The nitroplast and its relatives support a universal model of features predicting gene retention in endosymbiont and organelle genomes

Iain G. Johnston[1, 2, *]

1. Department of Mathematics, University of Bergen, Bergen, Norway
2. Computational Biology Unit, University of Bergen, Bergen, Norway
* correspondence to iain.johnston@uib.no

Abstract

Endosymbiotic relationships have shaped eukaryotic life. As endosymbionts coevolve with their host, towards full integration as organelles, their genomes tend to shrink, with genes being completely lost or transferred to the host nucleus. Modern endosymbionts and organelles show diverse patterns of gene retention, and why some genes and not others are retained in these genomes is not fully understood. Recent bioinformatic study has explored hypothesized influences on these evolutionary processes, finding that hydrophobicity and amino acid chemistry predict patterns of gene retention, both in organelles across eukaryotes and in less mature endosymbiotic relationships. The exciting new discovery and elucidation of more endosymbiotic relationships affords an independent set of instances to test this theory. Here we compare the properties of retained genes in the recently reported nitroplast, two related cyanobacterial endosymbionts which form "spheroid bodies" in their host cells, and a range of other endosymbionts, with free-living relatives. We find that in each case, the symbiont's genome encodes proteins with higher hydrophobicity and lower ammonium pK_a than their free-living relative, supporting the data-derived model predicting the retention propensity of genes across endosymbiont and organelle genomes.

Introduction

Eukaryotic life has numerous independent examples of endosymbiotic relationships. These include integrated organelles like the mitochondrion and plastid acquired billions of years ago (Smith & Keeling, 2015), through acquisition of a cyanobacterium around 100 million years ago to form the chromatophore in *Paulinella* algae (Gabr et al., 2020), to more recent acquisitions of bacterial endosymbionts in insects (Husnik & Keeling, 2019). Other examples include the nitrogen-fixing endosymbiont in *Azolla* water ferns (Peters & Meeks, 1989; Ran et al., 2010), a cyanobacterial symbiont of diatoms (Flores et al., 2022), a denitrifying endosymbiont in a ciliate host (Graf et al., 2021), "spheroid body" compartments in diatoms (Nakayama et al., 2011), and a nitrogen-fixing symbiont accompanying a picoeukaryotic alga (Thompson et al., 2012) which has since been characterized as an integrated organelle dubbed the "nitroplast" (Coale et al., 2024). In each of these cases, the proto-endosymbiont originally possessed a full genome. As endosymbiotic relationships proceed and endosymbionts become more and more integrated organelles in the host cell, the endosymbiont genome tends to become reduced, with genes completely lost or transferred to the host nucleus (Giannakis et al., 2022; McCutcheon & Moran, 2012; Moran et al., 2009). In some cases this process has been complete, leaving mitochondrion-related organelles with no mitochondrial DNA (Hjort et al., 2010; Makiuchi & Nozaki, 2014). In other cases, a subset of genes is retained in the organelle or endosymbiont.

The retained subset of genes in organelles and endosymbionts varies dramatically across eukaryotes, and the features favouring gene retention are not completely understood (Butenko et al., 2024; García-Pascual et al., 2022; Giannakis et al., 2023, 2024; McCutcheon & Moran, 2012; Smith & Keeling, 2015). Hypotheses have often focused on mitochondria and plastids, and have included roles for hydrophobicity (making it harder for nuclear-encoded genes to be

imported to the organelle (Björkholm et al., 2015; von Heijne, 1986)); favouring local individual control of organelles (colocalization for redox regulation or CoRR (Allen, 2015)); the economics of maintaining and expressing genes from different compartments (Kelly, 2021), and others (quantitatively compared in (Giannakis et al., 2022)).

Recent data-driven work has shown that models containing the same features (including hydrophobicity and acid dissociation constants) predict retention profiles in mitochondria and plastids across eukaryotes (Giannakis et al., 2022; Grub et al., 2022). Strikingly, when trained on mitochondria, this model predicts plastid retention patterns (and vice versa), suggesting that similar principles may shape gene retention in the two cases. Specifically, genes encoding products with high hydrophobicity and low ammonium pK_a were more likely to be retained, along with a role for the centrality of a protein subunit in its complex (related to CoRR). Hydrophobicity and pK_a values were also shown to differ systematically between other endosymbionts and their free-living relatives, in a set of relationships in insects, algae, and protists (Husnik & Keeling, 2019) (Fig. 1A).

