bacterial symbionts, so that symbiont composition could be viewed as an intrapopulation polymorphism. Variations in bacterial symbiont frequencies among different individuals of the same population is well documented in insects, for which field campaigns have provided reliable data across both spatial and temporal scales [40, 41], but virtually unknown in protists. Our survey shows for the first time that laboratory strains, either descending from a single isolated cell or maintained long enough as to amount to the same thing, are not representative of the natural population from which they were derived, especially when it comes to "accessory" symbionts. This calls into question what conclusions can be drawn from distribution patterns based only on laboratory strains, for example concerning biogeography and prevalence, topics that in protist systems are plagued by a scarcity of data to begin with.

Our data also show that the diversity of accessory symbionts in *Euplotes*, especially *Rickettsiales* and *Holosporales*, has not been exhaustively characterized yet. At the same time, it is noteworthy that the three most common accessory symbionts detected here had been previously described, in the same host species. Even though we cannot claim yet that our knowledge on the diversity of clade B *Euplotes* symbionts is comprehensive, we can probably

speculate that such an understanding is within reasonable reach and that several of the most ecologically relevant symbionts have been characterized.

Concerning *Rickettsiales*, the common presence in *Euplotes* of several bacterial symbionts belonging to two of the three families of the order ("*Ca.* Midichloriaceae" and *Rickettsiaceae*) is consistent with previous results [5, 36]. Members of the third family, *Anaplasmataceae*, were not detected here (Online Resource 1), and are indeed conspicuously absent from symbiont screenings in all protists [1]. All three *Rickettsiales* families were originally described as parasites of terrestrial arthropods [42, 43], but *Anaplasmataceae* seems to be unable to colonize either unicellular eukaryotes, or aquatic environments. Considering the phylogenetic tree of the order (e.g., [11]), this can be assumed to be a derived, rather than ancestral, feature.

***Patterns of symbiont distribution in* Euplotes.** We know from previous studies on the evolutionary history of *Polynucleobacter* and "*Ca.* Protistobacter" that at least the former can replace the latter (as well as different strains of its own species) "often" over evolutionary times [8, 9]. *Polynucleobacter* and "*Ca.* Protistobacter" have never been found inside the same cytoplasm despite theoretical expectations that such co-occurrence should be observable in a transitional step [9]. Libraries from a few *Euplotes* cells collected here did include reads from both, which would be consistent with the hypothetical presence of two essential symbionts in very different amount in some host cells. Nevertheless, relative abundances for the less dominant symbiont were so low that they might be also explained by tiny cross-contaminations coupled with deep sequencing. On the other hand, the presence of both essential symbionts inside different individuals of the same *E. woodruffi* population (site SR2A) is strongly supported. This could only be observed by looking at individual host cells or by analysing large numbers of clonal cultures originated from the same population, neither of which is commonly done. How common the replacement of *Polynucleobacter* by "*Ca.* Protistobacter" is in absolute terms is unknown, but their coexistence in the same host population may be indicative of an ongoing takeover. We cannot however rule out the possibility that *Polynucleobacter*- and "*Ca.* Protistobacter"-carrying *Euplotes* belong to different strains with undistinguishable 18S rRNA gene sequences. Should more studies like this find similar situations in multiple target populations, it would suggest that replacements of essential symbionts in *Euplotes* happen over a much shorter timespan than expected (i.e. years and decades, not millennia or millions of years).

Some of the accessory symbionts also seem to be specific to, or at least show a very strong affinity for, one host species. "*Ca.* Bandiella numerosa" was detected in every *E. woodruffi* cell here, and congeneric bacteria were found in most laboratory strains of the same host species [5], questioning if this symbiont is indeed "accessory" or if it might play a more important role. Notably, however, while "*Ca.* Bandiella numerosa" may be exclusively present in *E. woodruffi*, extremely close bacterial relatives were found in hosts as different as marine corals [44] and placozoans [45].

