

Supplementary File S4

Phylogenetic tree analysis

Page 1 Figure legends

Page 3 (A) Tree by Maximum Likelihood method

Page 4 (B) Tree by Neighbor-Joining method

Page 5 (C) Tree by Minimum Evolution method

Page 6 (D) Tree by UPGMA method

Figure legends

(A) Molecular Phylogenetic analysis by Maximum Likelihood method. The evolutionary history was inferred based on the JTT matrix-based model [1]. The tree with the highest log likelihood (-15686.03) is shown. The percentage of trees in which the associated taxa clustered together is shown next to the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 209 amino acid sequences. All positions with less than 95% site coverage were eliminated. Evolutionary analyses were conducted in MEGA7 [2].

(B) The evolutionary history was inferred using the Neighbor-Joining method [3]. The optimal tree with the sum of branch length = 25.67950417 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [4]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [5] and are in the units of the number of amino acid substitutions per site. The analysis involved 209 amino acid sequences. All ambiguous positions were removed for each sequence pair. Evolutionary analyses were conducted in MEGA7 [2].

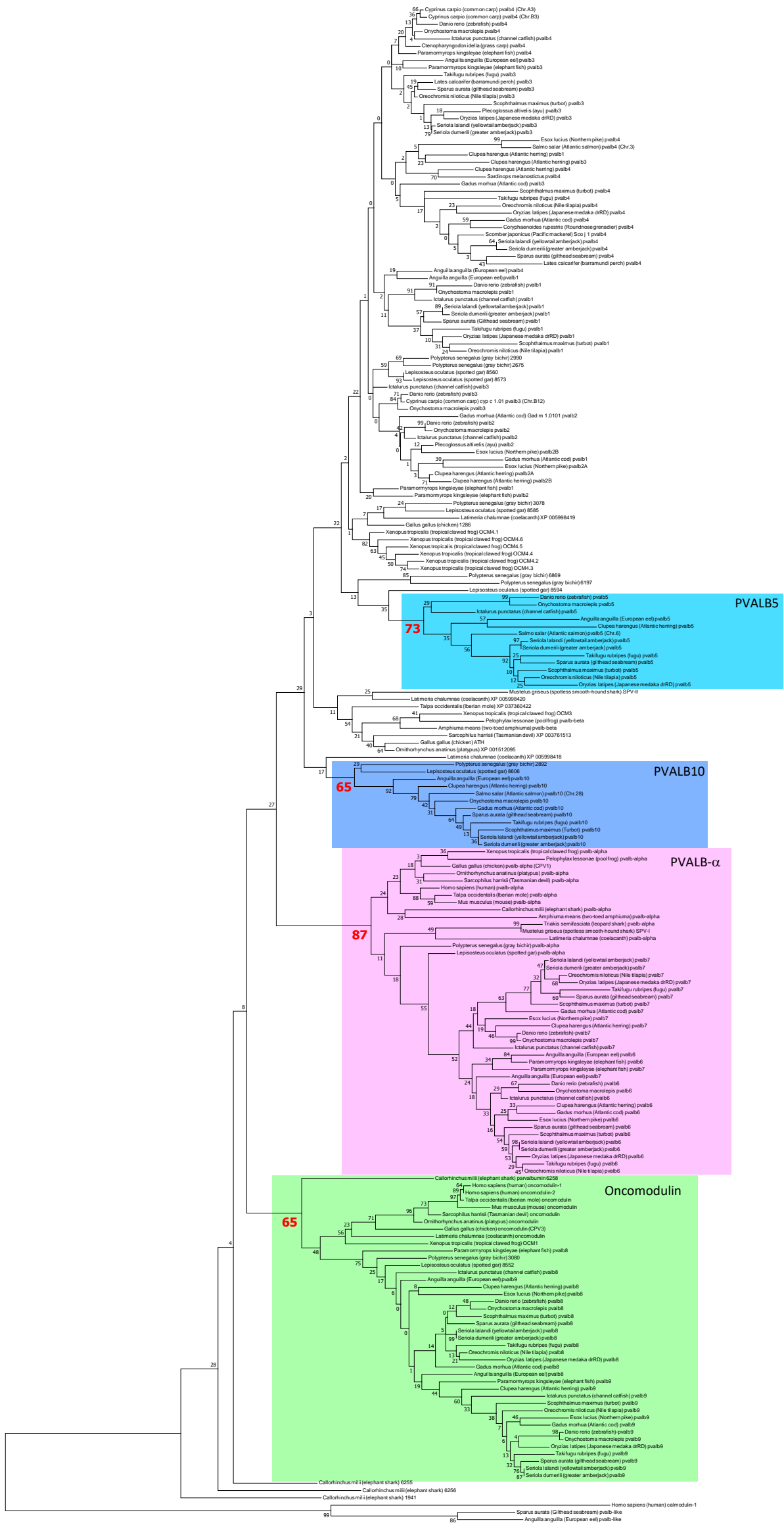
(C) The evolutionary history was inferred using the Minimum Evolution method [6]. The optimal tree with the sum of branch length = 25.15980363 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [4]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [5] and are in the units of the number of amino acid substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm [7] at a search level of 1. The Neighbor-joining algorithm [8] was used to generate the initial tree. The analysis involved 209 amino acid sequences. All ambiguous positions were removed for each sequence pair. Evolutionary analyses were conducted in MEGA7 [2].

