**生物信息学上机实验**

**实验4 蛋白质序列特征分析**

**实验目的：**

掌握蛋白质序列检索的操作方法，熟悉蛋白质基本性质分析，了解蛋白质结构分析和预测。

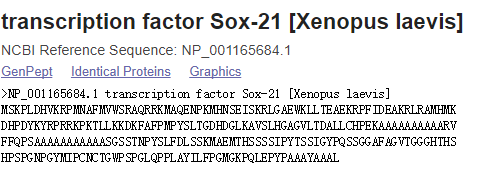
**实验内容：**

1. 通过NCBI生物信息网站检索人（Homo sapiens）、爪蟾（Xenopus laevis）和猕猴（Macaca mulatta）的转录因子SOX-21蛋白质序列，并保存其FASTA格式。

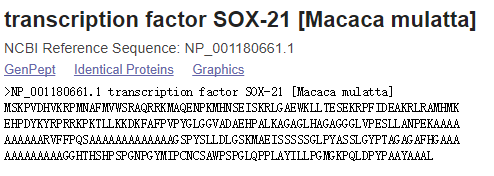
答：Homo sapiens：



Xenopus laevis：

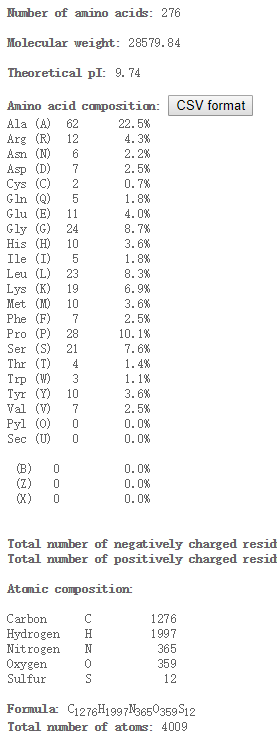


Macaca mulatta：

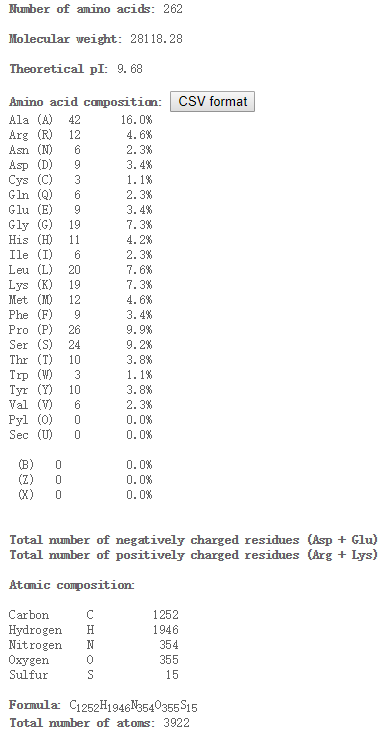


1. 利用ProParam工具https://web.expasy.org/protparam/进行以上三个物种的蛋白质序列的氨基酸组成、分子质量、等电点、氨基酸组成、原子总数等理化性质的分析，并比较三者之间的差异。