Other symbionts did not show a preference between *Euplotes* species, and were instead tied to a specific location. Midichloriaceae and holosporaceae as a whole are found in a variety of unrelated eukaryotes and their phylogeny does not match that of their hosts, so it is reasonable to assume that those with a broader host range, like LLMS and UHS2, are the rule rather than the exception.

Another clear pattern, here made more striking by the fact that it was shown both among and within populations, is the apparent competitive exclusion between "*Ca.* Anadelfobacter veles" and "*Ca.* Cyrtobacter comes". Both were originally described from *E. platystoma*, but from different strains [36], and in single cell sampling these midichloriaceae never seem to share the same cytoplasm, despite being common symbionts. This suggests some strong selection against their co-occurrence. Interaction dynamics between accessory bacterial symbionts of eukaryotes are a largely unexplored field. Most of the studies on this topic have been performed on accessory symbionts of aphids, for some of which competitive interactions have been shown, leading to drops in abundance, lowering of essential symbiont density, and weakening of functions useful to the host [46, 47]. Negative correlations between two species of accessory bacterial symbionts have been reported both in aphids and in the chestnut weevil, but the reasons behind this pattern remain to be clarified [40, 48]. All the previous findings in insects agree that competition is driven by many different factors, including benefits and costs tradeoffs, way of transmission, environmental pressure, number of symbionts, and genotypes of both bacteria and hosts [41, 46, 47]. We do not know what the competition between "*Ca.* Anadelfobacter veles" and "*Ca.* Cyrtobacter comes" stems from, but their occasional absence from *E. platystoma* cells and strains makes it unlikely to be host-driven, and suggests in turn that it might be actively triggered by one or both of the bacteria.

Opportunistic symbionts. While the symbiotic status of the aforementioned bacteria is quite certain, either because of their known effect on the host, their affiliation to exclusively intracellular lineages, and/or their absence from our environmental controls, two taxa previously characterized as symbionts are probably best described as opportunistic inhabitants of *Euplotes* cytoplasm. This is not surprising in the case of *Francisella*, because the entire genus is generally considered to be opportunistic and facultatively intracellular [49]. A handful of *Francisella* species have been recovered from marine *Euplotes* [4, 25], usually without any reported effect on the host, and *Francisella philomiragia*, found here in brackish *Euplotes*, could be added to that list. At the same time its presence in the environmental background community suggests that its specificity as a symbiont is at best tenuous.

The situation in *Devosia* is more complex. The genus is large and diverse, and includes free-living as well as host-associated species [50]. However, the few well-characterized *Devosia* lineages in *Euplotes* did have a strong effect on their host, taking the role usually performed by *Polynucleobacter* in one *E. platystoma* strain [16], and being equally essential in the unrelated, marine *Euplotes magnicirratus* [14]. The phylogenetic relationship between *Devosia* species

found in *Euplotes* suggested the existence of an *Euplotes*-specific clade [16]. Here, however, *Devosia* displayed a distribution pattern not dissimilar from the opportunistic *Francisella*, both within the putative *Euplotes*-specific clade and in the rest of the genus *Devosia*. The topic needs to be further explored, but it is possible that *Devosia* is also, generally speaking, opportunistic, and that previously reported cases (at least the one in *E. platystoma*) represent rare instances of an opportunist accidentally replacing another symbiont, much as free-living *Polynucleobacter* strains often do to other betaproteobacteria in clade B *Euplotes* species.

