Ancient symbiosis, recent symbionts: all essential endobacteria of the ciliate *Euplotes* are cyclically replaced

Vittorio Boscaro[a*], Mitchell J. Syberg-Olsen[a], Nicholas A. T. Irwin[a,b], Emma E. George[a], Claudia Vannini[c], Filip Husnik[a,d], Patrick J. Keeling[a]

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

[b] University of Oxford, Merton College, United Kingdom

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

[d] Okinawa Institute of Science and Technology, Japan

*Corresponding author: Vittorio Boscaro. Email address: vittorio.boscaro@botany.ubc.ca

Keywords: Endosymbiosis; *Polynucleobacter*; mutualism; coevolution

The authors declare no competing interests.

Symbiotic relationships are ubiquitous, ecologically important, and widely variable in the degree to which the partners are bound to each other[1]. Obligate intracellular endosymbionts represent one extreme, having evolved a strong co-dependence with their host which is often interpreted as mutualistic. The mutually obligate symbiosis between the betaproteobacterium *Polynucleobacter* and the clade B subgroup of the ciliate *Euplotes* challenges this view[2,3]: while freshwater *Euplotes* long ago became dependent on endosymbionts, the extant relationships between individual hosts and *Polynucleobacter* arose in parallel many times from different but closely related free-living bacteria[3]. The host benefits from the relationship, but each newly established symbiont is driven to extinction in a cycle of establishment-degeneration-replacement. To a lower extent, similar replacement events have been observed in insects[4–6], by far the most investigated models for bacteria-eukaryote symbioses[7]. In insect systems, however, an ancient symbiont coevolving with the host usually still exists, corroborating the idea that long-term mutualism remains the rule. Could this be the case for *Euplotes* too?

A small percentage of *Euplotes* populations indeed lack *Polynucleobacter* and harbour instead one of two other essential symbionts[8]: "*Candidatus* Protistobacter heckmanni" and "*Candidatus* Devosia symbiotica". No free-living *Protistobacter* are known, whereas *Devosia* comprises both free-living and symbiotic representatives, including an essential symbiont of the marine *Euplotes magnicirratus*[8] ("*Candidatus* Devosia euplotis"). The exclusively symbiotic *Protistobacter* could then be the ancestral symbiont replaced in most hosts by *Polynucleobacter*; alternatively, an even older symbiosis with *Devosia* predated the split between freshwater and marine *Euplotes*. To find support for either scenario, we have characterised the genomes of multiple symbiotic *Protistobacter* and *Devosia* strains (accession numbers: SAMN25125324-7).

One apparently universal signature of obligate endosymbionts is genome erosion[7,9]. Ancient symbionts possess tiny, compact genomes, while recently established symbionts have larger genomes that are typically rich in pseudogenes and repetitive elements (Fig. 1A). Two genomic sequence drafts of *Protistobacter* and two of *Devosia* (from both freshwater and marine *Euplotes*) investigated here are large and enriched in pseudogenes and repetitive elements (Fig. 1B), even more so than those of recently-established endosymbiotic *Polynucleobacter* strains. Moreover, different strains of *Protistobacter* and *Devosia* vary considerably in the number of mobile elements in their genomes, which argues against a long, shared evolutionary history as stable symbionts. Overall, these bacteria resemble neither free-living organisms (due to their many non-functional genes) nor ancient and streamlined symbionts with small, gene-rich genomes, but instead fit the prediction for recently-established endosymbionts undergoing genome erosion.

Ancient obligate symbionts coevolved with and share the phylogenetic history of their hosts, while independently established symbionts do not (Fig. 1A). With only two available strains, the phylogeny of symbiotic *Devosia* cannot be assessed, but we could compare the phylogenetic relationships among multiple strains of *Protistobacter* (two with complete genomes and two with genes extracted from metagenomic contigs) with that of *Euplotes* (Fig. 1C). The branching order of the bacteria conflicts with

that of the ciliate hosts, regardless of the phylogenetic method or model employed. AU tests performed on both trees rejected alternative topologies constrained to match those of the symbiotic partners (*Euplotes*: *p*-value < 10^{-8}; bacteria: *p*-value = 0.0037). Extant *Protistobacter*, like *Polynucleobacter*[3], did not co-differentiate with their hosts, and are the descendants of independently established symbioses.

