

Bioinformatics Programming 101 activity hand-out

0. Please “clone” the repository for the class programming by issuing this command: `git clone https://github.com/htruong/perlbio101` then `cd perlbio101`

1. First steps

- Try to run `hello-world.pl`.
- Variables are an useful concept in programming. Inspect `hello-person.pl` to see how they are used to make personalized greetings.
- `if` statements are used to test variables to make sure the variables being tested have certain properties. Try to run `hello-person-2.pl` and see what it does. Try to look at `hello-person-2.pl`. Modify it so it greets me (Huan), Gavin or you with the secret greetings (They are called **backdoors**, but don't be evil).
- You can manipulate variables too. Have a look at `hello-person-3.pl`. It automatically changes the first character of any name to uppercase :). But then it breaks down when the first character is already in uppercase! How would you fix it?
- Computers are good at doing things repeatedly. Modify `hello-world-100.pl` and `hello-world-while.pl` so that the program prints Hello 100 times with a `for` and with a `while` loop.

2. A fastq file reader

- A **fastq** file (`.fq`) is a text-based file for storing both a biological sequence (usually nucleotide sequence) and its corresponding quality scores. We often care more about the sequences but not the quality scores. We will build a program to convert the **fastq** to a **fasta** file that contains only the sequences by ‘trimming’ all the low-quality reads.
- Have a look at fastq on **Wikipedia** to understand the file format, and then open `lab1.fq` to inspect it.
- Derive a strategy to convert a **fastq** to **fasta** (Imagine you hire an assistant. How would you describe to her how to covert `lab1.fq` to `lab1.fas`).
- To start I have created some step-by-step problem-solving solution for a fastq reader. Look at `fastq-reader-1.pl` through `fastq-reader-7.pl` and solve each problem step-by-step. **If you already knew how to programming, write this program in your favorite language.**

3. A list comparer

- Every once in a while, there would be a person asking me that they have a problem that has been bothering them for days: They wanted Excel to find common things from two relatively messy, unformatted lists. We will try to come up with a program that reads two lists and tell us what are common among those two lists using arrays.