

Mar 24, 2021

# Protocols for "Chromosome-level genome assembly of the humpback puffer, *Tetraodon palembangensis*"

Rui Zhang<sup>1</sup>, Chang Li<sup>1</sup>, Mengjun Yu<sup>1</sup>, Xiaoyun Huang<sup>1</sup>, Mengqi Zhang<sup>1</sup>, Shanshan Liu<sup>1</sup>, Shanshan Pan<sup>1</sup>, Weizhen Xue<sup>1</sup>, Congyan Wang<sup>1</sup>, Chunyan Mao<sup>1</sup>, He Zhang<sup>2</sup>, Guangyi Fan<sup>1</sup>

<sup>1</sup>BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China;

<sup>2</sup>BGI-Qingdao, BGI-Shenzhen, Qingdao 266555, China. Department of Biology, Hong Kong Baptist University, Hong Kong, China

1 Works for me [dx.doi.org/10.17504/protocols.io.bs8inhue](https://dx.doi.org/10.17504/protocols.io.bs8inhue)

BGI GIGA 1 more workspace

zhangrui7

SUBMIT TO PLOS ONE

## ABSTRACT

In this study we assembled the first chromosome-level genome assembly of the humpback puffer (*Tetraodon palembangensis*), a species of poisonous freshwater pufferfish mainly distributed in Southeast Asia. This utilized Oxford Nanopore and BGISEQ-500 sequencing, stLFR libraries and Hi-C assembly, and the protocols for nucleic acid extraction, library construction and sequencing are all made available here. The resulting genome size was 362 Mb with a ~1.78 Mb contig N50 and ~15.8 Mb scaffold N50. Based on the assembled genome, ~61.5Mb (18.11%) repeat sequences were identified, 19,925 genes were annotated, and 90.01% of these genes could be predicted with function. Finally, a phylogenetic tree of ten teleost fish species was successfully constructed using this and public data. These protocols are made available for others carrying vertebrate genome assembly, and the open data and published study for comparative genome studies.

## DOI

[dx.doi.org/10.17504/protocols.io.bs8inhue](https://dx.doi.org/10.17504/protocols.io.bs8inhue)

## COLLECTION CITATION

Rui Zhang, Chang Li, Mengjun Yu, Xiaoyun Huang, Mengqi Zhang, Shanshan Liu, Shanshan Pan, Weizhen Xue, Congyan Wang, Chunyan Mao, He Zhang, Guangyi Fan 2021. Protocols for "Chromosome-level genome assembly of the humpback puffer, *Tetraodon palembangensis*". **protocols.io**  
<https://dx.doi.org/10.17504/protocols.io.bs8inhue>

## KEYWORDS

genomics, Hi-C, Nanopore, stLFR, Vertebrate genomes

## LICENSE

This is an open access collection distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

## CREATED

Mar 11, 2021

## LAST MODIFIED

Mar 24, 2021

#### OWNERSHIP HISTORY

Mar 11, 2021  Hongling Zhou

Mar 24, 2021  zhangrui7



#### COLLECTION INTEGER ID



48106



#### ABSTRACT



In this study we assembled the first chromosome-level genome assembly of the humpback puffer (*Tetraodon palembangensis*), a species of poisonous freshwater pufferfish mainly distributed in Southeast Asia. This utilized Oxford Nanopore and BGISEQ-500 sequencing, stLFR libraries and Hi-C assembly, and the protocols for nucleic acid extraction, library construction and sequencing are all made available here. The resulting genome size was 362 Mb with a ~1.78 Mb contig N50 and ~15.8 Mb scaffold N50. Based on the assembled genome, ~61.5Mb (18.11%) repeat sequences were identified, 19,925 genes were annotated, and 90.01% of these genes could be predicted with function. Finally, a phylogenetic tree of ten teleost fish species was successfully constructed using this and public data. These protocols are made available for others carrying vertebrate genome assembly, and the open data and published study for comparative genome studies.



## FILES



- 




Hi-C library preparation for the *Lateolabrax maculatus* genome  
**Version 1**  
 by Chang Li
- 




RNA isolation with TRIzol  
**Version 1**  
 by Magdalena Julkowska, King Abdullah University of Science and Technology
- 


BGISEQ-500 WGS library construction  
**Version 1**  
 by Xinming Liang, Beijing Genomics Institute
- 


The pipeline of Hi-C assembly  
**Version 1**  
 by Kailong Ma
- 


stLFR library construction  
**Version 1**  
 by lichunhua
- 


RNA and DNA extraction (Qiagen) of frozen tissue to nanopore sequencing  
**Version 1**  
 by Eric Pederson, Uppsala University
- 


BGISEQ-500 Sequencing  
**Version 1**  
 by Xinming Liang, Beijing Genomics Institute
- 


🔗 Oxford Nanopore sequencing and library construction  
**Version 1**  
 by zhangrui7