Following widespread expectations about obligate symbioses, it was natural to seek descendants of an ancestral mutualist in *Euplotes*, subsequently replaced in many instances by *Polynucleobacter*. *Protistobacter* and *Devosia* were possible candidates, but they all display the hallmarks of recent establishment and genomic erosion, which strongly suggests that, just like *Polynucleobacter*, these other "essential" symbionts are continuously replacing each other over relatively short evolutionary time spans, even within the same host species (Fig. 1C). New strains undergo genomic decay due to relaxed selection and drift until they are themselves replaced and driven to extinction, as it must have happened to any "original" symbiont, which we can now assume left no descendent. As a result, this mutually obligate symbiosis is not – and never was – a mutualism. *Euplotes* require endocellular bacteria, but any selective pressure on maintaining the genome of the symbionts functional over long periods is weakened by the continuous availability of new potential symbionts, probably entering the host by escaping the digestive vacuole.

What *Euplotes* requires its symbionts for remains unknown, but it must be something that all three bacteria can provide. Nutritional supplementation, where symbionts synthesize essential nutrients missing from the host diet, is otherwise common but seems unlikely in an omnivorous predator. It is possible that the symbiosis required no gain-of-function. If, for example, a ciliate host harboring non-essential cytoplasmic bacteria would lose a universal metabolic pathway, the intracellular bacteria might compensate for the loss, changing an otherwise lethal mutation into a neutral one that might become fixed. From that point onwards, the host would be locked in a relationship with a symbiont that can provide the lost function. The lack of unusual pathways conserved among the essential symbionts of *Euplotes*, or of genes absent in their close free-living relatives, seems to point in this direction. Why these particular bacteria are the only ones involved in the process is also not known, but it may be due to a mix of chance and pre-adaptation, with the common free-living *Polynucleobacter* and *Devosia*[10] simply being taken up (e.g. eaten) more frequently, and the rare *Protistobacter* possibly sharing some predisposing trait with the related *Polynucleobacter*.

The *Euplotes-Polynucleobacter* system is one of the most straightforward examples of repeated symbiont replacements, and we now show this extends to all the essential symbionts in this host-complex. Demonstrating cyclic replacements requires a lot of genomic data, and we predict that as more systems are investigated, it will come to represent more of a rule than an exception. Moreover, each of these cases provides valuable and rare replicates of the same symbiotic event: *Polynucleobacter*, *Protistobacter*, and *Devosia* can be seen as different "treatments" in this evolutionary experiment, where host and selective pressures are similar but the starting points are different.

Acknowledgments

The work was funded by the Gordon and Betty Moore Foundation (https://doi.org/10.37807GMBF9201).

References

1. Moran, N.A. (2006). Symbiosis. Curr. Biol. *16*, PR866–871.
2. Boscaro, V., Felletti, M., Vannini, C., Ackerman, M.S., Chain, P.S.G., Malfatti, S., Vergez, L.M., Shin, M., Doak, T.G., Lynch, M., and Petroni, G. (2013). *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. Proc. Natl. Acad. Sci. U.S.A. *110*, 18590–18595.
3. Boscaro, V., Kolisko, M., Felletti, M., Vannini, C., Lynn, D.H., and Keeling, P.J. (2017). Parallel genome reduction in symbionts descended from closely related free-living bacteria. Nat. Ecol. Evol. *1*, 1160–1167.
4. Chong, R.A., and Moran, N.A. (2018). Evolutionary loss and replacement of *Buchnera*, the obligate endosymbiont of aphids. ISME J. *12*, 898–908.
5. Dial, D.T., Weglarz, K.M., Aremu, A.O., Havill, N.P., Pearson, T.A., Burke, G.R., and von Dohlen, C.D. (2021). Transitional genomes and nutritional role reversals identified for dual symbionts of adelgids. ISME J. *12*, 898–908.
6. Mao, M., and Bennett, G.M. (2020). Symbiont replacements reset the co-evolutionary relationship between insects and their heritable bacteria. *ISME J. 14*, 1384–1395.
7. McCutcheon, J.P., Boyd, B.M., and Dale, C. (2019). The life of an insect endosymbiont from the cradle to the grave. Curr. Biol. *29*, R485–R495.
8. Boscaro, V., Husnik, F., Vannini, C., and Keeling, P.J. (2019). Symbionts of the ciliate *Euplotes*: diversity, patterns and potential as models for bacteria-eukaryote endosymbioses. Proc. R. Soc. B. *286*, 20190693.
9. McCutcheon, J.P. (2021). The genomics and cell biology of host-beneficial intracellular infections. Annu. Rev. Cell. Develop. Biol. *37*, 115–142.
10. Savio, D., Sinclair, L., Ijaz, U.M., Parajka, J., Reischer, G.H., Stadler, P., Blaschke, A.P., Blöschl, G., Mach, R.L., Kirschner, A.K.T., et al. (2015). Bacterial diversity along a 2600 km river continuum. Environ. Microbiol. *17*, 4994–5007.

