

天津医科大学实验课教案首页

(共 3 页、第 1 页)

课程名称：分子生物计算		实验名称：实验 8 翻译 DNA 序列	
教师姓名：伊现富	职称：讲师	教学日期：2015 年 11 月 2 日 10:00-12:00	
授课对象：生物医学工程与技术学院 2013 级生信班（本）			实验人数：28
实验类型（验证型、综合型、设计型、创新型）：验证型			实验分组：一人一机
学时数：2		教材版本：Perl 语言在生物信息学中的应用——基础篇	

实验目的与要求：

- 了解阅读框的概念。
- 熟悉 FASTA 格式。
- 掌握：散列的使用；子程序的应用。

实验内容及学时分配：

- (100') 实验操作：编写 Perl 程序把 FASTA 文件中的 DNA 序列翻译成蛋白质。

主要仪器和实验材料：

- 主要仪器：一台安装有 Perl 语言（Linux 操作系统）的计算机。

实验重点、难点及解决策略：

- 重点难点：散列的使用。
- 解决策略：通过演示进行学习，通过练习熟练掌握。

思考题：

- 如何使用正则表达式表征密码子？
- 比较 Perl 语言中常见的三种数据类型。
- 举例说明常见的序列格式。
- 举例说明 6 种阅读框。

参考资料：

- Beginning Perl for Bioinformatics, James Tisdall, O'Reilly Media, 2001.
- Perl 语言入门（第六版），Randal L. Schwartz, brian d foy & Tom Phoenix 著，盛春 译，东南大学出版社，2012。
- Mastering Perl for Bioinformatics, James Tisdall, O'Reilly Media, 2003.
- 维基百科等网络资源。

主任签字：

年 月 日

教务处制

一、实验操作 (100 分钟)

1. 相关的子程序

- `codon2aa`: 把密码子翻译成氨基酸
- `dna2peptide`: 把 DNA 翻译成多肽
- `get_file_data`, `extract_sequence_from_fasta_data`: 从 FASTA 文件中提取序列数据
- `print_sequence`: 输出格式化的序列数据
- `revcom`: 获取反向互补序列
- `translate_frame`: 翻译 DNA 序列的一个阅读框

2. 把 DNA 序列翻译成蛋白质

```
#!/usr/bin/perl

use strict; use warnings; use BeginPerlBioinfo;

my $dna      = 'CGACGTCTTCGTACGGGACTAGCTCGTGTCGGTCGC';
my $protein  = '';
my $codon;

for ( my $i = 0 ; $i < ( length($dna) - 2 ) ; $i += 3 ) {
    $codon = substr( $dna, $i, 3 );
    $protein .= codon2aa($codon);
}
print "I translated the DNA\n\n$dna\n\n into the protein\n\n$protein\n\n";
```

3. 读取 FASTA 文件并提取其中的序列数据

```
#!/usr/bin/perl

use strict; use warnings; use BeginPerlBioinfo;

my @file_data = ();
my $dna       = '';

@file_data = get_file_data("sample.dna");
$dna = extract_sequence_from_fasta_data(@file_data);
print_sequence( $dna, 25 );
```

4. 把 FASTA 文件中的 DNA 翻译成蛋白质并进行格式化输出

```
#!/usr/bin/perl

use strict; use warnings; use BeginPerlBioinfo;

my @file_data = ();
my $dna       = '';
my $protein   = '';

@file_data = get_file_data("sample.dna");
$dna = extract_sequence_from_fasta_data(@file_data);
$protein = dna2peptide($dna);
print_sequence( $protein, 25 );
```

5. 从六个阅读框上翻译 DNA 序列

```
#!/usr/bin/perl

use strict; use warnings; use BeginPerlBioinfo;

my @file_data = ();
my $dna       = '';
my $revcom    = '';
my $protein   = '';

@file_data = get_file_data("sample.dna");
$dna = extract_sequence_from_fasta_data(@file_data);

print "\n -----Reading Frame 1-----\n\n";
$protein = translate_frame( $dna, 1 );
print_sequence( $protein, 70 );

print "\n -----Reading Frame 2-----\n\n";
$protein = translate_frame( $dna, 2 );
print_sequence( $protein, 70 );

print "\n -----Reading Frame 3-----\n\n";
$protein = translate_frame( $dna, 3 );
print_sequence( $protein, 70 );

$revcom = revcom($dna);

print "\n -----Reading Frame 4-----\n\n";
$protein = translate_frame( $revcom, 1 );
print_sequence( $protein, 70 );

print "\n -----Reading Frame 5-----\n\n";
$protein = translate_frame( $revcom, 2 );
print_sequence( $protein, 70 );

print "\n -----Reading Frame 6-----\n\n";
$protein = translate_frame( $revcom, 3 );
print_sequence( $protein, 70 );
```