The ongoing elucidation of examples along the spectrum from endosymbiont to mature organelle, including the nitroplast (Coale et al., 2024) and its cyanobacterial relatives (Nakayama & Inagaki, 2017) allow an independent test of this "universal" model. In this note, we ask whether these other relationships, reflecting a spectrum of maturity of endosymbiosis, support this picture.

Methods

Following the pipeline from (Giannakis et al., 2022), we obtained coding sequence records for the collection of genomes in endosymbionts, organelles, and free-living relatives in Table 1. This set was originally chosen from a comprehensive review (Husnik & Keeling, 2019); we included *Wolbachia* as a famous, though not obligate, endosymbiont example. Close free-living relatives were identified from phylogenetic analysis in the references cited therein and confirmed with NCBI Common Taxonomy Tree (Federhen, 2012). For the Rickettsiales examples, most close relatives were also endosymbionts (often parasites), so we took statistics from a sister clade *Ca*. Pelagibacter ubique, the ubiquitous marine bacterium (Rappé et al., 2002). We also included mitochondria and chloroplasts from different species for comparison, compared to modern-day Rickettsia and cyanobacterial examples (Keeling, 2010; Roger et al., 2017). We computed statistics for the protein corresponding to each gene in each record, specifically taking the mean hydrophobicity and mean carboxyl and ammonium pK_a values across amino acid residues in each sequence, using lookup tables from https://www.sigmaaldrich.com/NO/en/technical-documents/technical-article/protein-biology/protein-structural-analysis/amino-acid-reference-chart . Analysis was performed in Biopython (Cock et al., 2009) and R (R Core Team & Team, 2022) with libraries *ggplot2* (Wickham, 2016) and *ggpubr* (Kassambara, 2020) for visualization. Code for the analysis and visualization is freely available at https://github.com/StochasticBiology/endosymbiont-gene-loss.

Results

Genes retained in the nitroplast, spheroid body endosymbionts, and *Richelia* symbiont showed substantial increased hydrophobicity compared to their free-living relatives (Fig. 1B). The spheroid bodies and *Richelia* showed a hydrophobicity increase on a similar scale to that seen in the *Paulinella* chromatophore (Fig. 1A). The increase was slightly greater in the nitroplast, on a similar scale to the nitrogen-fixing *Nostoc azollae* symbiont in the *Azolla* water fern.

