A quantitative analysis of the FIV and HIV genome using bioinformatics software

Meredith Horn, Charlotte Siu, and Xiao-Wen Cheng
Department of Microbiology, Miami University, Oxford, OH

special thanks is awarded to the Miami University’s CBG and Virginia tech for providing the FASTA and LALIGN software

methods

The DNA sequencing machine used is called Illumina MiSeq Benchtop next-generation sequencer, which follows the next-generation DNA sequencing method. Before the DNA of HIV and FIV could be sequenced, the DNA samples were needed to be amplified by PCR. This could increase the number of copies of the same DNA available for the sequencer to detect (Cacho et al., 2011).

If this experiment, the master mix of the PCR consists of 1 ul of big dye dilution buffer (table 1), 0.5 ul of the primer (table 1), 4.5 ul of the gag DNA/table (1), 5.5 ul of molecular grade water (table 1). This adds up to a total volume of 10ul. The same recipe was used for both the HIV and FIV gag. After preparing the master mix, we programmed the standard cycle sequencing protocol on the thermocycler. Stop one of the cycle lasted for 1 minute at 96 degree celsius. Stop 2 of the cycle lasted for 10 seconds at 96 degree celsius. Stop 3 of the cycle lasted 5 seconds at 50 degree celsius. Stop 4 of the cycle lasted 4 minutes at 60 degree celsius. At the end of step 4, the cycle would go back to step 2 and complete step 3 and 4 again. The cycle did this repeatedly for 24 hours. After completing the PCR run, the master mix was transferred to a 1.5ml eppendorf tube. 1ul of 1.5M NaOAc: EDTA and 80 ul 95% ethanol were added. All supernatants were removed with a pipette, and 100ul of 70% ethanol was added. The sample was stored at -20 degree celsius. After the purification and PCR process, the sample was handed over to a technician to run the next generation sequencing machine. After a couple of hours, the results were out.

The gag genome sequences received were “nucleotide Match” on NCBI. The results of the blast shows that the sequences were accurate because it shows a 100 percent match to the HIV and FIV gag genome sequences in their records. The HIV and FIV gag sequences obtained were entered into the FASTA software, a software designed to compare DNA sequences. The FASTA software computed the E value, bit score as well as the percent identity of the two sequences. The results were recorded and presented below.

Figure 1- the software indicates that there is a specific region of the DNA in HIV and FIV that requires attention, since the E value is particularly low which means it is very similar and it is unlikely that it is by chance.

results

The E value of the overall sequences is 3.1x10^-9, and a bit score of 51.9. The percent identity of the overall sequences is 54.4 percent (table 2).

Table 1-the ingredients used for preparing the PCR master mix

The E value of the overall sequences is 3.1x10^-9, and a bit score of 51.9. The percent identity of the overall sequences is 54.4 percent (table 2).