

天津医科大学理论课教案首页

(共 4 页、第 1 页)

课程名称：生物信息学	课程内容/章节：基因识别与 RNA 序列分析
教师姓名：伊现富	职称：讲师
授课对象：生物医学工程学院 2011 级生医班（本）	教学日期：2014 年 10 月 13 日 13:30-15:20
授课方式：理论讲授	学时数：2
	教材版本：生物信息学（自编教材）

教学目的与要求（分掌握、熟悉、了解、自学四个层次）：

- 掌握基因识别的方法；mRNA 选择性剪接的主要机制；miRNA 预测和 miRNA 靶基因预测的方法。
- 熟悉原核基因和真核基因的结构特点；miRNA 的特点、生成过程和作用方式；miRNA 的相关数据库和分析工具。
- 了解基因识别的分析工具；选择性剪接的相关数据库与分析工具。
- 自学基因识别分析工具的使用方法。

授课内容及学时分配：

- (5') 回顾与导入：回顾序列基本信息和特征信息分析的主要内容，引出基因识别的内容；总结 RNA 的主要类别，引出 RNA 分析的内容。
- (45') 基因识别：介绍基因和基因识别的基本概念，介绍原核基因和真核基因的结构特点并进行比较，讲解基因识别的主要方法与策略，介绍基因识别的常用工具。
- (20') mRNA 选择性剪接：介绍剪接和选择性剪接的基本概念，讲解 mRNA 选择性剪接的主要机制，介绍相关的数据库和分析工具。
- (25') miRNA 及其靶基因预测：介绍 miRNA 的特点、生成过程、作用方式和生物学功能，讲解 miRNA 预测和 miRNA 靶基因预测的主要方法，介绍常用的数据库和分析工具。
- (5') 总结与答疑：总结授课内容中的知识点，解答学生疑问。

教学重点、难点及解决策略：

- 重点：原核基因和真核基因的结构特点；基因识别方法；mRNA 选择性剪接的主要机制。
- 难点：基因识别中“信号”特征和“内容”特征的区别；mRNA 选择性剪接各种机制之间的区别。
- 解决策略：通过示意图和实例帮助学生理解，通过对比加深记忆。

专业外语词汇或术语：

基因识别 (gene prediction/finding)	选择性剪接 (alternative splicing)
间接识别法 (extrinsic approach)	微 RNA (miRNA, microRNA)
从头算法 (<i>ab initio</i> approach)	非编码 RNA (ncRNA, non-coding RNA)

辅助教学情况：

- 多媒体：原核基因和真核基因的结构，基因识别的策略，mRNA 选择性剪接的机制和实例，miRNA 的生成过程和作用方式。
- 板书：基因识别的方法与策略。

复习思考题：

- | | |
|-----------------------|--------------------------------|
| • 简述原核基因和真核基因结构的异同。 | • 简述 miRNA 特点、生成过程和作用方式。 |
| • 简述基因识别的三大类方法。 | • 简述 miRNA 预测和 miRNA 靶基因预测的方法。 |
| • 简述 mRNA 选择性剪接的主要机制。 | |

参考资料：

- 朱玉贤，李毅，郑晓峰。现代分子生物学（第 3 版），高等教育出版社，2007。
- 李霞，李亦学，廖飞。生物信息学，人民卫生出版社，2010。
- 王明怡，杨益，吴平。生物信息学（中译本，第 2 版），科学出版社，2004。
- 维基百科。

主任签字：

年 月 日

教务处制

一、回顾与导入 (5 分钟)

1. 序列分析

- DNA 序列分析：基本信息，序列特征，基因识别
- RNA 序列分析：mRNA 选择性剪接，miRNA 与靶基因

2. RNA 的分类 (RNA 既是携带遗传信息的主要生物大分子，也是重要的功能单位)

- 编码 RNA：mRNA
- 非编码 RNA：tRNA, rRNA; miRNA, siRNA, lncRNA

3. ncRNA 的分类 (转录后不编码蛋白质的 RNA 分子的统称)

- 基础结构性 ncRNA/看家 ncRNA：tRNA, rRNA, snRNA, snoRNA
- 调节性 ncRNA
 - sRNA：<200nt, miRNA、siRNA、piRNA (已经开展了广泛的研究)
 - lncRNA：>200nt, 长链非编码 RNA (引起关注，研究正逐步深入)

Non-coding RNA	Length (nt)	Species	Function
Ribosomal RNA (rRNA)	120~4700	All	Translation
Transfer RNA (tRNA)	70~100	All	Translation
Small nuclear RNA (snRNA)	70~350	Eukaryote	Splicing, mRNA processing
Small nucleolar RNA (snoRNA)	70~300	Eukaryote, archaea	RNA modification, rRNA processing
miRNA	21~25	Eukaryote	Translational regulation
siRNA	21~25	Eukaryote	Protection against viral infection
piRNA	24~30	Eukaryote	Genome stabilization
Long ncRNA	several hundreds~ several hundred thousands	Eukaryote	Transcription, splicing, transport regulation