Endosymbiont / organelle	Free-living/non-organelle relative	Notes and references
Mitochondrion (*Reclinomonas americana* and *Plasmodium falciparum*) (NC_001823.1 and NC_037526.1)	*Rickettsia typhi* (CP003398.1)	Bacterial-derived organelle found across almost all eukaryotes (Roger et al., 2017; Smith & Keeling, 2015)
Plastid (*Chondrus crispus* and *Hydnora visseri*) (NC_020795.1 and NC_029358.1)	*Synechococcus* PCC 7002 (CP000951)	Bacterial-derived organelle found across photosynthetic (and other) eukaryotes (Keeling, 2010; Smith & Keeling, 2015)
Paulinella chromatophore (CP000815.1)	*Synechococcus* PCC 7002 (CP000951)	Cyanobacterium-derived organelle in an alga (Lhee et al., 2019)
Nitroplast (UCYN-A, *Ca.* Atelocyanobacterium thalassa) (CP001842.1)	*Crocosphaera watsonii* (GCF_000235665.1)	Nitrogen-fixing organelle in algae (Coale et al., 2024; Thompson et al., 2012)
Epithemia turgida spheroid body (AP012549)	*Rippkaea orientalis* (GCF_000021805.1)	Cyanobacterium-derived compartment in diatom (Nakayama & Inagaki, 2017); closely related to *Rhopalodia gibberula* spheroid body and related to nitroplast (Qiu et al., 2021)
Rhopalodia gibberula spheroid body (AP018341.1)	*Cyanothece sp.* PCC 8801 (CP001287.1)	Cyanobacterium-derived compartment in diatom (Nakayama & Inagaki, 2017); closely related to *Epithemia turgida* spheroid body and related to nitroplast (Qiu et al., 2021). *Rippkaea* is a free-living relative; comparison with another free-living relative *Cyanothece* is included to link with (Giannakis et al., 2022).
Ca. Azoamicus ciliaticola (NZ_LR794158.1)	*Legionella clemsonensis* (NZ_CP016397)	Denitrifying endosymbiont in an anaerobic ciliate (Graf et al., 2021); most relatives, including Legionella, are largely intracellular
Nostoc azollae (CP002059.1)	*Raphidiopsis brookii* (ACYB01000001.1)	Nitrogen-fixing cyanobacterium in a water fern (Ran et al., 2010)
Richelia intracellularis (GCA_000350105.1)	*Richelia sinica* (GCF_019056575.1)	Cyanobacterial symbiont in diatom (Flores et al., 2022)
Nasuia deltocephalinicola (CP013211.1)	*Herbaspirillum seropedicae* (CP002039.1)	Bacterial endosymbiont of insects (Bennett & Moran, 2013)
Ca. Sulcia muelleri (CP001981.1)	*Porphyromonas gingivalis* (AE015924.1)	Bacterial endosymbiont of insects (McCutcheon & Moran, 2007); "free-living" relative does invade cells but can survive independently in oral cavity.
Ca. Tremblaya phenacola (CP003982.1)	*Sodalis praecaptivus* (CP006569.1)	Bacterial endosymbiont of insects (Enomoto et al., 2017)
Ca. Hodgkinia cicadicola (CP008699)	*Rhizobium etli* (CP007641.1)	α-proteobacterial symbiont of cicadas (McCutcheon et al., 2009)
Ca. Pinguicoccus supinus (CP039370.1)	*Coraliomargarita akajimensis* (CP001998.1)	Bacterial endosymbiont in ciliate (Serra et al., 2020); partner is not closest sequence found, but is closest annotated sequence in putative phylogeny
Ca. Fokinia solitaria (CP025989.1)	*Ca.* Pelagibacter ubique (CP000084.1)	Rickettsiales endosymbiont (*Ca.* Midichloriaceae family) in ciliate (Floriano et al., 2018); like *Wolbachia,* all closest relatives are intracellular Rickettsiales – relative taken from a sister group.
Wigglesworthia glossinidia (*GCF_000247565.1*)	*Pantoea agglomerans* (GCF_019048385.1)	Gammaproteobacterial endosymbiont of tsetse fly (Akman et al., 2002)
Buchnera aphidicola (GCF_003099975.1)	*Pantoea agglomerans* (GCF_019048385.1)	Gammaproteobacterial endosymbiont of aphids (van Ham et al., 2003)
Wolbachia pipentis (GCF_014107475.1)	*Ca.* Pelagibacter ubique (CP000084.1)	Rickettsiales endosymbiont, can exist as insect endosymbiont or independently (Werren et al., 2008); like Fokinia, all closest relatives are intracellular Rickettsiales – relative taken from a sister group.

Table 1. Pairs of endosymbionts and free-living relative, and organelles and non-organelle relatives, used for comparison in this study, with NCBI accessions and references supporting the choice of relative. The species chosen for mitochondria and plastids correspond to very high (*R. americana*, *C. crispus*) and very low (*P. falciparum*, *H. visseri*) organelle DNA gene counts.