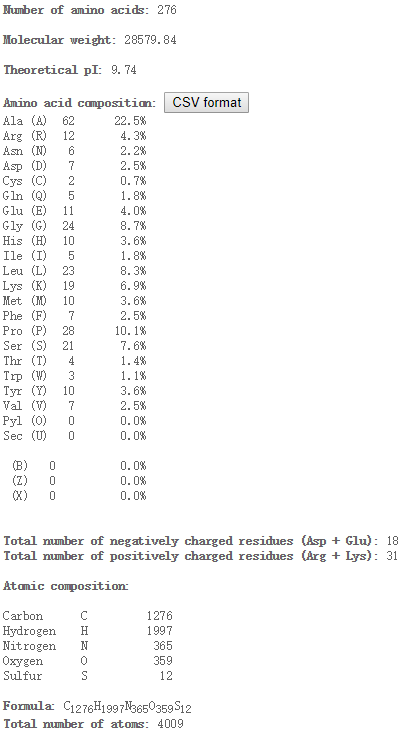
答：Homo sapiens：



Xenopus laevis：



Macaca mulatta：

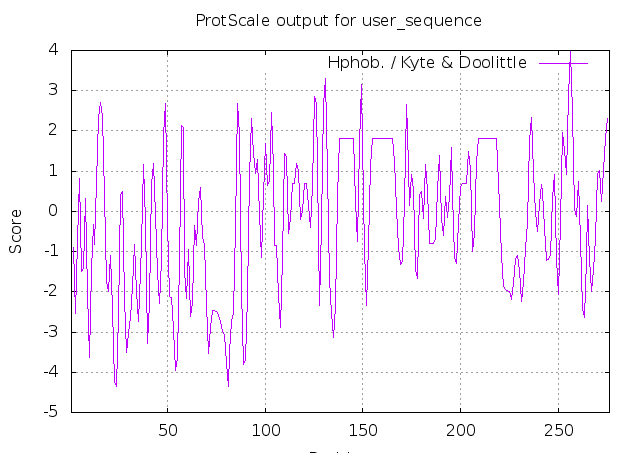


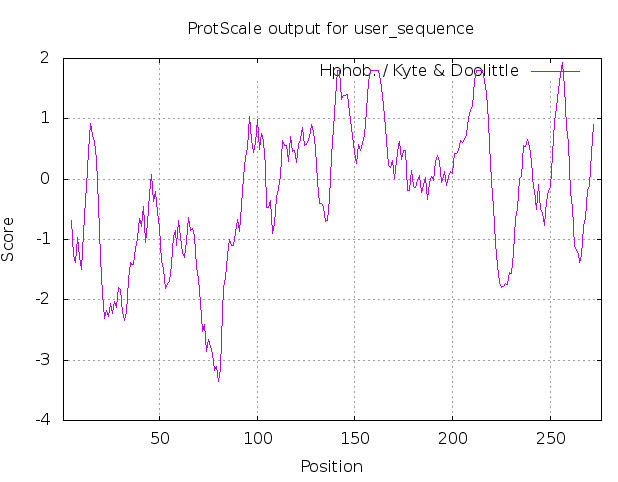
差异比较：人（Homo sapiens）与猕猴（Macaca mulatta）转录因子SOX-21蛋白质序列完全相同，且与爪蟾（Xenopus laevis）的序列仅在Ala(A)氨基酸含量上存在明显差异，整体上极为相似。说明该序列在进化过程中极具保守性。

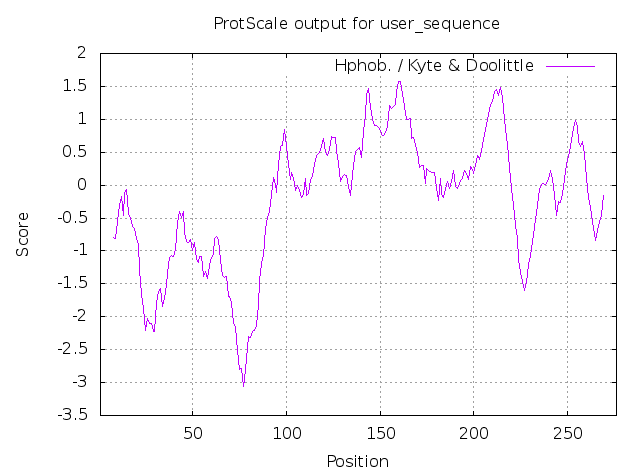
1. 利用ProtScale工具https://web.expasy.org/protscale/对人（Homo sapiens）转录因子SOX-21蛋白质序列的疏水区进行分析。

参数设置：氨基酸标度：[Hphob. / Kyte & Doolittle](http://web.expasy.org/protscale/pscale/Hphob.Doolittle.html) 方法；窗口尺寸：3，9，15。权重：50%；权重变化：linear；是否归一化：No。