二、基因识别 (45 分钟)

在介绍基本概念的基础上，通过比较原核基因和真核基因的异同，讲解基因识别的主要策略及各种方法在原核和真核基因识别中的具体应用。

1. 基本概念

- 基因：产生一条多肽链或功能 RNA 所需的全部核苷酸序列 (强调既包括编码区，也包括非编码区)
- 基因识别：识别 DNA 序列上具有生物学特征的片段



2. 【重点】基因结构 (通过示意图形象化展示、比较原核和真核的基因结构)

- 共同点：都包括编码区和非编码区
- 原核基因：连续基因
- 真核基因：不连续性



3. 【重点】识别方法

- 间接识别法：mRNA/蛋白质序列 \Rightarrow DNA 序列
- 从头预测法：基因预测，基于“信号”和“内容”两类特征
- 比较基因组学的方法：比较相关物种的 DNA 序列

4. 【难点】基因预测

• “信号”和“内容”

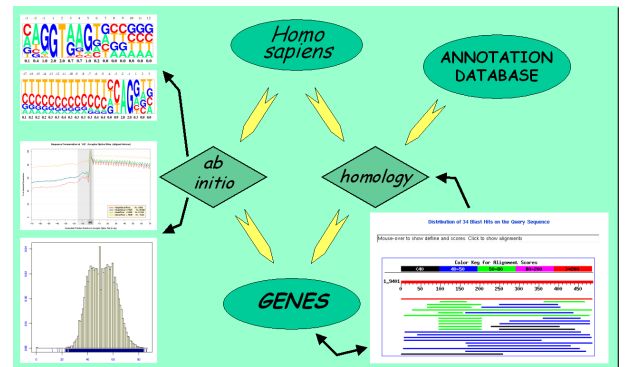
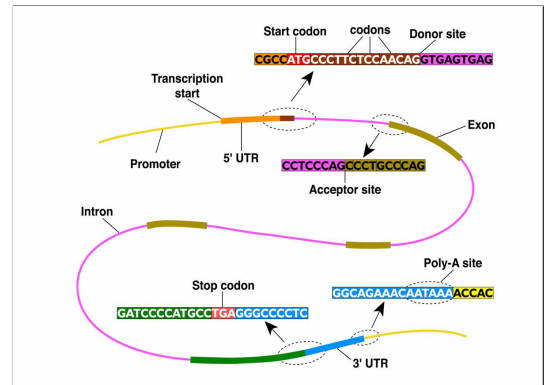
- 信号：不连续的局部序列模体，一般都有一致性序列；如启动子，剪接供体和受体位点，起始和终止密码子，polyA 位点
- 内容：不同长度的扩展序列，没有一致性序列，但具有把自己与周围 DNA 区分开来的保守特征；如密码子使用偏好性，双联密码子出现频率，基因组等值区

• 原核基因

- 信号：启动子序列，转录因子结合位点
- 内容：连续的开放阅读框，统计学特征
- 总结：信号容易识别，内容容易判别，预测能达到相对较高的精度

• 真核基因

- 信号：启动子区特征序列，供体和受体位点，起始和终止密码子，polyA 序列；确定外显子的边界，识别编码区域
- 内容：密码子使用偏好性，双联密码子出现频率，基因组等值区；区分外显子、内含子和基因间区域
- 总结：信号复杂，内容难判别，预测相当有挑战性；联合信号和内容检测以及同源性搜索，提高识别效率



5. 识别策略

6. 识别工具 (强调分析工具的适用范围)

- 识别原核基因：GeneMarkS, Glimmer
- 识别真核基因：GENSCAN

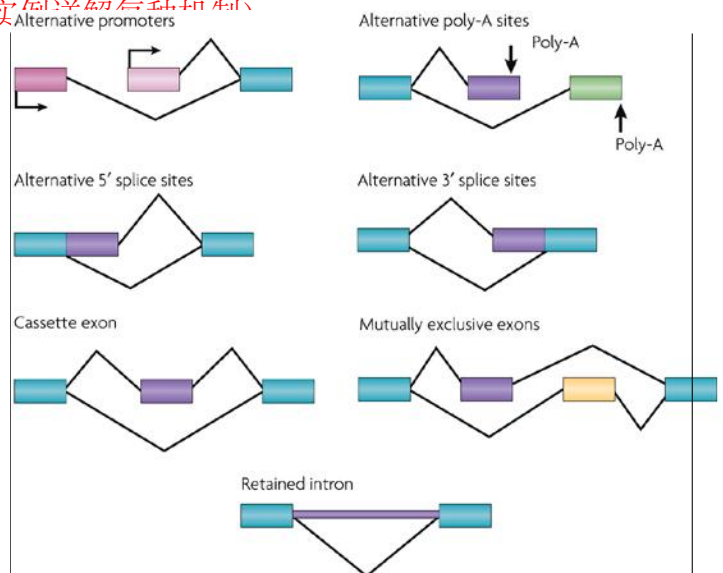
三、mRNA 选择性剪接 (20 分钟)

