

Transcriptome Sequencing of 100 Fishes from the Yangtze River

Yu Huang^{1,2,3}, Guillermo Ortí³, Lily Hughes³, Xinhui Zhang¹, Min Wang¹, Qiong Shi^{1,2}

BGI Education Center, University of Chinese Academy of Science, Shenzhen 518083, China

BGI Academy of Marine Sciences, BGI Marine, Shenzhen 518083, China

Department of Biological Sciences, George Washington University, Washington, DC 20052, USA

SEQUENCING FISHES IN THE YANGTZE RIVER

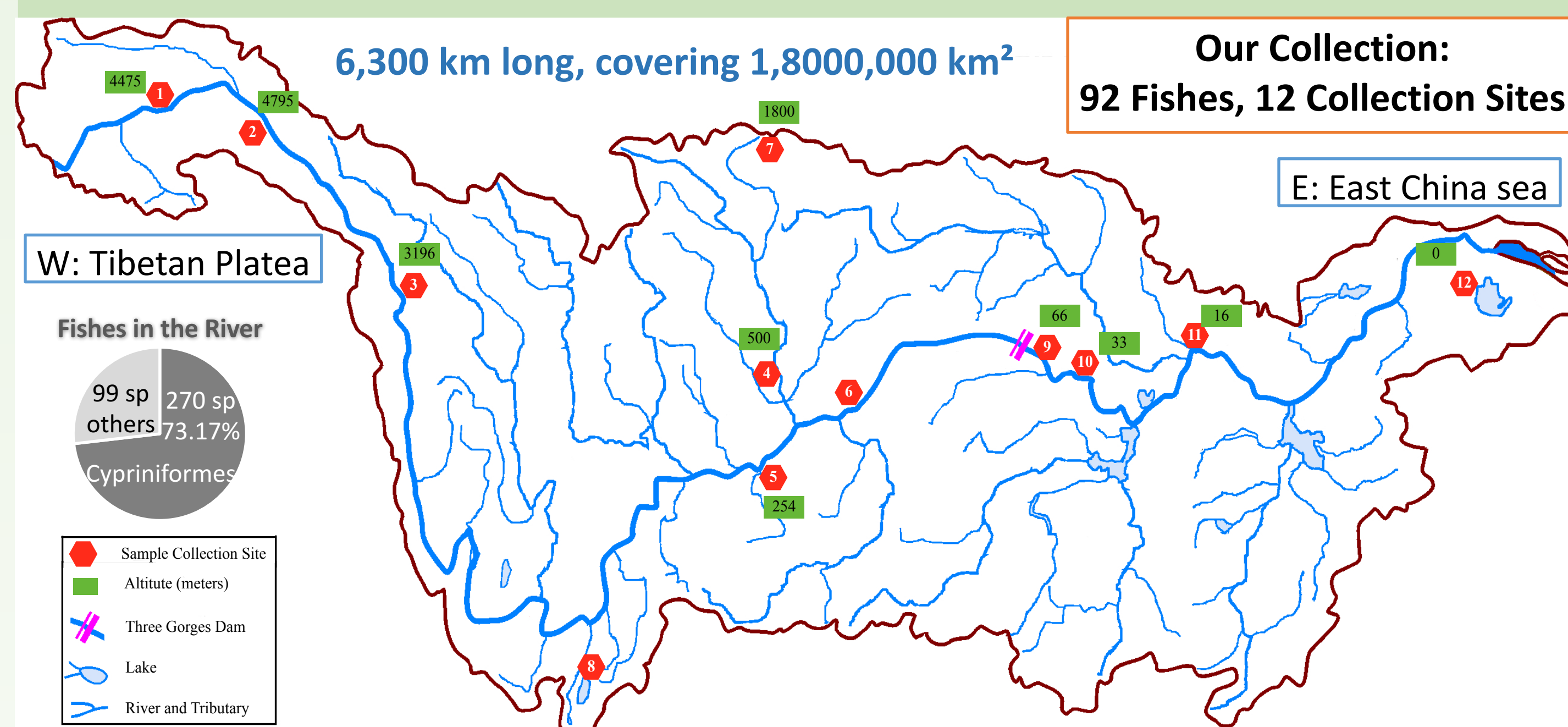


Figure 1. The Yangtze River Basin and sample collection sites in this study.

