

中国科学院北京基因组研究所

2007 年招收攻读硕士学位研究生入学统一考试专业试题

科目名称：生物信息学

科目代码：937

考试时间：2007 年 1 月 21 日下午 2: 00-5: 00

考生须知：

1. 本试卷满分为 150 分，全部考试时间总计 180 分钟。
2. 所有答案必须写在答题纸上，写在试题纸上或草稿纸上一律无效。

一、简要说明以下数据库存储的主要内容及其特色：（共 10 小题，每题 5 分，共 50 分）

- 1) PubMed,
- 2) SWISS-PROT,
- 3) InterPro,
- 4) OMIM,
- 5) TRANSFAC,
- 6) RepBase,
- 7) UTRdb,
- 8) ICN,
- 9) TrEMBL,
- 10) PFAM,

二、简要说明以下算法的用途：（共 4 小题，每题 5 分，共 20 分）

- 1) Weight matrix,
- 2) Baum-Welch,
- 3) progressive alignment,
- 4) Smith-Waterman,

三、简要说明以下软件的用途：（共 6 小题，每题 5 分，共 30 分）

- 1) SRS,
- 2) BLAT,
- 3) PAML,
- 4) CAP3,
- 5) FgeneSH,
- 6) HMMER,

四、简要介绍隐马模型，并用它描述基因编码区识别问题。（25 分）

五、写出 Viterbi 算法的递推公式，并举例说明其用途。（25 分）