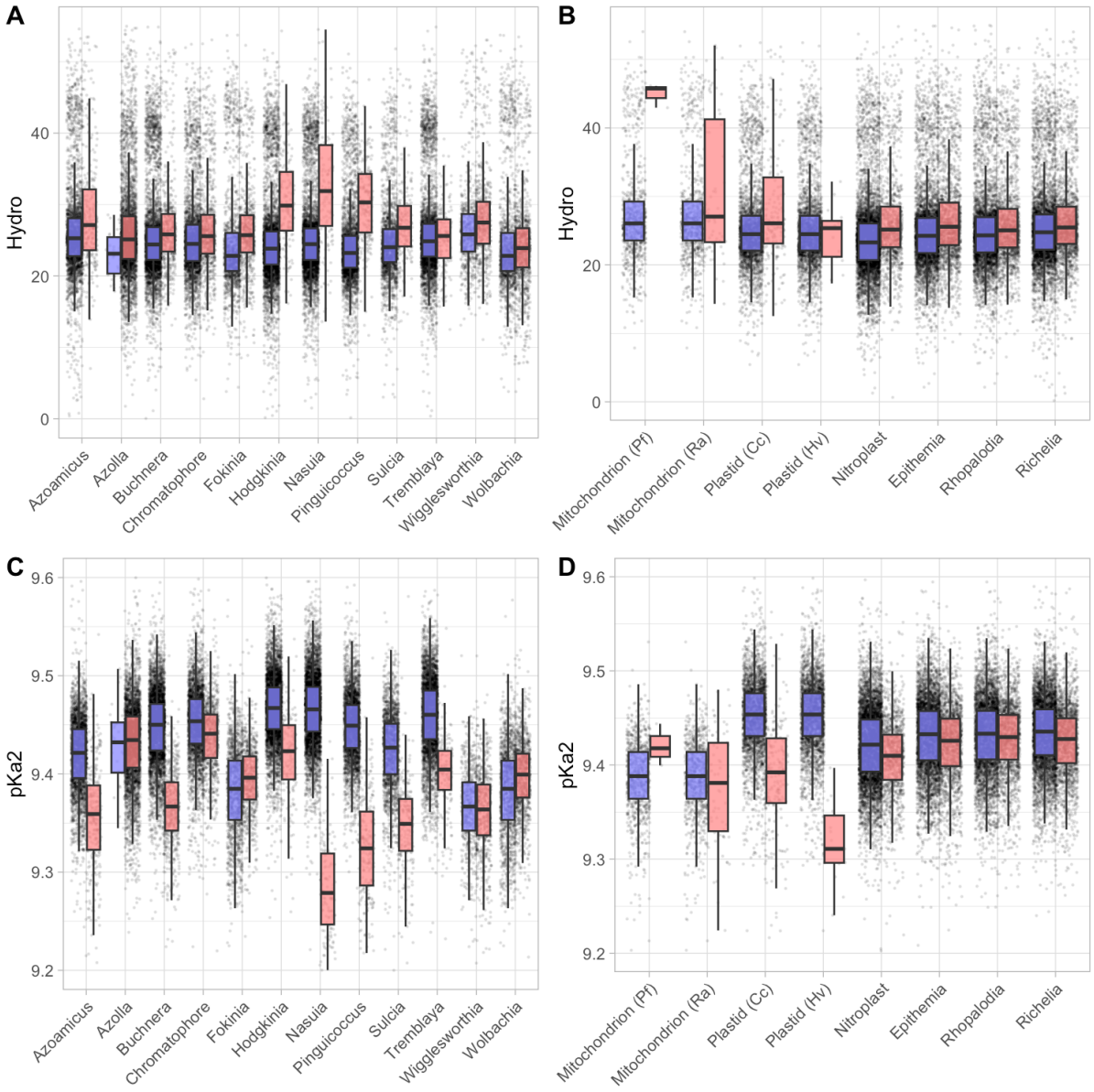

Figure 1. **Differences between endosymbiont and free-living gene profiles consistently agree with model predictions.** (A-B) Hydrophobicity and (C-D) ammonium pK_a distributions in genes retained in endosymbionts and organelles (red) and free-living close relatives (blue). Individual genes are shown as jittered points; boxplots give a summary distribution. Pf, *Plasmodium falciparum*; Ra, *Reclinomonas americana*; Cc, *Chondrus crispus*; Hv, *Hydnora visseri*.