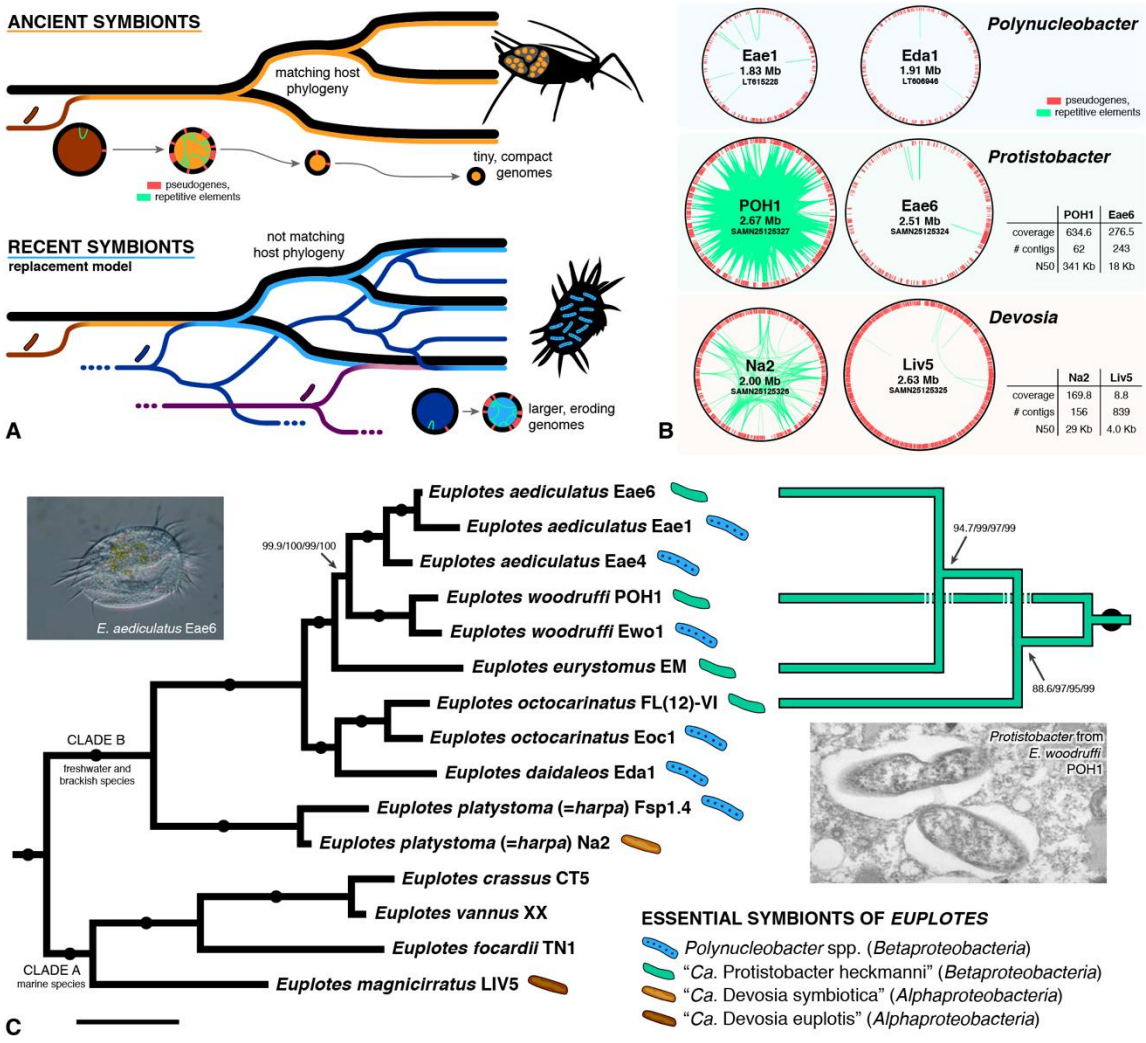

Figure 1. (A) Ancient, stable endosymbionts (e.g., *Buchnera* in aphids) can be distinguished from recent ones (e.g., *Polynucleobacter* in *Euplotes*) based on genomic features and the extent of co-differentiation with their hosts. (B) The newly characterized genomes of *Protistobacter* and *Devosia* strains are slightly larger and show even more conspicuous signs of genome erosion than symbiotic *Polynucleobacter*. (C) Phylogenomic branching order differs between hosts (tree on the left) and symbionts (cladogram on the right) in strains of *Protistobacter*-harbouring *Euplotes*. Numbers associated to nodes represent, from left to right: SH-aLRT values, ultrafast bootstrap support, non-parametric bootstrap support (1,000 pseudoreplicates), and ultrafast bootstrap support obtained with an alternative model (LG+C20+F+G4); black dots mark fully supported nodes. The bar stands for an expected divergence of 0.05 for the *Euplotes* tree.

Ancient symbiosis, recent symbionts: all essential endobacteria of the ciliate *Euplotes* are cyclically replaced

Supplementary Methods

New data in this paper come from six *Euplotes* strains harbouring either *Protistobacter* (Eae6, POH1, EM, FL(12)-VI) or symbiotic *Devosia* (freshwater, Na2, and marine, LIV5). All ciliates and bacterial strains were identified at the species level in previous works[8], which also described DNA isolation and Illumina sequencing. MinION reads were additionally generated starting from the same extracted DNA. Illumina read quality trimming was performed using Trimmomatic[11] v0.36. Preliminary genome assemblies were generated from metagenomic data using SPAdes[12], then contigs from the host and environmental bacteria were removed using a combination of taxonomic, coverage and GC-content cut-offs using BlobTools[13] v1.0. Quality-trimmed reads were remapped to the resulting contigs using Bowtie 2[14] v2.2.6. MinION reads were also mapped to the contigs using GraphMap[15] v0.5.2. Mapped Illumina and MinION reads were then used together to generate a hybrid assembly with Unicycler[16] v0.4.6. Bandage[17] v0.8.1 was used throughout the pipeline to visually assess quality and contiguity of the assemblies. Gene annotation was performed with Prokka[18] v1.12 and pseudogenes detected with Pseudofinder[19] v3.0. Transposons and other repeated elements were identified with ISfinder[20] and BLASTP.