Reference list

1. Husnik F, Tashyreva D, Boscaro V, George EE, Lukeš J, Keeling PJ (2021) Bacterial and archaeal symbioses with protists: functional and evolutionary comparisons with animal models. Curr Biol 31:862-877
2. Görtz HD (2006) Symbiotic associations between ciliates and prokaryotes. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds) The Prokaryotes. Springer, New York, pp 364-402
3. Hafkine WM (1890) Maladies infectieuses des paramecies. Ann Inst Pasteur Paris 4:148-162
4. Schrallhammer M, Schweikert M, Vallesi A, Verni F, Petroni G (2011) Detection of a novel subspecies of *Francisella noatunensis* as endosymbiont of the ciliate *Euplotes raikovi*. Microb Ecol 61:455-464
5. Boscaro V, Husnik F, Vannini C, Keeling PJ (2019) Symbionts of the ciliate *Euplotes*: diversity, patterns and potential as models for bacteria–eukaryote endosymbioses. Proc R Soc B 286:20190693
6. Graf JS, Schorn S, Kitzinger K et al (2021) Anaerobic endosymbiont generates energy for ciliate host by denitrification. Nature 591:445-450
7. Muñoz-Gómez SA, Kreutz M, Hess S (2021) A microbial eukaryote with a unique combination of purple bacteria and green algae as endosymbionts. Sci Adv 7:eabg4102
8. Vannini C, Ferrantini F, Ristori A, Verni F, Petroni G (2012) Betaproteobacterial symbionts of the ciliate *Euplotes*: origin and tangled evolutionary path of an obligate microbial association. Environ Microbiol 14:2553-2563
9. Boscaro V, Kolisko M, Felletti M, Vannini C, Lynn DH, Keeling PJ (2017) Parallel genome reduction in symbionts descended from closely related free-living bacteria. Nat Ecol Evol 1:1160-1167
10. Seah BKB, Schwaha T, Volland JM, Huettel B, Dubilier N, Gruber-Vodicka HR (2017) Specificity in diversity: single origin of a widespread ciliate-bacteria symbiosis. Proc R Soc B 284:20170764
11. Castelli M, Sabaneyeva E, Lanzoni O et al (2019) *Deianiraea*, an extracellular bacterium associated with the ciliate *Paramecium*, suggests an alternative scenario for the evolution of *Rickettsiales*. ISME J 13:2280-2294

12. Serra V, Gammuto L, Nitla V et al (2020) Morphology, ultrastructure, genomics, and phylogeny of *Euplotes vanleeuwenhoeki* sp. nov. and its ultra-reduced endosymbiont "*Candidatus* Pinguicoccus supinus" sp. nov. Sci Rep 10:2031
13. Potekhin A, Schweikert M, Nekrasova I et al (2018) Complex life cycle, broad host range and adaptation strategy of the intranuclear *Paramecium* symbiont *Preeria caryophila* comb. nov. FEMS Microbiol Ecol 94:fiy076
14. Vannini C, Schena A, Verni F, Rosati G (2004) *Euplotes magnicirratus* (Ciliophora, Hypotrichia) depends on its bacterial symbionts "*Candidatus* Devosia euplotis" for a successful digestive process. Aquat Microb Ecol 36:19-28
15. Beinart RA, Beaudoin DJ, Bernhard JM, Edgcomb VP (2018) Insights into the metabolic functioning of a multipartner ciliate symbiosis from oxygen-depleted sediments. Mol Ecol 27:1794-1807
16. Boscaro V, Fokin SI, Petroni G, Verni F, Keeling PJ, Vannini C (2018) Symbiont replacement between bacteria of different classes reveals additional layers of complexity in the evolution of symbiosis in the ciliate *Euplotes*. Protist 169:43-52
17. Volland JM, Schintlmeister A, Zambalos H et al (2018) NanoSIMS and tissue autoradiography reveal symbiont carbon fixation and organic carbon transfer to giant ciliate host. ISME J 12:714-727
18. Syberg-Olsen MJ, Irwin NAT, Vannini C, Erra F, Di Giuseppe G, Boscaro V, Keeling PJ (2016) Biogeography and character evolution of the ciliate genus *Euplotes* (Spirotrichea, Euplotia), with description of *Euplotes curdsi* sp. nov. PLoS One 11:e0165442
19. Di Giuseppe G, Erra F, Dini F et al (2011) Antarctic and Arctic populations of the ciliate *Euplotes nobilii* show common pheromone-mediated cell-cell signaling and cross-mating. Proc Natl Acad Sci USA 108:3181-3186
20. Bracht JR, Fang WW, Goldman AD, Dolzhenko E, Stein EM, Landweber LF (2013) Genomes on the edge: programmed genome instability in ciliates. Cell 152:406-416
21. Faktorová, D, Nisbet RER, Fernández Robledo JA et al (2020) Genetic tool development in marine protists: emerging model organisms for experimental cell biology. Nat Methods 17:481-494
22. Heckmann K, Ten Hagen R, Görtz HD (1983) Freshwater *Euplotes* species with 9 type I cirrus pattern depend upon endosymbionts. J Protozool 30:284-289
23. Vannini C, Sigona C, Hahn MWH, Petroni G, Fujishima M (2017) High degree of specificity in the association between symbiotic betaproteobacteria and the host *Euplotes*. Eur J Protistol 59:124-132
24. Vannini C, Ferrantini F, Verni F, Petroni G (2013) A new obligate bacterial symbiont colonizing the ciliate *Euplotes* in brackish and freshwater: "*Candidatus* Protistobacter heckmanni". Aquat Microb Ecol 70:233-243