Ammonium pK_a values were found to predict gene retention patterns in mitochondria and chloroplasts, but were not explicitly examined previously in other endosymbionts in (Giannakis et al., 2022). Fig. 1C shows the trends across the relationships explored in that study. With three exceptions (*Azolla*, *Fokinia* and *Wolbachia*, the latter two of which are in the same order), ammonium pK_a values are lower (sometimes dramatically so) in endosymbionts than in free-living relatives, matching the behaviour expected from the universal model. Plastids also show this behaviour; the *Plasmodium* mitochondrion we consider instead has a higher average ammonium pK_a. This is not inconsistent with the universal model picture: the very high

difference in hydrophobicity in the *Plasmodium* mitochondria overcomes the pK_a term in the predictive model, so that the three genes are predicted to have a high retention index. In the set of newly-considered relationships in this study (nitroplasts, spheroid bodies, and *Richelia*), each endosymbiont also showed lower ammonium pK_a values than its free-living relative (Fig. 1D), again on a similar scale to the chromatophore, with this effect stronger for the nitroplast than for the spheroid bodies.

The gene-by-gene correlation across our dataset of hydrophobicity and ammonium pK_a value is weak ($r^2 = 0.022$), suggesting that Fig. 1A-B and 1C-D are not just reporting the same effect twice over; the behaviour in hydrophobicity is largely independent of the behaviour in pK_a. This reflects the fact that in the original model selection process for organelle gene retention, the two features were selected together, suggesting that they provide independent information about gene retention propensity.

Significance testing for the individual comparisons in Fig. 1 is not directly meaningful, as the full sets of genes from each organisms are being reported – there is no sampling noise to account for, so statements about mean differences are not subject to meaningful uncertainty. The more interesting hypothesis test relates to the observation of partnership comparisons, against the null hypothesis that hydrophobicity and pK_a do not differ between symbionts and relatives. If our symbiont-relative pairs are treated as independent, the probability of these eight new observations (four partnerships, for hydrophobicity and pK_a) all agreeing with the theory under the null hypothesis is $1/2^8 \approx 0.004$. If the two spheroid body partnerships are regarded as reflecting the same case, the probability becomes $1/2^6 \approx 0.016$.

Discussion

From the study of mitochondria alone, a model involving hydrophobicity and amino acid biochemistry was found to predict gene retention patterns (Giannakis et al., 2022; Johnston & Williams, 2016). The same model with the same parameters (positive effect for hydrophobicity, negative effect for ammonium pK_a) also predicts plastid gene retention (Giannakis et al., 2022; Grub et al., 2022). We have found here that the same influences separate genes retained in endosymbionts across a range of maturities, from recent insect acquisitions to the more integrated and established chromatophore and nitroplast.

Why these features? Hydrophobicity was originally argued to challenge protein import to the organelle from the remote encoding of the nucleus (von Heijne, 1986), and has since been suggested to influence mistargeting of protein products (Björkholm et al., 2015). In many of the relationships we consider, it is far from clear whether symbiont genes have been transferred to the nucleus, so whether hydrophobicity acts as a barrier to transfer is less well-posed. However, it can likely still act as a barrier to *loss*. All our cases do seem to involve reduction of the symbiont genome, likely due in part to redundancy, where host-encoded proteins can be used by the symbiont. For this to be the case, host-encoded proteins still require import to the endosymbiont, so the argument that hard-to-import machinery is more likely to be retained can still be used.

We previously and very speculatively suggested that links to pK_a could relate to the necessity of assembling proteins in a cellular compartment where pH may be different (Giannakis et al., 2022). pK_a reports how easily protons are lost from amino acids under different pH conditions, and hence necessarily influences the dynamics of peptide formation in translation (Watts & Forster, 2010). This influence leads to differences in peptide formation dynamics in different pH environments (Johansson et al., 2011). The differences in compartmental properties – including

pH – as endosymbiotic relationships evolve could conceivably therefore mean that the inside-compartment ease of assembling proteins is greater for those with particular pK_a profiles. However, further and more detailed investigation is needed to explore this hypothesis.

Of course, the consideration of two features alone cannot describe all the possible mechanisms and influences shaping endosymbiont genomes across relationships. The performance of models considering these features for mitochondrial and plastid gene retention is reasonable (Spearman's ρ around 0.5-0.6 for mtDNA and ptDNA genes outside the training sets (Giannakis et al., 2022)), but the effect sizes are smaller in these less mature endosymbiotic cases and the predictive power of such models will be more limited. This note intends only to highlight that these exciting emerging cases provide further independent support for these features having some possible (not complete) influence over endosymbiont genome evolution, not that the question is resolved!