Phylogenomic marker genes were identified using BLAST, HMMER[21], and exonerate[22], and were manually curated using single-gene trees. Main phylogenomic analyses and AU tests were performed using IQ-TREE[23] (*Euplotes*: 79,840 aminoacids, LG+F+R4 model, 10 other spirotrichean ciliates as outgroup. *Protistobacter*: 30,874 aminoacids, LG+F+R10 model, 39 *Polynucleobacter* and 31 other *Burkholderiaceae* bacteria as outgroup).

Additional references

11. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
12. Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. *19*, 455–477.
13. Laetsch, D.R., and Blaxter, M.L. (2017). BlobTools: interrogation of genome assemblies. F1000 Research *6*, 1287.

14. Langmead, B. and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Met. *9*, 357–359.
15. Sović, I., Šikić, M, Wilm, A., Fenlon, S.N., Chen, S., and Nagarajan, N. (2016). Fast and sensitive mapping of nanopore sequencing reads with GraphMap. Nat. Comm. *7*, 11307.
16. Wick, R.R., Judd, L.M., Gorrie, C.L., and Holt, K.E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. *13*, e1005595.
17. Wick, R.R., Schultz, M.B., Zobel, J., and Holt, K.E. (2015). Bandage: interactive visualization of *de novo* genome assemblies. Bioinformatics *31*, 3350–3352.
18. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068–2069.
19. Syberg-Olsen, M.J., Garber, A.I., Keeling, P.J., McCutcheon, J.P., and Husnik, F. (2021) Pseudofinder: detection of pseudogenes in prokaryotic genomes. bioRxiv, doi: 10.1101/2021.10.07.463580
20. Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. *34*, D32–D36.
21. Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. *39*, S29–S37.
22. Slater, G.St.C., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics *6*, 31.
23. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2014). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molec. Biol. Evol. *32*, 268–274.