(D) The evolutionary history was inferred using the UPGMA method [9]. The optimal tree with the sum of branch length = 24.98273800 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches [4]. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method [5] and are in the units of the number of amino acid substitutions per site. The analysis involved 209 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 109 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 [2].

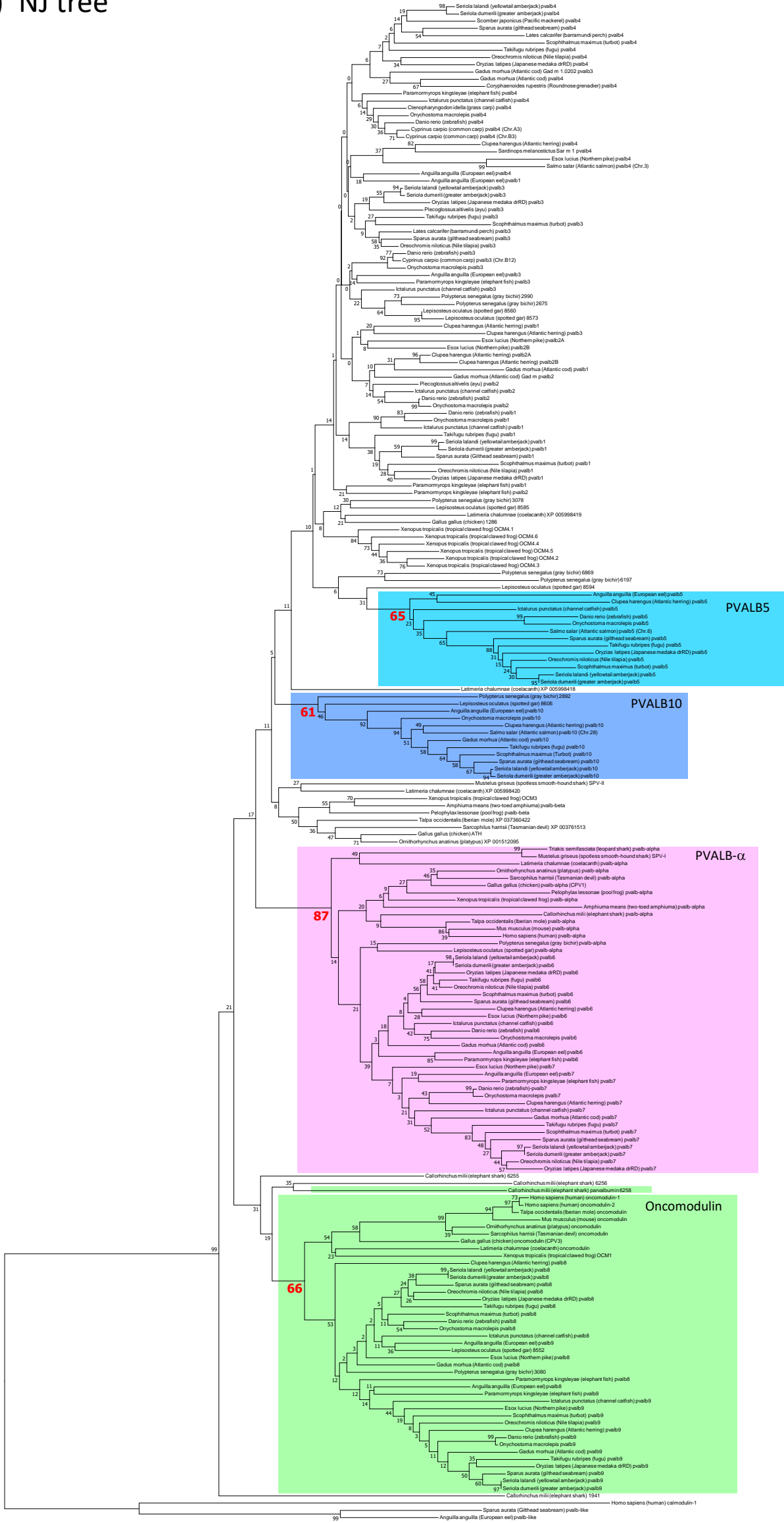
References

1. Jones D.T., Taylor W.R., and Thornton J.M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8: 275-282.
2. Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870-1874.
3. Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.
4. Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.
5. Zuckerkandl E. and Pauling L. (1965). Evolutionary divergence and convergence in proteins. Edited in *Evolving Genes and Proteins* by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York.
6. Rzhetsky A. and Nei M. (1992). A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution* 9:945-967.
7. Nei M. and Kumar S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
8. Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.
9. Sneath P.H.A. and Sokal R.R. (1973). *Numerical Taxonomy*. Freeman, San Francisco.