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement no. 805046 [EvoConBiO] to IGJ).

References

Akman, L., Yamashita, A., Watanabe, H., Oshima, K., Shiba, T., Hattori, M., & Aksoy, S. (2002). Genome sequence of the endocellular obligate symbiont of tsetse flies, Wigglesworthia glossinidia. *Nature Genetics*, *32*(3), 402–407. https://doi.org/10.1038/ng986

Allen, J. F. (2015). Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proceedings of the National Academy of Sciences*, *112*(33), 10231–10238. https://doi.org/10.1073/pnas.1500012112

Bennett, G. M., & Moran, N. A. (2013). Small, Smaller, Smallest: The Origins and Evolution of Ancient Dual Symbioses in a Phloem-Feeding Insect. *Genome Biology and Evolution*, *5*(9), 1675–1688. https://doi.org/10.1093/gbe/evt118

Björkholm, P., Harish, A., Hagström, E., Ernst, A. M., & Andersson, S. G. (2015). Mitochondrial genomes are retained by selective constraints on protein targeting. *Proceedings of the National Academy of Sciences*, *112*(33), 10154–10161.

Butenko, A., Lukeš, J., Speijer, D., & Wideman, J. G. (2024). Mitochondrial genomes revisited: Why do different lineages retain different genes? *BMC Biology*, *22*(1), 15. https://doi.org/10.1186/s12915-024-01824-1

Coale, T. H., Loconte, V., Turk-Kubo, K. A., Vanslembrouck, B., Mak, W. K. E., Cheung, S., Ekman, A., Chen, J.-H., Hagino, K., Takano, Y., Nishimura, T., Adachi, M., Le Gros, M., Larabell, C., & Zehr, J. P. (2024). Nitrogen-fixing organelle in a marine alga. *Science*, *384*(6692), 217–222. https://doi.org/10.1126/science.adk1075

Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., & de Hoon, M. J. L. (2009). Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, *25*(11), 1422–1423. https://doi.org/10.1093/bioinformatics/btp163

Enomoto, S., Chari, A., Clayton, A. L., & Dale, C. (2017). Quorum sensing attenuates virulence in Sodalis praecaptivus. *Cell Host & Microbe*, *21*(5), 629–636.

Federhen, S. (2012). The NCBI Taxonomy database. *Nucleic Acids Research*, *40*(D1), D136–D143. https://doi.org/10.1093/nar/gkr1178

Flores, E., Romanovicz, D. K., Nieves-Morión, M., Foster, R. A., & Villareal, T. A. (2022). Adaptation to an Intracellular Lifestyle by a Nitrogen-Fixing, Heterocyst-Forming Cyanobacterial Endosymbiont of a Diatom. *Frontiers in Microbiology*, *13*. https://doi.org/10.3389/fmicb.2022.799362

Floriano, A. M., Castelli, M., Krenek, S., Berendonk, T. U., Bazzocchi, C., Petroni, G., & Sassera, D. (2018). The Genome Sequence of "Candidatus Fokinia solitaria": Insights on Reductive Evolution in

Rickettsiales. *Genome Biology and Evolution*, *10*(4), 1120–1126. https://doi.org/10.1093/gbe/evy072

Gabr, A., Grossman, A. R., & Bhattacharya, D. (2020). Paulinella, a model for understanding plastid primary endosymbiosis. *Journal of Phycology*, *56*(4), 837–843. https://doi.org/10.1111/jpy.13003

García-Pascual, B., Nordbotten, J. M., & Johnston, I. G. (2022). *Cellular and environmental dynamics influence species-specific extents of organelle gene retention* (p. 2022.10.17.512581). bioRxiv. https://doi.org/10.1101/2022.10.17.512581

Giannakis, K., Arrowsmith, S. J., Richards, L., Gasparini, S., Chustecki, J. M., Røyrvik, E. C., & Johnston, I. G. (2022). Evolutionary inference across eukaryotes identifies universal features shaping organelle gene retention. *Cell Systems*, *13*(11), 874-884.e5. https://doi.org/10.1016/j.cels.2022.08.007

Giannakis, K., Richards, L., Dauda, K., & Johnston, I. (2023). Connecting species-specific extents of genome reduction in mitochondria and plastids. *bioRxiv*, 2023–12.