1. 基本概念

- 剪接：移除内含子、合并外显子
- 选择性剪接：一个 mRNA 前体 \Rightarrow 不同 mRNA 剪接异构体

2. 【重点、难点】主要机制 (通过示意图和实验验证)

- 外显子跳跃：外显子被移除或保留
- 互斥外显子：两个外显子只有一个保留下来
- 5' 选择性剪接：使用不同的 5' 端的供体位点
- 3' 选择性剪接：使用不同的 3' 端的受体位点
- 内含子保留：内含子作为外显子保留下来
- 选择性起始：在不同的位点起始转录
- 选择性终止：使用不同的 polyA 位点



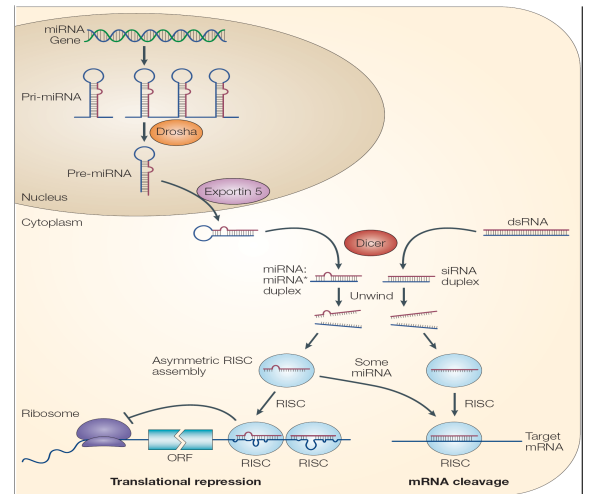
3. 相关资源 (提醒注意数据库的时效性)

- 数据库：ASTD, ASAP, ASPicDB
- 分析工具：ESEfinder, RESCUE-ESE

四、miRNA 及其靶基因预测 (25 分钟)

1. miRNA 简介 (通过示意图形象化展示相关内容)

- 真核生物中广泛存在的一种长约 20 到 24 个核苷酸的内源性非编码单链 RNA 分子
- 生成过程: 300 ~ 1000nt 的双链 pri-miRNA \Rightarrow 70 ~ 90nt 的双链具有茎环结构的 pre-miRNA \Rightarrow 20 ~ 24nt 的单链成熟 miRNA
- 作用方式: 完全互补型——结合在 mRNA 的编码区中, 导致靶基因 mRNA 降解, 在植物中比较常见; 不完全互补型——结合在 mRNA 的 3' UTR, 导致靶基因 mRNA 的翻译受到抑制
- 生物学功能: 调控个体发育、细胞分化、组织发育、肿瘤发生发展、……



2. miRNA 的特征

- 序列: 不具有开放阅读框, 不编码蛋白质; 成熟的 miRNA 5' 端为单一磷酸基团, 3' 端为羟基
- 表达: 具有时序性和组织特异性
- 调控: miRNA 与靶基因间呈多对多的关系
- 物理位置: 倾向于成簇地出现在染色体上
- 进化: 在物种间高度保守

3. miRNA 预测

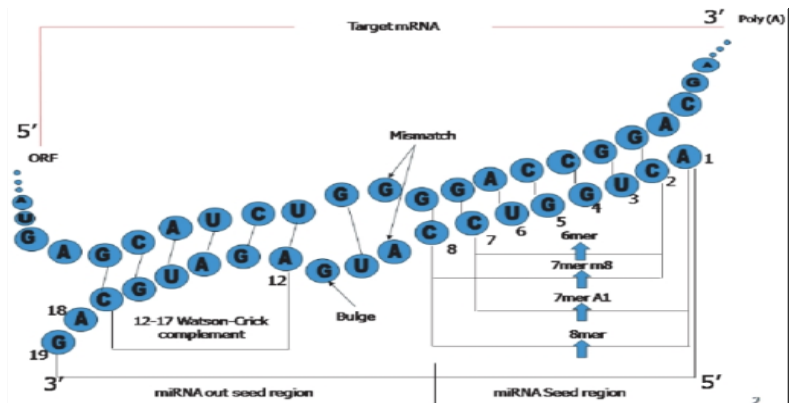
- 同源片段搜索
- 基于比较基因组学
- 基于序列和结构特征打分
- 结合作用靶标
- 基于机器学习

4. miRNA 靶基因预测

- 基于种子区域互补和保守性
- 基于机器学习

5. 相关资源

- 数据库: miRBase, TarBase, miRGen
- 分析工具: MiRscan, MiPred, miRfinder; miRanda, TargetScan, PicTar, miTarget



五、总结与答疑 (5 分钟)

- 基因识别: 原核和真核的基因结构, 基因识别方法
- mRNA 选择性剪接: 主要机制
- miRNA: 特征、生成过程和作用方式, miRNA 及其靶基因预测方法