

EasyCGTree V4.2 使用说明（简洁版）

一、EasyCGTree V4.2 主要特点：

- 1、使用 HMMER 进行同源基因检索，MUSCLE5 进行序列比对，trimAl 筛选保守区，FastTree/IQ-TREE 构建系统发育树；
- 2、自带常用基因集的 profile HMM 文件，目前包括 bac120、ar122、rp1 和 rp2；
- 3、可使用个性化的 gene clusters 进行序列比对和构建 profile HMM 文件，后续可直接使用；
- 4、可构建 supermatrix 和 supertree 两种策略的进化树。

二、安装：安装好 perl 语言环境，解压软件包即可运行。

EasyCGTree 软件包和 HMM 数据分别下载，将 HMM 数据解压后“.hmm”文件放入 HMM 文件夹。

三、使用方法：

- 1、**数据准备：**将所需基因组的 CDS 氨基酸序列，或者 DNA 序列，或者二者混合（每株菌的序列类型须统一）的 FASTA 格式文件，放入一个文件夹，如“myGenome”，放入 EasyCGTree-WinV4 文件夹。
- 2、**运行：**假设 EasyCGTree 的主文件夹放在 D 盘，myGenome 放在 EasyCGTree-WinV4 文件夹内。打开 cmd 或终端，输入（每行结束按回车键）：

```
> D:
> cd ./EasyCGTree-WinV4
> perl EasyCGTree.pl -input myGenome
```

然后即可得到使用 bac120 基因集，trimAl 的“strict”严谨度筛选保守区，FastTree 构建的 supermatrix 系统发育树。

想使用其他基因集，在最后一行命令添加“-hmm ar122/rp1/rp2 或其他”即可（默认 bac120）。

运行太慢，想增加线程数，添加“-thread 数字”（默认 2）。

想使用 IQ-TREE 进行进化树构建，添加“-tree_app iqtree”（默认 fasttree）。

想构建 supertree 进化树，添加“-tree st”（默认 sm）。

想改变 trimAl 的严谨程度，添加“-trim gappyout/strictplus”（默认 strict）。

想修改 FastTree/IQ-TREE 的参数，可在“tree_app-options.txt”文件中进行。

更多参数，请输入“perl EasyCGTree.pl -h”查看。（**请注意参数和值之间要留空格**）

四、引用： Dao-Feng Zhang, Wei He, Zongze Shao, Iftikhar Ahmed, Yuqin Zhang, Wen-Jun Li & Zhe Zhao.

EasyCGTree: a pipeline for prokaryotic phylogenomic analysis based on core gene sets. BMC Bioinformatics 2023 24:390. <https://doi.org/10.1186/s12859-023-05527-2>。

五、下载地址：<https://gitee.com/zdf1987/EasyCGTree4/releases>

欢迎报告 bug 或提出宝贵建议（zdf@hhu.edu.cn，张道锋，河海大学）