Sample Collection (Figure 1 & 2)

- **Dissection:** multiple types of tissues (muscle, gill, eye, liver, heart and brain)
- **Species identification:** morphology and DNA barcoding
- **Shipping:** snap frozen in liquid nitrogen
- **Stored:** ultra-low refrigerator (-80 °C)

Transcriptome Sequencing

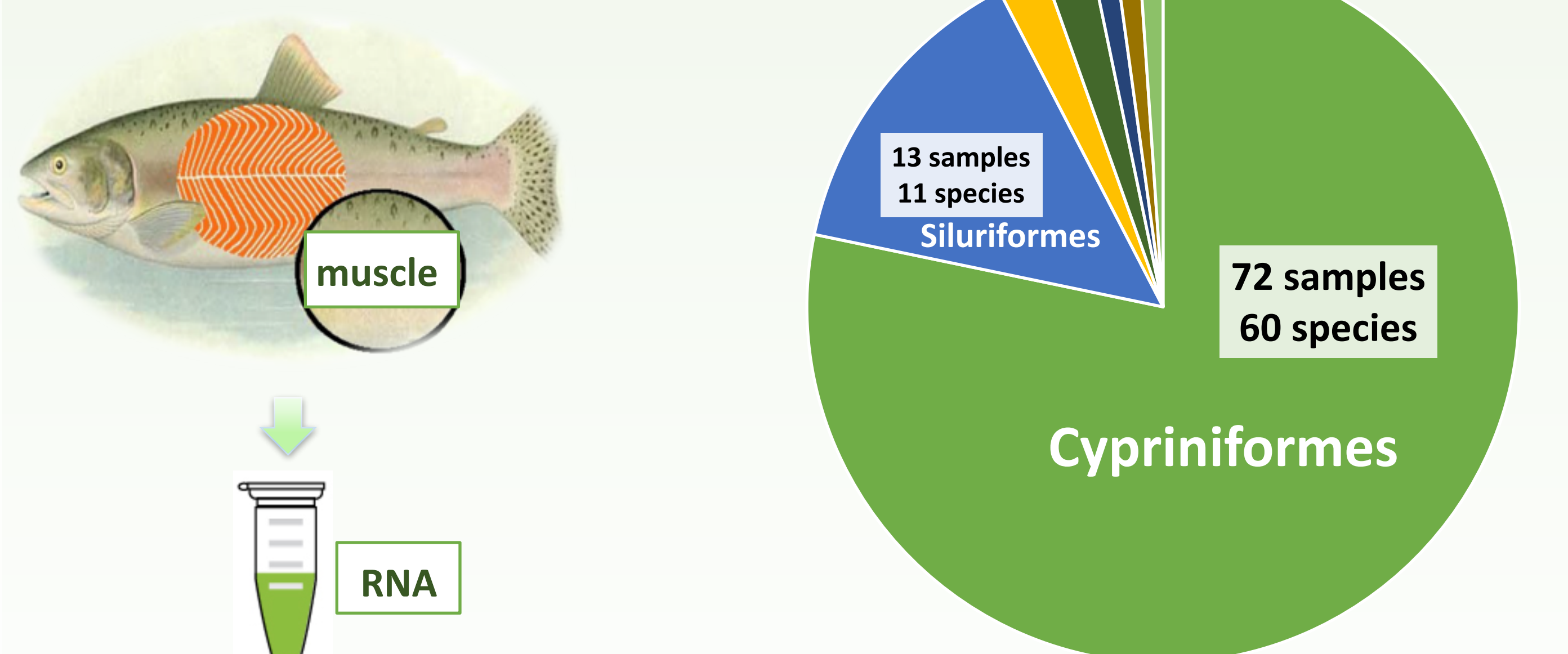
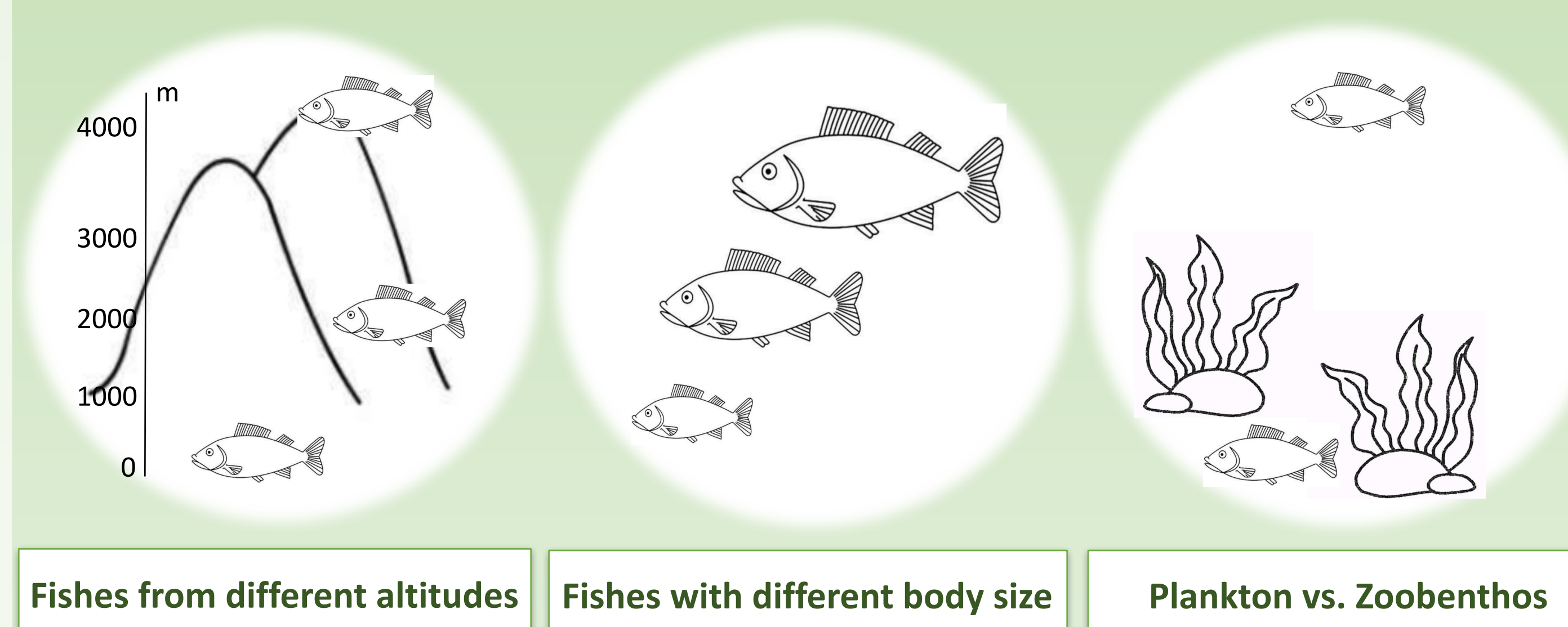


