

宁波大学 2018 年博士研究生招生考试初试试题(B 卷)

(答案必须写在考点提供的答题纸上)

科目代码： 3824 科目名称： 生物信息学

一、名词解释（每小题 4 分，共 20 分）。

1. 序列比对
2. 同源性
3. 相似性打分矩阵
4. 表达序列标签
5. 转录组学

二、选择题（每小题 2 分，共 20 分）。

1. 分别以疾病和基因为中心，阐述遗传变异介导的疾病（表型）相关基因情况，以及变异介导的基因参与不同疾病（表型）情况的数据库是_____。
A. dbSNP B. KEGG C. OMIM D. Refseq
2. BLAST 搜索中-e 的含义是什么？
A. 在随机的情况下，其它序列与目标序列相似度要大于 S 的可能性
B. 在随机的情况下，其它序列与目标序列相似度要小于 S 的可能性
C. 其它序列与目标序列的相似性分值
D. 其它序列与目标序列的匹配长度
3. 当前蛋白质三级结构预测的主要方法有_____。
A. 神经网络法 B. 从头预测法 C. 比较建模法 D. 折叠识别法
4. 蛋白质信号肽的预测工具有_____。
A. nnpredict B. PredictProtein C. SingalD D. SingalP
5. 下面哪个不属于基因本体（Gene Ontology）数据库中对基因功能描述的分类？
A. 生物过程 B. 细胞组分 C. 基因通路 D. 分子功能
6. ChIP-seq 主要用于检测_____。
A. RNA 表达 B. DNA 突变 C. DNA-蛋白相互作用 D. DNA-DNA 相互作用
7. 下面哪个软件不是应用于短序列比对基因组的？
A. Tophat B. MEGA C. BWA D. Bowtie

宁波大学 2018 年博士研究生招生考试初试试题(B 卷)

(答案必须写在考点提供的答题纸上)

科目代码： 3824 科目名称： 生物信息学

8. 如果要检测全基因组长非编码 RNA 的表达，宜选择何种高通量技术_____。

- A. 去除 rRNA 建库方式的 RNA-seq B. 去除带 polyA 建库方式的 RNA-seq
C. 小 RNA 测序 D. 基因芯片

9. CLIP-seq 主要用于检测_____。

- A. RNA-RNA 相互作用 B. RNA-DNA 相互作用 C. RNA-蛋白相互作用
D. DNA-蛋白相互作用

10. 外显子测序主要用于检测_____。

- A. 外显子 RNA 表达 B. 外显子 DNA 突变、插入及缺失
C. 外显子与内含子的剪切 D. 外显子区域的组蛋白修饰

三、简答题（每小题 7 分，共 35 分）。

1. 在基因组注释中，为什么与蛋白编码基因预测相比，RNA 基因的预测更为困难？
2. 简述至少 3 种差异表达基因的筛选方法。
3. 简述蛋白编码基因功能预测的方法。
4. 从头开始预测蛋白编码基因主要包含哪两个步骤？
5. 精准医学的定义是什么？其重要基础是什么？

四、综合分析题（25 分）。

2016年，美国哈佛大学和麻省理工学院的Broad研究所Itay Tirosh等在Nature Letter发表题为“Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma”的文章，该研究通过对来自6个*IDH1*或*IDH2*突变的人少突胶质细胞瘤的4347个单细胞进行RNA-seq分析，发现人少突胶质细胞瘤细胞可以分为两大类，一类呈分化状态，而另一类较少的细胞则未分化、且与神经干细胞相关。请回答以下问题：

- 1) 什么是单细胞测序技术？（4分）
- 2) 如何基于单细胞转录组测序数据获取基因的表达值？（8分）
- 3) 基于单细胞测序的结果，可以采用何种统计学方法对单细胞进行分类？（5分）
- 4) 怎样判断其中一类细胞与神经干细胞相关？（8分）