

Review by [Sophie Abby](#), 21 Feb 2023 13:46

I thank the authors for this round of revision, I am happy with the replies to my comments and I only had a few minor points, mostly on wording/clarifications, typos corrections, and on the availability of gene trees.

We appreciate reviewer's comment over the manuscript, specially the phylogenetic section. We feel they really improved the manuscript.

Minor points:

- I believe it would be better to amend the Figshare repository to include there the gene trees, and/or to have them directly added as Sup Figures (at least for those discussed in main text – see also my comment below). It would probably be a more stable storage than on iTOL. In addition, as they are now displayed in iTOL, these trees are difficult to exploit/read and analyse, since the name of the species does not appear at tips. It is an issue for the support of the evolutionary scenarios proposed e.g., for Vitamin B.

Species names, together with the previous gene ID, are now available in all generated trees. ITOL repository has been updated accordingly. In addition, as suggested by the reviewer, we added the newick files to the Figshare repository.

- L. 172: “half-gaps” => the authors could be more explicit here regarding the filtering criteria used by Gblocks.

We completed the sentence as follows: “i) positions with more than 50% of the sequences being gaps in the OrthoFinder concatenated alignment (107812 positions) were filtered out with Gblocks v0.91b (21499 selected positions in 143 blocks)” (L171-173)

- Figure 2: the font of the figure is small, especially for the tree inset and its legend, as well as the numbers above the bars in the upset plot. May it be possible to gain some space by e.g. contracting a bit the barplot with the COPs numbers? Or any other figure rearrangement? Maybe also abbreviate the genera names along the tree to save some space?

We compressed the COPs plot and increase font sizes to a maximum which we felt doesn't compromise readability or plotting quality (eg. letters/numbers overlap).

- L. 282: “clustering pattern” based on what? It should be detailed here or reformulated (I assume, based on presence/absence of the genes?).

No, we refer to the single-gene trees clustering patten, specially the *C. burnetii*/*Coxiella-LE* lineage and their basal clades. However, we removed the sentence regarding the “clustering pattern” since we noticed it is redundant with the information given some lines below: “In addition, the topology of single-gene trees from the different biosynthetic pathways support their ancestral status in the *C. burnetii* and *Coxiella-LEs* lineage...” (old L298-301)

- L.287: Should “or gain(ed)” be removed? Otherwise I don’t understand the proposed scenario. I assumed ancestral presence and differential losses, is it what the authors are suggesting? Or given the following sentence on line 288, is that suggested that there were indeed (re)acquisitions via LGT? It should be clarified. Also, if such scenario is proposed, it would be important to provide a figure of the corresponding gene tree, at least as Supplementary Figure.

We acknowledge the reviewer to pointing out on that paragraph. Indeed, reviewer’s comment is right, we proposed that traits were in the ancestor and lost in different lineages. The confusion comes to the mention to biotin and our mistake on being unclear regarding “patchy” distributions of a trait (absence/presence) and its phylogenetic origin. Also, we added two single-gene tree phylogenies (lipA and ribF) as Supplementary Figures to illustrate that Coxiellaceae species cluster as in the species tree (phylogenomic) except for some topological differences caused by the position of *Coxiella* MAGs. We hope now this section is more clear (L280-288).

In addition, aren’t the statements on Lines 296-299 somehow contradicting that on line 288?

“As an example, single-gene phylogenies of the Biotin (B7) pathway generally present different topologies from the species tree”

VS

“While pantothenate (B5), pyridoxine (B6), thiamine (B1), biotin (B7), riboflavin (B2), and lipoic acid pathways are strongly retained in this group, nicotinic acid (B3) and folic acid (B9) are only complete in *C. burnetii* and *Coxiella*-LEs. In addition, the topology of single-gene trees from the different biosynthetic pathways support their ancestral status in the *C. burnetii* and *Coxiella*-LEs lineage (Supplementary Data).”

This section needs some clarifications.

It might be possible that the word “retained” created some confusion, specially after the addition of the results from single-gene phylogenies. Our intention was to describe the presence of the pathways on the mentioned cluster independently of how they arrived to that status (presence/absence) or their phylogenetic origin. Also, we mentioned the biotin since it is a well known case of HGT where single-gene trees and species trees does not correlate. We added the phylogenetic tree of bioD as an example in both the Supplementary Figures and Supplementary data. We modified the text (L294-305) to clarify that and highlight the fact that despite the HGT event, biotin genes form a monophyletic clade in the *C. burnetii* and *Coxiella*-LE lineage, hence, supporting that the biotin pathway was present in their last common ancestor.

Typos:

L. 70: “te” => “the”

Corrected

L. 71: “use” => “uses”

Corrected

L. 116: “paired-end sequences” maybe?

Corrected

L. 163: “..” => “.”

Corrected

L. 172: “pruned” => would “filtered” be more appropriate?

Changed

L. 194: “The IPC 2.0 web-server”

Corrected

L. 212: “to several single-copy genes” => “to that of several single-copy genes”

Corrected

Figure S2: “Species” on the right hand

Duplicated sequence deleted.

Figure 3: is it “Thiamine” (as spelled in main text)?

Both names are correct. We changed the figure as suggested for homogenization.

Line 331: remove “.”?

Removed.

Line 341: verb missing?

Part of the sentence was lost. Now is: “Among the predicted HGT (Table S7), **it is noticeable** the presence of a Sodium Hydrogen/Multiple resistance and pH (Sha/Mrp) antiporter.”

Figure S4: why not present species names as in the other figures? It is hard to read like this...

All genomes from Figure S4 are strains from *C. burnettii*. This information is given in the legend. We prefer not to repeat the species name over all strains.

L. 366: “One proposed”... unclear what this refers to.

We tried to clarify the sentence. Now it reads: “One proposed specific adaption of acidophilic bacteria concerns the modification of their proteomes, through the enrichment of proteins in basic residues. Hence, it is expected that proteomes from acidophilic bacteria present lower average isoelectric point (pI) than non-acidophilic ones (Baker-Austin and Dopson 2007).”

Review by [Adam Ossowicki](#), 03 Mar 2023 08:44

In this article the authors explore the evolutionary history of genus *Coxiella* genomes in the context of the microorganism lifestyle. The article itself underwent evolution documented in the evaluation process of PCI Microbiology. The manuscript is written in a clear way, the methodology is sound and well described, the choice of bioinformatic tools is well justified, visualizations are clear. I cannot find any major drawbacks in the manuscript. I also cannot wait for the continuation of this study. It is not necessary, but can still add to the manuscript proposed way to validate your findings. I also agree with one of the previous reviewers that the paper would benefit from a figure illustrating possible evolutionary scenarios and acquire gained/loss traits being a simple synthesis of your work (graphical abstract style).

We thank the Reviewer for his/her time and comments which helped to improve the manuscript. We also appreciate his/her kind words.