

Assignment of Perl

- Write a program to print Fibonacci sequence. The length of output sequence is specified by the first command line parameter. (fibonacci.pl)
- Try to optimize the pos_annotate.pl as much as you can. (pos_annotateV3.pl)

All assignments should be sent to chengang@genomics.cn

Perl in Bioinformatics I

Gang Chen
chengang@genomics.cn

October 17, 2015

Outline

- 1 Perl Use Cases
- 2 Perl Modules
- 3 SNP Annotation

Next

1

Perl Use Cases

- System Administration
- CGI Programming

2

Perl Modules

3

SNP Annotation

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

Perl One Liners

Options

- -e

- -n

```
while (<>) {  
    # your code goes here  
}
```

- -p

```
while (<>) {  
    # your code goes here  
} continue {  
    print or die "-p destination: $!\n";  
}
```

Process file content

Adding Line Number to file content

```
perl -ne 'print "$. $_"' names.txt  
perl -pe '$_ = "$. $_"' names.txt  
perl -pe 'tr/a-z/A-Z/' names.txt  
perl -le 'print "PRIME" if (1 x shift) !~ /^(11+)  
\1+$/' 19
```

References

- Useful One-Line Scripts for Perl :

`http://www.catonmat.net/download/perl1line.txt`

Chinese: `https://github.com/vinian/perl1line.txt/blob/master/perl1line-ch.txt`



Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

CGI Programming

CGI

Common Gateway Interface (CGI) is a standard environment for web servers to interface with executable programs installed on a server that generate web pages dynamically.

Next

1 Perl Use Cases

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

Web Scraping in Perl

LWP

The libwww-perl collection is a set of Perl modules which provides a simple and consistent application programming interface (API) to the World-Wide Web.
<https://metacpan.org/pod/LWP>

Example: lwp.pl

Files

- pubmedids.txt: A list of pubmed ids
- lwp.pl: Get publication titles of these ids from PubMed.

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- **Management of Perl Modules**
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

Searching

www.CPAN.org

Installation

- cpan command: Linux, Mac OS and Windows(Strawberry)
- perl -MCPAN -e shell
- cpanm
- from source
- ppm: ActivePerl

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- **Development of Perl Modules**
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

Development

see PUBMED.pm and run.pl

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- **Submit your modules to CPAN**

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

Why?

- Easy the installation of your module
- Share your work with the community
- Increase your importance in the community
- Get response form the community to improve your module

How?

<http://www.cpan.org/modules/04pause.html>

Next

- 1 Perl Use Cases
- 2 Perl Modules
- 3 **SNP Annotation**
 - A Perl Script for SNP Annotation
 - SNPAnno module

SNP Annotation

Task

Given a list of genome positions, add corresponding gene symbol to each position.

Files

- Example input file: pos.txt
- Annotation database: refGene.txt
- Example script: pos_annotate.pl
- Optimized Script: pos_annotateV2.pl

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

pos_annotate.pl

pos_annotateV2.pl

Next

1 Perl Use Cases

- System Administration
- CGI Programming

2 Perl Modules

- Usage of Perl Modules
- Management of Perl Modules
- Development of Perl Modules
- Submit your modules to CPAN

3 SNP Annotation

- A Perl Script for SNP Annotation
- SNPAnno module

SNPAnno module

- SNPAnno.pm
- run_SNPAnno.pl

Thanks!