

Wellcome Trust Sanger Institute (SC) in their Cichlid diversity sequencing WTMGM student project.

By using SOAPaligner/soap2 (V2.21) (Li et al. 2009), we mapped all the raw reads to the reference genome, complete mitochondrial genome of *Alticorpus geoffreyi* (Genbank accession: NC_028033) (Qi et al. 2016). We assembled these reads which could map to the reference genome by SPAdes3 (V3.1.0) (Bankevich et al. 2012), and got the complete circular mitochondrial genome of *C. obliquidens*. Also, we used DOGMA (Wyman et al. 2004) and tRNAscan-SE 2.0 (Lowe & Eddy 1997) to annotate the complete mitochondrial genome.

The mitochondrial genome of *C. obliquidens* is 16,581 bp in length, contains 37 genes, including 2 ribosomal RNA genes (rRNA), 22 transfer RNA genes (tRNA) and 13 protein-coding genes (PCGs). The lengths of 22 tRNA genes range from 67 bp (tRNA^{Cys} and tRNA^{Ser}) to 74 bp (tRNA^{Leu} and tRNA^{Lys}), whereas 16S rRNA is 1676 bp and 12S rRNA is 941 bp. Twenty-eight genes contain 2 rRNA, 14 tRNA and 12 PCGs are H-strand, while the remaining 9 genes including 1 PCGs (ND6) and 8 tRNA are L-strand. All PCGs in *C. obliquidens* are started with ATG and stopped with TAN or AGA, except for COX1 started with GTG. The GC content of this mitochondrial genome is 45.94% (27.43% A, 26.63% T, 30.08% C, 15.86% G), similar to *A. geoffreyi* (45.82%) from the same family, Cichlidae.

Furthermore, we used MEGA7 (V7.0.25) (Kumar et al. 2016) to construct the phylogenetic tree on the complete mitochondrial genomes of *C. obliquidens*, 2 species with the same phylum, 2 species from the same class and other 10 closely related species by Maximum likelihood method (Figure 1). Those results would facilitate our understanding of the mitochondrial evolution of Cichlidae family.

Acknowledgements

The authors would like to thank Wellcome Trust Sanger Institute (SC) for generating and uploading the raw data of the *C. obliquidens* used in this study. This work was sponsored by Qing Lan project in Jiangsu Province.

Disclosure statement

The authors declare no conflict of interest.

Funding

This work was sponsored by Qing Lan project in Jiangsu Province.

References

- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, et al. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 19:455–477.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol.* 33:1870–1874.
- Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics.* 25:1966–1967.
- Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
- Meyer BS, Matschiner M, Salzburger W. 2015. A tribal level phylogeny of Lake Tanganyika cichlid fishes based on a genomic multi-marker approach. *Mol Phylogenet Evol.* 83:56–71.
- Qi DS, Tao J, Huang XJ, Jiang J, Wang M, Zhang LQ, Mei XY. 2016. The complete mitochondrial genome of the *Alticorpus geoffreyi*. *Mitochondrial DNA Part B Resources.* 1:404–405.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 20:3252–3255.