🌐 Repetitive element annotation protocol for the eastern banjo frog V.1

🦔 In 1 collection

Qiye Li[1,2], Qunfei Guo[1,3], Yang Zhou[1], Huishuang Tan[1,4], Terry Bertozzi[5,6], Yuanzhen Zhu[1,7], Ji Li[2,8], Stephen Donnellan[5], Guojie Zhang[2,8,9,10]

[1]BGI-Shenzhen, Shenzhen 518083, China;
[2]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650223, China;
[3]College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China;
[4]Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 611731, China;
[5]South Australian Museum, North Terrace, Adelaide 5000, Australia;
[6]School of Biological Sciences, University of Adelaide, North Terrace, Adelaide 5005, Australia;
[7]School of Basic Medicine, Qingdao University, Qingdao 266071, China;
[8]China National Genebank, BGI-Shenzhen, Shenzhen 518120, China;
[9]Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, 650223, Kunming, China;
[10]Section for Ecology and Evolution, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark

Qunfei Guo

VERSION 1

FEB 27, 2023

Protocol status: Working
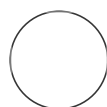
ABSTRACT

This pipeline is the repetitive element annotation by using homology-based, *de novo* predictions and Tandem Repeats Finder to identify repetitive elements in the *Limnodynastes dumerilii dumerilii* genome assembly.

1   **The following sub-steps are the process for homology-based predictions to identify repetitive elements in the _Limnodynastes dumerilii dumerilii_ genome assembly.**

1.1   Identify known repetitive elements by aligning the _Limnodynastes dumerilii dumerilii_ genome sequences against the Repbase-derived RepeatMasker libraries using RepeatMasker (v4.1.0).

> **Note**
>
> Use the parameters "-nolow -norna -no is" when running RepeatMasker.

1.2   Identify known repetitive elements by aligning the _Limnodynastes dumerilii dumerilii_ genome sequences against the transposable element protein database using RepeatProteinMask (an application within the RepeatMasker package)

> **Note**
>
> Use the parameters "-noLowSimple -pvalue 0.0001 -engine ncbi" when running RepeatProteinMask.

2   **The following sub-steps are the process for _de novo_ predictions to identify repetitive elements in the _Limnodynastes dumerilii dumerilii_ genome assembly.**

2.1   Execute RepeatModeler (v2.0, RRID:SCR_015027) on the _Limnodynastes dumerilii dumerilii_ assembly to build a _de novo_ repeat library for this species.

2.2   Employ RepeatMasker (v4.1.0) to align the *Limnodynastes dumerilii dumerilii* genome sequences against the *de novo* library for repetitive element identification.

3   **Identify tandem repeats in the *Limnodynastes dumerilii dumerilii* genome assembly by Tandem Repeats Finder (v4.09).**

> **Note**
>
> Use the parameters "Match=2 Mismatch=7 Delta=7 PM=80 PI=10 Minscore=50 MaxPeriod=2000" when running Tandem Repeats Finder.