25. Vallesi A, Sjödin A, Petrelli D et al (2019) A new species of the γ-proteobacterium *Francisella*, *F. adeliensis* sp. nov., endocytobiont in an antarctic marine ciliate and potential evolutionary forerunner of pathogenic species. Microb Ecol 77:587-596
26. Rossi A, Bellone A, Fokin SI, Boscaro V, Vannini C (2019) Detecting associations between ciliated protists and prokaryotes with culture-independent single-cell microbiomics: a proof-of-concept study. Microb Ecol 78:232-242
27. Herlemann DPR, Labrenz M, Juergens K, Bertilsson S, Waniek JJ, Anderrson AF (2011) Transition in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J 5:1571-1579
28. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402
29. Bolyen E, Rideout JR, Dillon MR et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852-857
30. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13:581-583
31. Glöckner FO, Yilmaz P, Quast C et al (2017) 25 years of serving the community with ribosomal RNA gene reference databases and tools. J Biotechnol 261:169-176
32. Werner JJ, Koren O, Hugenholtz P et al (2012) Impact of training sets on classification of high-throughput bacterial 16S rRNA gene surveys. ISME J 6:94-103
33. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772-780
34. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268-274
35. Hahn MW, Jezberová J, Koll U, Saueressig-Beck T, Schmidt J (2016) Complete ecological isolation and cryptic diversity in *Polynucleobacter* bacteria not resolved by 16S rRNA gene sequences. ISME J 10:1642-1655
36. Vannini C, Ferrantini F, Schleifer KH, Ludwig W, Verni F, Petroni G (2010) "*Candidatus* Anadelfobacter veles" and "*Candidatus* Cyrtobacter comes", two new rickettsiales species hosted by the protist ciliate *Euplotes harpa* (Ciliophora, Spirotrichea). Appl Environ Microbiol 76:4047-4054
37. Boscaro V, Vannini C, Fokin SI, Verni F, Petroni G (2012) Characterization of "*Candidatus* Nebulobacter yamunensis" from the cytoplasm of *Euplotes aediculatus* (Ciliophora, Spirotrichea) and emended description of the family *Francisellaceae*. Syst Appl Microbiol 35:432-440