答：

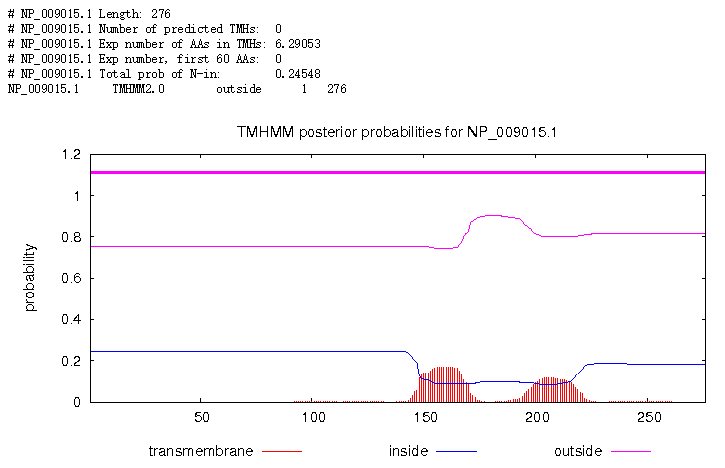
Window size：3

Window size：9

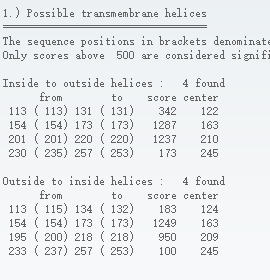
Window size：15

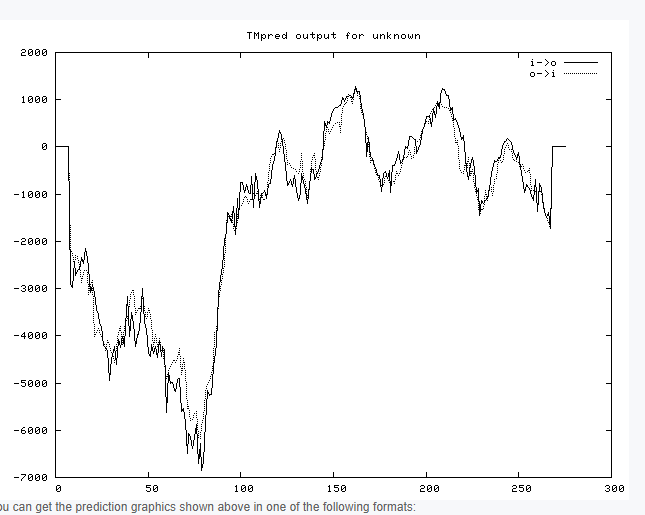
1. 利用TMHMM、TMPRED等工具对人（Homo sapiens）转录因子SOX-21蛋白质进行跨膜分析，并比较两个工具分析结果的差异。

答：TMHMM：



TMPRED：





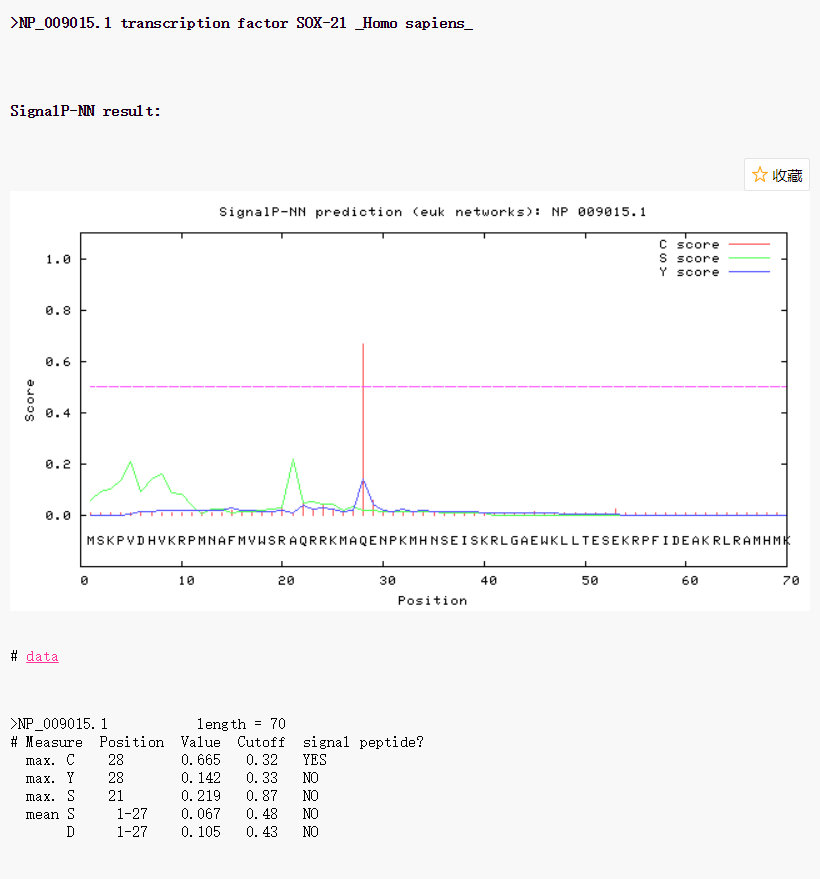
**分析：**TMHMM分析在140~180,190~220处有可能俩个跨膜域

TMPRED分析在120~125,140~170,200~220,230~250处可能有四个跨膜域。

但俩工具均显示跨膜域存在的可能性较低。

1. 利用SignalIP工具分析蛋白质序列的信号肽及其剪切位点。

参数设置：算法模型：Neural networks；其他为默认。

答：