Review of: "Marine gregarine genomes reveal the breadth of apicomplexan diversity and provide new insights on gliding motility"

Miroslav Oborník[1]

1 The Czech Academy of Sciences

Potential competing interests: The author(s) declared that no potential competing interests exist.

The manuscript presents the draft genomes of two related eugregarine specimens identified as Porospora gigantea, a parasite of lobsters. The genomes were sequenced using Next Seq 500 Illumina technology. Despite the massive bacterial contamination, the authors obtained a relatively trustable genome assembly after filtering the contamination out. Both genomes appeared to be approximately 9 Mb in size, each coding for about 5300 genes. The authors used the data to investigate the structure of genomes, the evolutionary history of both specimens, and the composition and evolution of the glideosome machinery. The paper is well written, methods are appropriate, and results are well described and discussed. However, I still have some comments and objections:

1. There were high bacterial contaminations in the data. Wasn´t it possible to clear gregarines out of bacteria? When I consider the substantial differences in the cell size between gregarines and bacteria, there should be a physical method to separate them.
2. The authors suggested that the two specimens represent two separated species. They built their suggestion based on the nucleotide divergence of genomes, even though SSU rRNA markers differ between both samples only in a single nucleotide. It is not clear if the divergence of genomes is caused by coding, noncoding, intron, or repetitive sequences. I would be careful to make such suggestions before finding the particular source of the nucleotide divergence.
3. Personally, I do not like the graphical form of the phylogenomic tree in figure 4. It may imply that gregarines constitute an advanced group in the apicomplexan tree, but in fact, it is caused by the flexible node between gregarines and the core Apicomplexa, which is shown in a misleading position. It should be turned opposite, with gregarines appearing in the proximity of Cryptosporidium and core apicomplexans moving up.
4. It is not clear on what ground the tree in figure 5A was constructed. In this tree, gregarines are shown as an advanced group with no phylogenetic relation to Cryptosporidium. The core Apicomplexa are shown as an early branching group (the earliest are only cryptosporidia and chromerids). In my opinion, this way of presenting can be misleading.
5. Gregarines appeared polyphyletic in the tree in Figure S6. It should be discussed.
6. The orientation of all trees should be the same (e.g., all roots down).
7. In Table one, only the number of contigs is shown for C. velia, but the number of chromosomes in this alga is known (4 chromosomes) (Vazač J et al. Protist 169: 53-63.)

In summary, this is a well-written paper describing genomes of two specimens of P. gigantea, their evolution, and glideosome. The results are well presented and discussed. In my opinion, it requires minor revisions to be suitable for publication.