Giannakis, K., Richards, L., & Johnston, I. G. (2024). Ecological predictors of organelle genome evolution: Phylogenetic correlations with taxonomically broad, sparse, unsystematized data. *Systematic Biology*, syae009. https://doi.org/10.1093/sysbio/syae009

Graf, J. S., Schorn, S., Kitzinger, K., Ahmerkamp, S., Woehle, C., Huettel, B., Schubert, C. J., Kuypers, M. M. M., & Milucka, J. (2021). Anaerobic endosymbiont generates energy for ciliate host by denitrification. *Nature*, *591*(7850), 445–450. https://doi.org/10.1038/s41586-021-03297-6

Grub, L. K., Tsyba, N., & Patel, M. R. (2022). Should I stay, or should I go? Gene retention in organellar genomes. *Cell Systems*, *13*(11), 861–863. https://doi.org/10.1016/j.cels.2022.10.006

Hjort, K., Goldberg, A. V., Tsaousis, A. D., Hirt, R. P., & Embley, T. M. (2010). Diversity and reductive evolution of mitochondria among microbial eukaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1541), 713–727. https://doi.org/10.1098/rstb.2009.0224

Husnik, F., & Keeling, P. J. (2019). The fate of obligate endosymbionts: Reduction, integration, or extinction. *Current Opinion in Genetics & Development*, *58–59*, 1–8. https://doi.org/10.1016/j.gde.2019.07.014

Johansson, M., Ieong, K.-W., Trobro, S., Strazewski, P., Åqvist, J., Pavlov, M. Y., & Ehrenberg, M. (2011). pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proceedings of the National Academy of Sciences*, *108*(1), 79–84. https://doi.org/10.1073/pnas.1012612107

Johnston, I. G., & Williams, B. P. (2016). Evolutionary Inference across Eukaryotes Identifies Specific Pressures Favoring Mitochondrial Gene Retention. *Cell Systems*, *2*(2), 101–111. https://doi.org/10.1016/j.cels.2016.01.013

Kassambara, A. (2020). Ggpubr:"ggplot2" based publication ready plots. *R Package Version 0.4. 0*, *438*.

Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1541), 729–748. https://doi.org/10.1098/rstb.2009.0103

Kelly, S. (2021). The economics of organellar gene loss and endosymbiotic gene transfer. *Genome Biology*, *22*(1), 345. https://doi.org/10.1186/s13059-021-02567-w

Lhee, D., Ha, J.-S., Kim, S., Park, M. G., Bhattacharya, D., & Yoon, H. S. (2019). Evolutionary dynamics of the chromatophore genome in three photosynthetic Paulinella species. *Scientific Reports*, *9*(1), 2560. https://doi.org/10.1038/s41598-019-38621-8

Makiuchi, T., & Nozaki, T. (2014). Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. *Biochimie*, *100*, 3–17.

McCutcheon, J. P., McDonald, B. R., & Moran, N. A. (2009). Origin of an Alternative Genetic Code in the Extremely Small and GC–Rich Genome of a Bacterial Symbiont. *PLOS Genetics*, *5*(7), e1000565. https://doi.org/10.1371/journal.pgen.1000565

McCutcheon, J. P., & Moran, N. A. (2007). Parallel genomic evolution and metabolic interdependence in an ancient symbiosis. *Proceedings of the National Academy of Sciences*, *104*(49), 19392–19397. https://doi.org/10.1073/pnas.0708855104

McCutcheon, J. P., & Moran, N. A. (2012). Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology*, *10*(1), 13–26.