Figure 2. Classification of the collected samples

Questions to be Addressed



RESOLVING CYPRINIFORMES PHYLOGENY

Questions

- Test those ambiguous hypothesis and unsettled nodes in the current phylogeny^[1]
- Distinguish orthologs vs. paralogs (lineage-specific WGD in several taxa of Cypriniformes^[2, 3])
- The origin of Cypriniformes (historical biogeography)

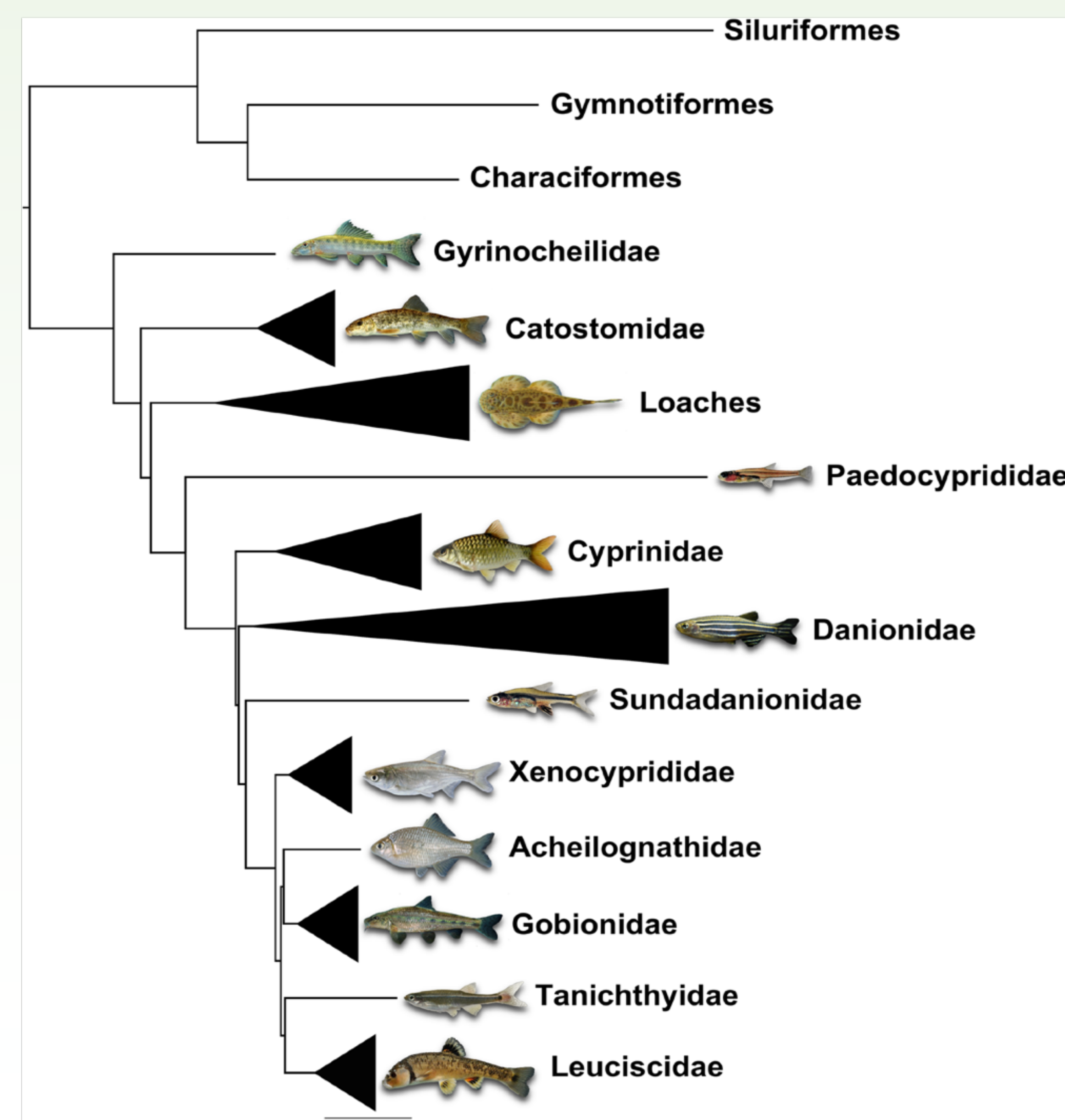
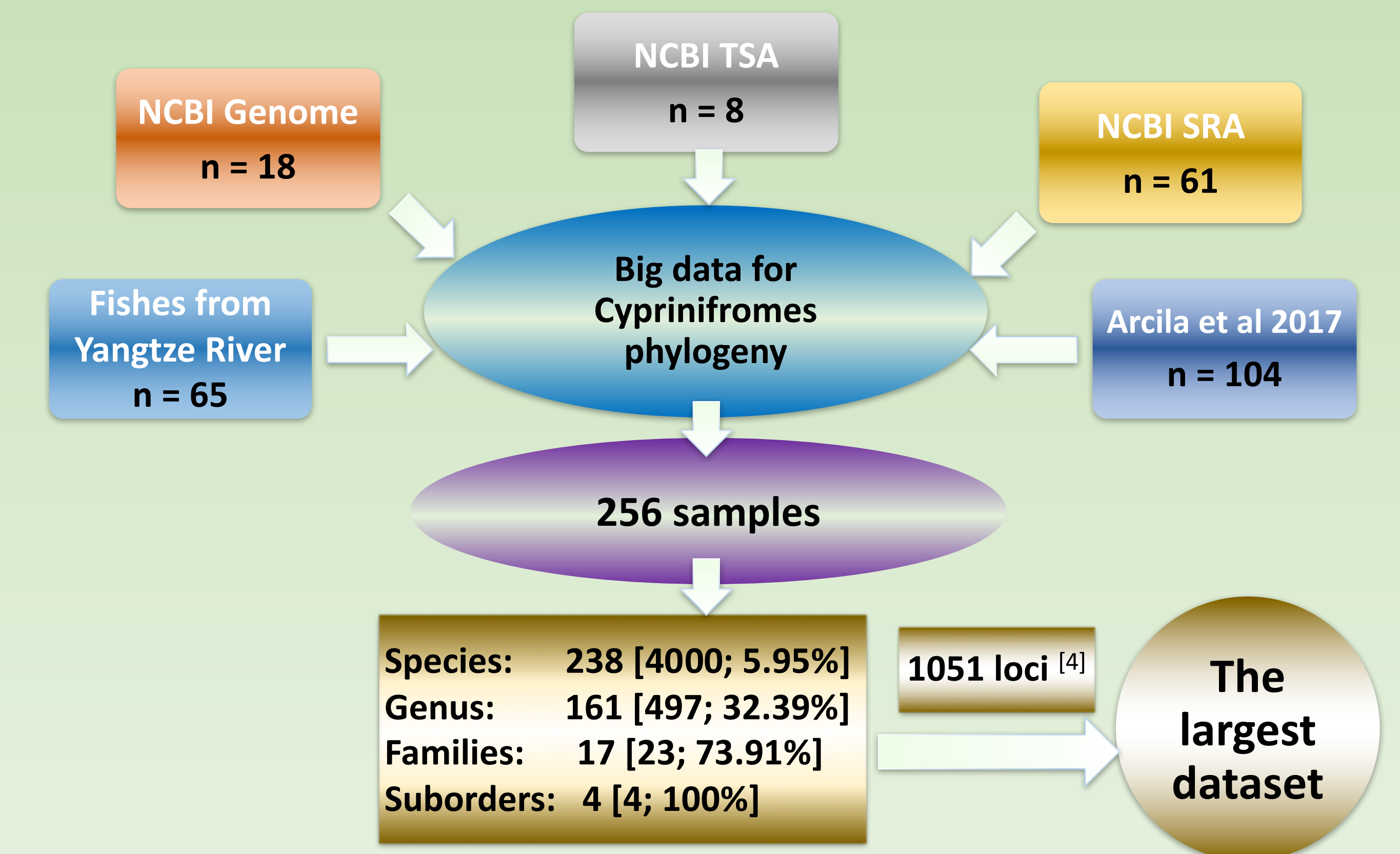