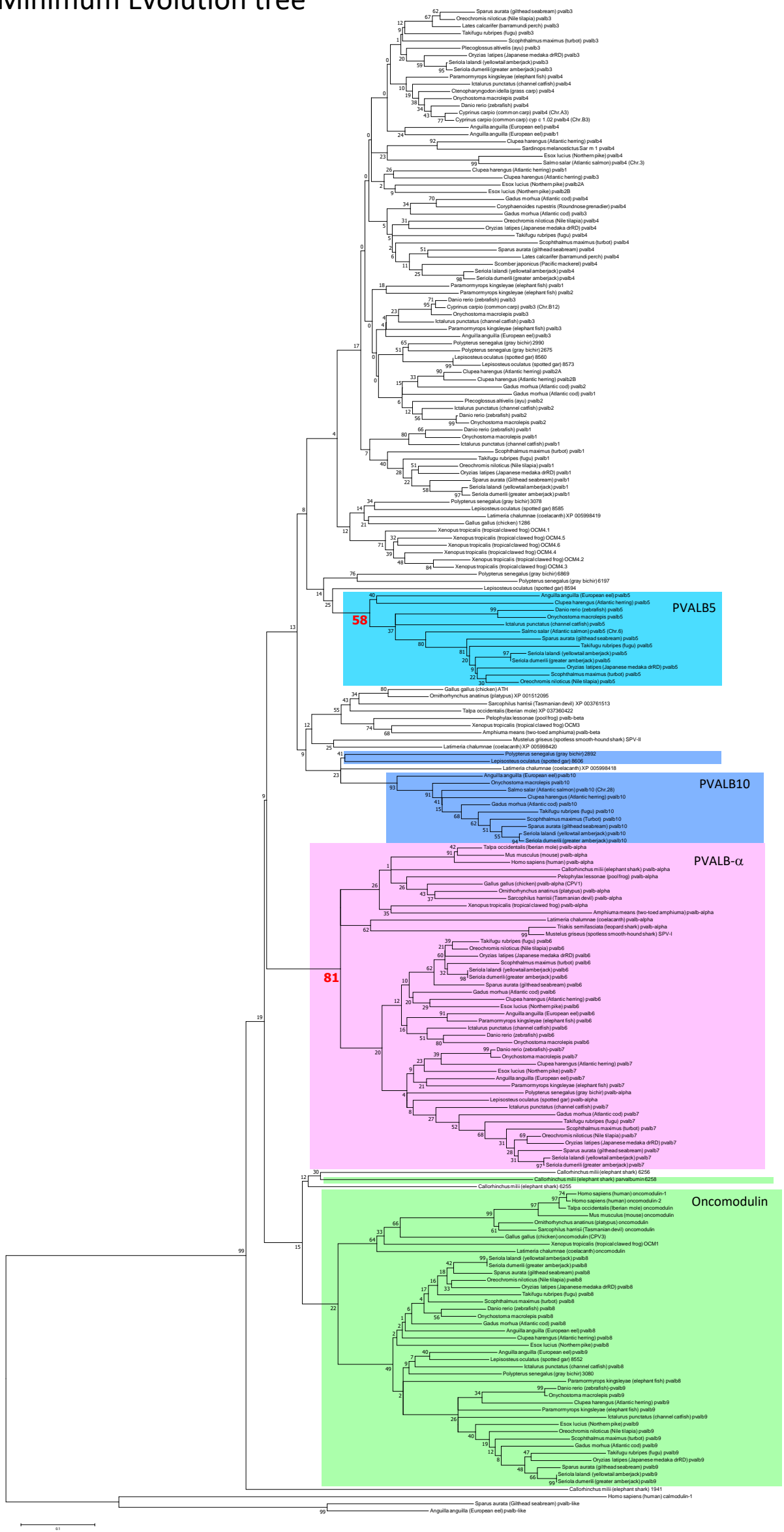
(A) ML tree



(B) NJ tree



(C) Minimum Evolution tree



(D) UPGMA tree

