基因家族鉴定:

利用 Pfam 数据库(http://pfam.sanger.ac.uk/) 下载模型(PF03106)

- Pineapple genomics database 下载菠萝基因组数据
- HMMER3.1 (文献为 3.0) 在菠萝的蛋白序列中进行搜索鉴定
- ~biosoft/hmmer-3.1/binaries/**hmmbuild** WRKY1.hmm PF03106_seed.txt

利用下载的数据建立模型,输出 WRKY1.hmm

~/biosoft/hmmer-3.1/binaries/**hmmsearch** --cut_tc --domtblout WRKY1.out WRKY1.hmm

Acomosus_321_v3.protein.fa

利用模型在蛋白序列中搜索 WRKY 基因家族的蛋白,输出 WRKY.out

筛选出需要的 ID

~/biosoft/seqtk/seqtk subseq Athaliana_447_Araport11.protein.fa WRKY_At1_id1.txt >

WRKY_At1.qua.fa

根据提取的 ID 在 fa 文件中筛选出需要的序列

基因家族结构和保守 motif 分析:

- ①、进行多序列比对,建立进化树
- 1.打开 ClustW (https://www.genome.jp/tools-bin/clustalw)
- 2.点击"选择文件"
- 3.可以设置参数
- 4.点击"开始多序列比对"

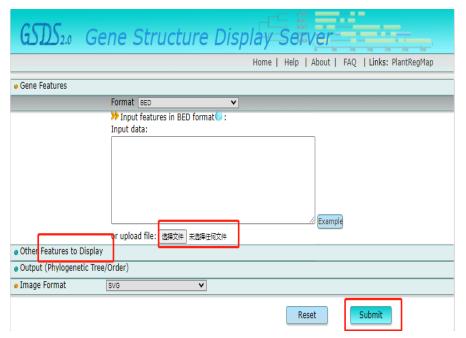


Multiple Sequence Alignment by CLUSTALW

ETE3	MAFFT	CLUSTALW	PRRN
			Help
General Setting Para			
Output Format:			
Pairwise Alignment: FAST/APPROXIMATE SLOW/ACCURATE			
Enter your sequences (with labels) below (copy & paste): PROTEIN O DNA			
Support Formats: FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF			
Or give the file name containing your query			
选择文件 未选择任何文			
ALIGNOSTI PROGRAMA	211		
Execute Multiple Align	nment Reset		
	More Deta	il Parameters	
Pairwise Alignment Parameters:			
For FAST/APPE	OVIMATE.		
K-tuple(wo		ize: 5 , Gap Penalty: 3	\neg
		oring Method: PERCENT V	_
Nulliber of	Top Diagonais: 3 , Sc	oring Method: PERCENT	
For SLOW/ACC	URATE:		
Gap Open	Penalty: 10.0 , Gap Exten	sion Penalty: 0.1	
Select Weig	ht Matrix: BLOSUM (for P	ROTEIN) V	
(Note that only p	arameters for the algorithm	n specified by the above "Pairw	vise Alignment" are
Multiple Alignment I	arameters:		
Gap Open I	Penalty: 10 , Gap Exten	sion Penalty: 0.05	
Weight Tra	nsition: OYES (Value: 0.	5), NO	

②、构建基因结构

- 1.打开 Gene Structure Display Server (GSDS: http://gsds.cbi.pku.edu.cn)
- 2.选择文件格式;
- 3.选择文件
- 4.可以添加另外想要展示的结构
- 5.点击"提交"



- ③、鉴定保守的 motifs
- 1.打开 MEME (http://meme-suite.org/tools/meme)
- 2.点击"选择文件"
- 3.设置参数
- 4.点击"开始搜索"