Moran, N. A., McLaughlin, H. J., & Sorek, R. (2009). The Dynamics and Time Scale of Ongoing Genomic Erosion in Symbiotic Bacteria. *Science*, *323*(5912), 379–382. https://doi.org/10.1126/science.1167140

Nakayama, T., Ikegami, Y., Nakayama, T., Ishida, K., Inagaki, Y., & Inouye, I. (2011). Spheroid bodies in rhopalodiacean diatoms were derived from a single endosymbiotic cyanobacterium. *Journal of Plant Research*, *124*(1), 93–97. https://doi.org/10.1007/s10265-010-0355-0

Nakayama, T., & Inagaki, Y. (2017). Genomic divergence within non-photosynthetic cyanobacterial endosymbionts in rhopalodiacean diatoms. *Scientific Reports*, *7*(1), 13075. https://doi.org/10.1038/s41598-017-13578-8

Peters, G. A., & Meeks, J. C. (1989). The Azolla-Anabaena symbiosis: Basic biology. *The Azolla-Anabaena Symbiosis: Basic Biology*, *40*, 193–210.

Qiu, G.-W., Jiang, H.-B., Lis, H., Li, Z.-K., Deng, B., Shang, J.-L., Sun, C.-Y., Keren, N., & Qiu, B.-S. (2021). A unique porin meditates iron-selective transport through cyanobacterial outer membranes. *Environmental Microbiology*, *23*(1), 376–390. https://doi.org/10.1111/1462-2920.15324

R Core Team, A., & Team, R. C. (2022). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2012*.

Ran, L., Larsson, J., Vigil-Stenman, T., Nylander, J. A. A., Ininbergs, K., Zheng, W.-W., Lapidus, A., Lowry, S., Haselkorn, R., & Bergman, B. (2010). Genome Erosion in a Nitrogen-Fixing Vertically Transmitted Endosymbiotic Multicellular Cyanobacterium. *PLOS ONE*, *5*(7), e11486. https://doi.org/10.1371/journal.pone.0011486

Rappé, M. S., Connon, S. A., Vergin, K. L., & Giovannoni, S. J. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature*, *418*(6898), 630–633. https://doi.org/10.1038/nature00917

Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Current Biology: CB*, *27*(21), R1177–R1192. https://doi.org/10.1016/j.cub.2017.09.015

Serra, V., Gammuto, L., Nitla, V., Castelli, M., Lanzoni, O., Sassera, D., Bandi, C., Sandeep, B. V., Verni, F., Modeo, L., & Petroni, G. (2020). Morphology, ultrastructure, genomics, and phylogeny of Euplotes vanleeuwenhoeki sp. Nov. And its ultra-reduced endosymbiont "Candidatus Pinguicoccus supinus" sp. Nov. *Scientific Reports*, *10*(1), 20311. https://doi.org/10.1038/s41598-020-76348-z

Smith, D. R., & Keeling, P. J. (2015). Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. *Proceedings of the National Academy of Sciences*, *112*(33), 10177–10184. https://doi.org/10.1073/pnas.1422049112

Thompson, A. W., Foster, R. A., Krupke, A., Carter, B. J., Musat, N., Vaulot, D., Kuypers, M. M. M., & Zehr, J. P. (2012). Unicellular Cyanobacterium Symbiotic with a Single-Celled Eukaryotic Alga. *Science*, *337*(6101), 1546–1550. https://doi.org/10.1126/science.1222700

van Ham, R. C. H. J., Kamerbeek, J., Palacios, C., Rausell, C., Abascal, F., Bastolla, U., Fernández, J. M., Jiménez, L., Postigo, M., Silva, F. J., Tamames, J., Viguera, E., Latorre, A., Valencia, A., Morán, F., & Moya, A. (2003). Reductive genome evolution in Buchnera aphidicola. *Proceedings of the National Academy of Sciences*, *100*(2), 581–586. https://doi.org/10.1073/pnas.0235981100

von Heijne, G. (1986). Why mitochondria need a genome. *FEBS Letters*, *198*(1), 1–4.

Watts, R. E., & Forster, A. C. (2010). Chemical Models of Peptide Formation in Translation. *Biochemistry*, *49*(10), 2177–2185. https://doi.org/10.1021/bi1000273

Werren, J. H., Baldo, L., & Clark, M. E. (2008). Wolbachia: Master manipulators of invertebrate biology. *Nature Reviews Microbiology*, *6*(10), 741–751. https://doi.org/10.1038/nrmicro1969

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag. https://link.springer.com/book/10.1007/978-0-387-98141-3