38. Lanzoni O, Plotnikov AO, Khlopko Y, Munz G, Petroni G, Potekhin A (2019) The core microbiome of sessile ciliate *Stentor coeruleus* is not shaped by the environment. Sci Rep 9:11356
39. Plotnikov AO, Balkin AS, Gogoleva NE, Lanzoni O, Khlopko YA, Cherkasov SV, Potekhin AA (2019) High-throughput sequencing of the 16S rRNA gene as a survey to analyze the microbiomes of free-living ciliates *Paramecium*. Microb Ecol 78:286-298
40. Toju H, Fukatsu T (2011) Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: relevance of local climate and host plants. Mol Ecol 20:853-868
41. Smith AH, Łukasik P, O'Connor MP *et al.* (2015) Patterns, causes and consequences of defensive microbiome dynamics across multiple scales. Mol Ecol 24:1135-1149
42. Dumler JS, Barbet AF, Bekker CP et al (2001) Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int J Syst Evol Microbiol 51:2145-2165
43. Sassera D, Beninati T, Bandi C, Bouman EAP, Sacchi L, Fabbi M, Lo N (2006) "*Candidatus* Midichloria mitochondrii", an endosymbiont of the tick *Ixodes ricinus* with a unique intramitochondrial lifestyle. Int J System Evol Microbiol 56:2535-2540
44. Klinges JG, Rosales SM, McMinds R et al (2019) Phylogenetic, genomic, and biogeographic characterization of a novel and ubiquitous marine invertebrate-associated *Rickettsiales* parasite, *Candidatus* Aquarickettsia rohweri, gen. nov., sp. nov. ISME J 13:2938-2953
45. Gruber-Vodicka HR, Leisch N, Kleiner M et al (2019) Two intracellular and cell type-specific bacterial symbionts in the placozoan *Trichoplax* H2. Nat Microbiol 4:1465-1474
46. Leclair M, Polin S, Jousseaume T et al (2017) Consequences of coinfection with protective symbionts on the host phenotype and symbiont titres in the pea aphid system. Insect Sci 24:798-808
47. Weldon SR, Russell JA, Oliver KM (2020) More is not always better: coinfections with defensive symbionts generate highly variable outcomes. Appl Environ Microbiol 86:e02537-19
48. Rock DI, Smith AH, Joffe J et al (2018) Context-dependent vertical transmission shapes strong endosymbiont community structure in the pea aphid, *Acyrthosiphon pisum*. Mol Ecol 27:2039-2056
49. Sjödin A, Svensson K, Ohrman C et al (2012) Genome characterisation of the genus *Francisella* reveals insight into similar evolutionary paths in pathogens of mammals and fish. BMC Genomics 13:268

50. Talwar C, Nagar S, Kumar R, Scaria J, Lal R, Negi RK (2020) Defining the environmental adaptations of genus *Devosia*: insights into its expansive short peptide transport system and positively selected genes. Sci Rep 10:1151

Figure Captions

Fig. 1 Overview of sampling and summary of symbiont distribution. A, localization of the two investigated areas, Marina di Vecchiano (MdV, in blue), and San Rossore (SR, in brown). B, presence of targeted bacterial taxa in processed *Euplotes* single cells, grouped by site and host species. Hatched squares represent relative abundances below 1% for essential, common, and uncommon symbionts. The presence of opportunistic bacteria is tracked both in *Euplotes* specimens and environmental controls, regardless of relative abundance. C, Synopsis of symbiont patterns arranged by host species. Individual cells are represented by wedges, so that co-occurring symbionts are found along the same radius. From the inside out, circles depict essential symbionts, "*Ca.* Anadelfobacter veles", "*Ca.* Cyrtobacter comes", "*Ca.* Bandiella numerosa", and the bacteria provisionally named LLMS and UHS2

Fig. 2 Maximum Likelihood tree of family "*Candidatus* Midichloriaceae", used to cluster ASVs into species-like assemblages and refine their classification. Reference sequences are in black and associated with accession numbers. Sequences in smaller font represented by alphanumerical strings represent ASVs obtained in this work (Online Resource 4). On the right, underlined, manually identified species-like assemblages are shown, together with their final identification. Short alphanumerical codes were used to provisionally name undescribed species. Coloured taxa were considered putative symbionts and are discussed in the text (see also Fig. 1), taxa in grey and marked by an asterisk were not. The bar stands for an estimated sequence divergence of 0.1