Figure 3. Phylogeny of Cypriniformes (Stout, et al. 2016).

Data Curation for Cypriniformes Phylogeny



Fish-T1K PROJECT

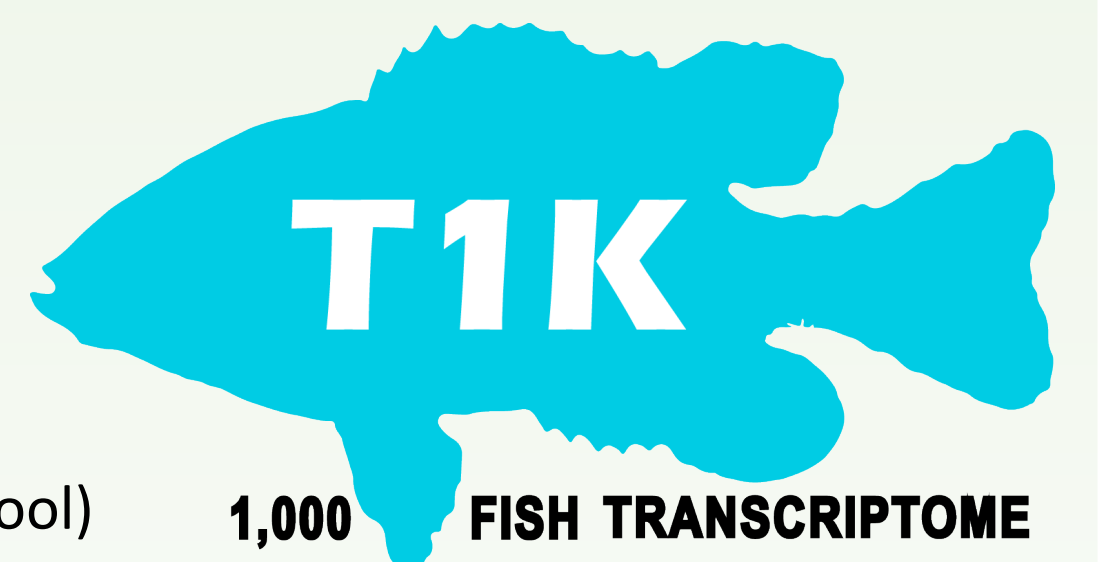
Fish-T1K^[5] is an collaborative and non-profit initiative to www.fisht1k.org

- generate transcriptome sequences for 1,000 diverse fish species
- establish a large-scale fish transcriptome database
- establish SOPs for genome quality sampling of fishes
- develop new tools for data analysis and management
- foster extensive global collaboration of fish omics study



Current Progress

- ~ 8,000 genome-level frozen fish tissues (stored in -80°C)
- available for genome and transcriptome sequencing
- ~ 400 transcriptomes (average data size > 3.5 Gb)
- data are deposited in www.fisht1k.org (provide with a BLAST tool)
- Eg: Our data has helped to construct the comprehensive phylogeny of ray-finned fishes^[6]



Fish-T1K
Yu Huang fiht1k@genomics.cn
huangyu@genomics.cn

ACKNOWLEDGEMENTS



THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

华大基因
BGI

China National GeneBank
CNGB 国家基因库

REFERENCES

1. Carla Stout, Milton Tan, Alan Lemmon, et al. Resolving Cypriniformes relationships using an anchored enrichment approach. *BMC Evol Biol*, 2016, 16:244.
2. Peng Xu, Xiaofeng Zhang, Xumin Wang, et al. Genome sequence and genetic diversity of the common carp, *Cyprinus carpio*. *Nat Genet*, 2014, 46: 1212-1219.
3. M.Vincent Hirt, Gloria Arratia, Weijin Chen, et al. Effects of gene choice, base composition and rate heterogeneity on inference and estimates of divergence times in cypriniform fishes. *Biol J Linn Soc*, 2017, 121: 319-339.
4. Dahiana Arcila, Guillermo Ortí, Richard Vari, et al. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nat Ecol Evol*, 2017, 1:0020.
5. Ying Sun, Yu Huang, Xiaofeng Li, et al. Fish-T1K (Transcriptomes of 1,000 Fishes) Project: large-scale transcriptome data for fish evolution studies. *GigaScience*, 2016, 5:18.
6. Lily Hughes, Guillermo Ortí, Yu Huang, et al. Comprehensive phylogeny of ray-finned fishes (Actinopterygii) based on transcriptomic and genomic data. *PNAS*, 2018, 115(24):6249-6254.