

任务练习

一个简单的基因家族分析



基因课精心制作

www.genek.cn

基本思路

用拟南芥中的FAD4基因序列，到芝麻、大豆基因集中搜索相似的序列，并构建基因树

在哪做？

用什么做？

安装 blast

准备输入数据

拟南芥FAD4序列

atha_FAD4.fa

芝麻蛋白集合

sind.fasta

大豆蛋白集合

gmax.fasta

分析步骤

第一步：blastdb 构建

第二步：blastp 比对

query

atha_FAD4.fa

database

sind.fasta

gmax.fasta

all.fasta

第三步：比对过滤并提取符合条件的序列的ID

第四步：提取序列

seqtk

第五步：构建基因树

多序列比对

muscle

建树