Programming 101 activity handout

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

1. Hello World

   - Try to run `hello-world.pl`
   - Variables are an useful concept in programming. Inspect `hello-person.pl` to see how it is used to make personalized greetings.
   - If statements are used to test variables. 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.
   - Computers are good at doing things repeatedly. Modify `hello-world-100.pl` and `hello-world-while.pl` so that it 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 *wikipedia* for the file format to see the format, and see lab1.fq to see what it is actually like.
   - Derive a strategy to convert a `fastq` to `fasta` (How would describe to a really dumb person how to do so?)
   - To start I have created some step-by-step problem-solving for a fastq reader. Look at `fastq-reader-1.pl` through `fastq-reader-7.pl` and solve each problem step-by-step.

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 genes from two relatively messy lists. We will try to come up with a program that reads two lists and tell us what is common in those two.