Introduction

It is a tale as old as time. Your friend, a free-loading, cat-loving, anti-veterinarian turns to you in their hour of need. Their cat is sick and needs a diagnosis stat. How can you possibly uncover the potentially fatal illness before it is too late? I will lay out a general framework for performing such a daunting task.

Methods

For the sake of brevity I will skip redundant analyses performed throughout the investigation to provide a streamlined pipeline for future analyses. Begin by using the suite of Blast+ (Zhang, J. 1997) tools to search for the possible proteins they code for. From the command line run blastx, searching in the non-redundant protein sequences database.

blastx -db nr -query <sequence\_file> -remote

For each sequence, screen the results for the sequences with the highest percent alignment and the highest e-value. Perform a brief literature review for the potential pathogens that may infect felines. You have officially provided a diagnosis. For fun, go to the accession number for the pathogen with the greatest probability of causing illness in the cat and download the fasta file for the coding regions of the pathogen’s genome. Create a query-able database and search for the protein, again, for fun. Now you are doubly sure the protein that you found is from the pathogen you discovered initially.

Results

My friend’s cat is sick with kitty COVID.

Discussion

While there may have been some redundant steps, the methods laid out above can be simplified to two simple steps of searching for potential proteins coded for by the mystery DNA samples and a quick follow-up of the literature.

References

Zhang, J. & Madden, T.L. (1997) "PowerBLAST: A new network BLAST application for interactive

or automated sequence analysis and annotation." Genome Res. 7:649-656.